

# Systematic screening for genes specifically expressed in the anterior neuroectoderm during early *Xenopus* development

NORIYUKI TAKAHASHI<sup>1,2</sup>, NAOKO TOCHIMOTO<sup>1</sup>, SHIN-YA OHMORI<sup>1,2,#</sup>, HIROSHI MAMADA<sup>1</sup>, MARI ITOH<sup>1,2</sup>, MASAKO INAMORI<sup>1,2</sup>, JUN SHINGA<sup>1,#</sup>, SHIN-ICHI OSADA<sup>1,2,#</sup> and MASANORI TAIRA<sup>1,2,\*</sup>

<sup>1</sup>Department of Biological Sciences, Graduate School of Science, University of Tokyo and <sup>2</sup>Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation, Tokyo, Japan

**ABSTRACT** A cDNA library derived from the anterior neuroectoderm (ANE) of *Xenopus* late-gastrula embryos was systematically screened to isolate novel developmental regulatory genes involved in early brain development. We isolated 1,706 5' expressed sequence tags (ESTs), which were subdivided into 1,383 clusters and categorized into 19 classes based on predicted functions according to their similarities to other known genes. Of these, 757 clusters that were considered possible novel regulatory genes or unknown genes were subjected to expression pattern analysis using whole-mount *in situ* hybridization. Genes from 69 clusters (9%) were expressed in the ANE region. Based on their expression patterns and predicted amino acid sequences, 25 genes were selected for further analysis as novel *Xenopus* genes expressed broadly or region-specifically in the ANE. Eighteen genes were expressed in postulated patterning centers in the neuroectoderm, including the anterior (four genes) and lateral (nine genes) neural ridges, the midbrain–hindbrain boundary region (one gene) and the midline region of the neural plate (two genes), whereas 13 genes were expressed in the eye anlagen. Therefore, early regionalization of the neuroectoderm appears to occur mainly in those neural patterning centers and the eye anlagen. We determined the entire coding regions of p54nrb, Semaphorin 6D and a novel gene designated scribble-related protein 1 (SCR1P). Interestingly, Semaphorin 6D is expressed in the mesoderm with a dorsoventral gradient, as well as in the ectoderm at the gastrula stage, implying a new role for this protein in development other than in axon guidance.

**KEY WORDS:** anterior neuroectoderm, expression pattern, scribble-related protein 1, p54nrb, Semaphorin 6D

## Introduction

Because the brain in vertebrates is an intricate organ, how the architecture of the brain is formed during embryogenesis is an intriguing question. To address this question, we focused on the molecular basis underlying an initial step of brain development. The formation of the central nervous system (CNS) including the brain is initiated upon neural induction and patterning by the gastrula organizer, called the “Spemann–Mangold organizer” in amphibians. The molecular nature of these organizer functions has been extensively investigated in amphibians using the African clawed frog, *Xenopus laevis* (Harland and Gerhart, 1997, De Robertis *et al.*, 2000). In *Xenopus*, neural induction appears to

occur at an early stage, from the blastula to gastrula stages, by bone morphogenetic protein (BMP) inhibitors that are secreted from the gastrula organizer (Munoz-Sanjuan and Brivanlou, 2002, De Robertis and Kuroda, 2004). Early patterning along the antero-posterior (AP) axis in the neuroectoderm is thought to be defined by posteriorizing factors, which are secreted from the trunk organizer (Mangold, 1933) and include retinoic acid (RA), fibroblast growth factors (FGFs) and Wnts (Taira *et al.*, 1997, Gamse and Sive, 2000). In the anterior neuroectoderm, which forms the forebrain and midbrain, posteriorizing signals are repressed by

*Abbreviations used in this paper:* AEM, anterior endomesoderm; ANE, anterior neuroectoderm; EST, expressed sequence tag; SCR1P, scribble-related protein.

\*Address correspondence to: Dr. Masanori Taira. Department of Biological Sciences, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Fax: +81-3-5841-4434. e-mail: m\_taira@biol.s.u-tokyo.ac.jp

# Present addresses: Dr. Osada. The 21st Century Center of Excellence (COE) Program, Akita University School of Medicine, Akita, Japan; Dr. Ohmori. Tsukuba Advanced Research Alliance (TARA) Center, University of Tsukuba, Ibaraki, Japan; Dr. Shinga. Research Center for Allergy and Immunology (RCAI), RIKEN Yokohama Institute, Yokohama City, Kanagawa, Japan.

**Accession numbers** for sequences cited in this paper: XSema6D4 (AB238227); Xp54nrb (AB238228); SCR1P (AB238229).

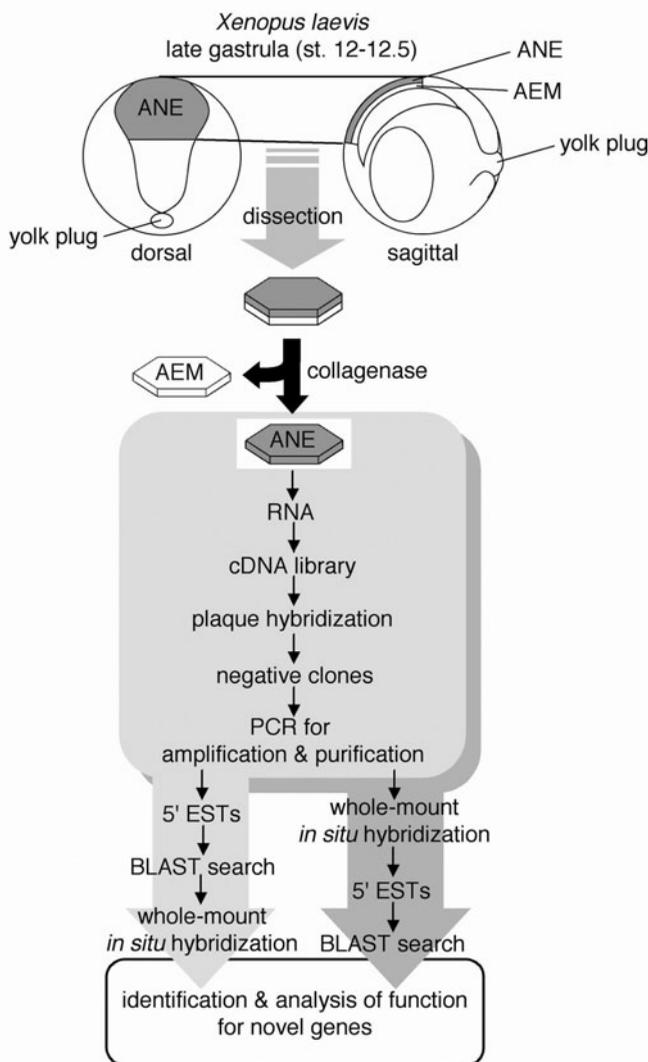
0214-6282/2005/\$25.00

© UBC Press  
Printed in Spain  
www.intjdevbiol.com

Wnt inhibitors secreted from the head organizer region (Niehrs, 1999) and by RA degradation enzyme expressed in the anterior region (Holleman *et al.*, 1998, Kudoh *et al.*, 2002).

The neuroectoderm subsequently forms the neural plate and then the neural tube. The neural tube is delimited into segmental AP structures such as the forebrain, midbrain, hindbrain and spinal cord and into dorsoventral (DV) structures such as the roof, alar and floor plates. These patterning processes of the neural plate and neural tube are apparently executed by so-called "patterning (or organizing) centers" that arise in the neuroectoderm and neural fold, such as the anterior neural ridge (ANR) and the midbrain–hindbrain boundary (MHB) as well as the lateral

edges and midline of the neural plate that will correspond to the roof plate and floor plate, respectively, in the neural tube. The molecular mechanisms that pattern the developing brain along the AP or DV axes have been extensively investigated in the chick and mouse (Lumsden and Krumlauf, 1996, Tanabe and Jessell, 1996, Puelles and Rubenstein, 2003). However, it is not yet known whether or how the gastrula organizer initiates the formation of organizing or patterning centers and how and when such patterning centers regionalize the neural plate and later the neural tube. To address these questions, *Xenopus laevis* is a useful model organism, because it is easy to manipulate the gastrula and neurula embryos and to examine gene expression patterns at various developmental stages from blastula to neurula by whole-mount *in situ* hybridization (WISH). Moreover, a detailed fate map of the neural plate in *Xenopus* shows that subdivided regions of the prospective brain region arise from the anterior neural plate and neural ridge (Eagleson and Harris, 1990). Therefore, as a first step, we used systematic screening to identify genes with expression patterns that suggest regionalization corresponding to this fate map. To increase cloning efficiency, we constructed regional cDNA libraries from the ANE and the underlying anterior endomesoderm (AEM) regions dissected from late-gastrula to early-neurula embryos. We have previously reported the systematic screening and analysis of the AEM cDNA library using expressed sequence tags (ESTs) and WISH, in which we successfully identified several AEM genes that are expressed differentially in the AEM (Shibata *et al.*, 2001). This implies that the AEM, or "head organizer region", is subdivided into regions with different properties. In this paper, we present the results of systematic screening of the ANE cDNA library to identify novel developmental regulatory genes with which to analyze the molecular mechanisms involved in early brain development. *XHR1*, one of the earliest genes expressed in the prospective MHB region, was isolated previously with this screen (Shinga *et al.*, 2001). Besides *XHR1*, we have isolated 25 candidate novel *Xenopus* developmental regulator genes, identified by their expression patterns and encoded protein structures. Importantly, most of these genes do not exhibit specific expression patterns corresponding to the detailed fate map, but are instead expressed in neural patterning centers, such as the neural ridge, MHB and the ventral and dorsal midlines of the neural tube and in the eye anlagen. These data imply that the early patterning of the presumptive brain region occurs mainly in the formation of signaling centers and eye anlagen at the late-gastrula to neural-plate stages.



**Fig. 1. Strategy for systematic screening of genes specifically expressed in the *Xenopus* ANE.** Screening for eliminating non-specific clones in the ANE region from an ANE library was performed along this scheme. Dissection of late-gastrula (stage 12–12.5) embryos was performed for isolation of ANE and AEM regions. Dissected pieces were treated with collagenase and separated into ANE and underlying AEM. A region-specific cDNA library was constructed from the ANE RNA. Further detailed processes are described in the text. AEM, anterior endomesoderm; ANE, anterior neuroectoderm.

## Results

### Systematic screening of an ANE cDNA library using ESTs

To obtain the *Xenopus* ANE with which to construct a regional cDNA library, the anterodorsal part of *Xenopus* late-gastrula embryos (stages 12–12.5) was dissected and separated into the ectodermal region and the underlying prechordal plate and anterior endodermal region by collagenase digestion (Fig. 1). Total RNA was extracted from about 600 pieces of the ANE and an ANE cDNA library was constructed using total RNA (Shinga *et al.*, 2001). To eliminate house-keeping genes and genes highly expressed in the trunk region at later stages, the library was negatively screened by plaque hybridization with mixed cDNA

probes prepared from total RNA of the trunk region of tailbud-stage embryos (stages 26–28), for which about 30% of plaques were positive. In the first round of screening, we randomly picked 433 hybridization-negative clones to screen for gene expression patterns using WISH. Some of the 433 clones were sequenced, in addition to 119 randomly picked clones, thereby constructing 215 informative 5' ESTs. In the second round of screening, we used additional plaque hybridization with cDNA probes to remove six genes that had repeatedly appeared when the AEM library was screened (Shibata *et al.*, 2001). Then we randomly picked and sequenced 1,824 hybridization-negative clones to obtain 1,491 informative 5' ESTs. Thus, we isolated 1,706 ANE 5' ESTs for our database, The *Xenopus laevis* Expression Pattern and Expressed Sequence Tags DataBase (XEXTDB).

The 1,706 ANE 5' ESTs were combined with 1,039 AEM 5' ESTs (Shibata *et al.*, 2001) and subjected to cluster analysis, which produced 2,179 total or 1,383 ANE 5'-EST-containing clusters (Fig. 2A). Of the 1,383 ANE clusters, 1,264 (91%) were ANE-specific and did not contain AEM clones. Of these 1,264 clusters, 1,124 (89%) were nonredundant (consisted of only one clone). These 1,124 clones were not necessarily unique clones,

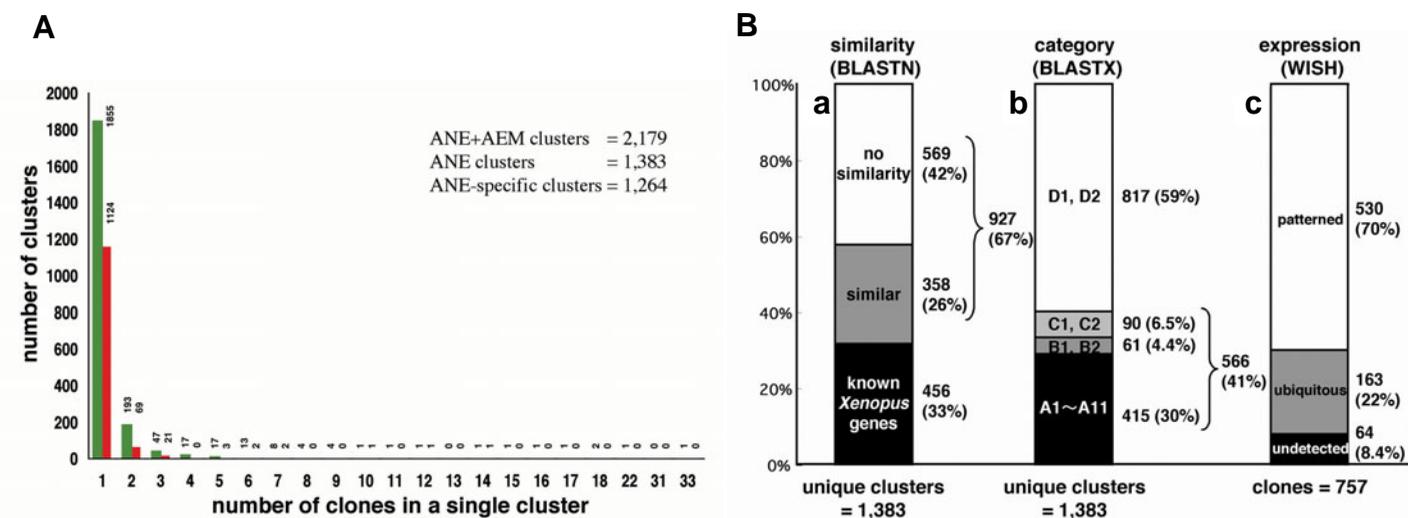
because the number of unique clusters will decrease when the entire lengths of the clones are sequenced. As shown in Fig. 2, however, the number of clusters decreased dramatically as the number of clones in each cluster increased, indicating that highly expressed genes had been efficiently removed from our ESTs by negative selection using plaque hybridization.

To classify the 1,383 clusters according to their similarity to other known genes, all ESTs in the clusters were subjected to similarity searches using the BLASTN and BLASTX programs against a public database (GenBank, release 139.0, December 2003). As shown in Fig. 2B, 456 (33%) of the 1,383 clusters correspond to known *Xenopus* genes as judged by the E-value (we empirically assumed that clones are identical to known genes when their E-values are either  $\leq e^{-75}$  for BLASTN or  $\leq e^{-50}$  for BLASTX). Five hundred and sixty-nine (42%) of 1,383 clusters showed significant similarity to known genes or proteins (their E-value are either  $\leq e^{-15}$  for BLASTN or  $\leq e^{-10}$  for BLASTX), whereas the remaining 817 clusters (59%) showed no similarity (their E-value are both  $> e^{-15}$  for BLASTN and  $> e^{-10}$  for BLASTX). Because 456 clusters (33%) showed strong similarity to known *Xenopus* genes, the remaining 927 clusters (67%) were

TABLE 1

## THE NUMBER OF DIFFERENT GENES EXPRESSED BY CLASS

Class	Total		Known <i>Xenopus</i> genes		
	No. of clusters	No. of clones	No. of clusters	No. of clones	
<b>(A) Functions that many kind of cell use</b>					
A1	Transportation and binding proteins for ions and other small molecules	15 (1.1%)	15	3 (0.2%)	3
A2	RNA processing, polymerizing, splicing and binding proteins, and enzymes (general transcription factors)	97 (7.0%)	119	50 (3.6%)	67
A3	Cell replication, histones, cyclins and allied kinases, DNA polymerases, topoisomerases, DNA modification	49 (3.5%)	79	38 (2.7%)	67
A4	Cytoskeleton and membrane proteins	31 (2.2%)	37	24 (1.7%)	29
A5	Protein synthesis co-factors, tRNA synthetases, ribosomal proteins	60 (4.3%)	84	37 (2.7%)	60
A6	Intermediary synthesis and catabolism enzymes	58 (4.2%)	68	25 (1.8%)	35
A7	Stress response, detoxification and cell defense proteins	15 (1.1%)	20	11 (0.8%)	15
A8	Protein degradation and processing, proteases	39 (2.8%)	49	22 (1.6%)	30
A9	Transportation and binding proteins for proteins and other macromolecules	25 (1.8%)	27	11 (0.8%)	13
A10	Apoptosis-related proteins	7 (0.5%)	8	2 (0.1%)	3
A11	Mitochondrial sequences	19 (1.4%)	77	13 (0.9%)	70
Subtotal		415 (30%)	583	236 (17%)	392
<b>(B) Cell-cell communication</b>					
B1	Signaling receptors, including cytokine and hormone receptors, and signaling ligands	21 (1.5%)	25	15 (1.1%)	19
B2	Intracellular signal transduction pathway molecules including kinases and signal intermediates	33 (2.4%)	40	23 (1.7%)	30
B3	Extracellular matrix proteins and cell adhesion	7 (0.5%)	8	4 (0.3%)	4
Subtotal		61 (4.4%)	73	42 (3.0%)	53
<b>(C) Transcription factors and other gene regulatory proteins</b>					
C1	Sequence-specific DNA-binding proteins	58 (4.2%)	72	39 (2.8%)	52
C2	Non-DNA binding proteins that perform positive or negative roles	11 (0.8%)	15	6 (0.4%)	9
C3	Chromatin proteins other than A3 with regulatory function	21 (1.5%)	44	17 (1.2%)	39
Subtotal		90 (6.5%)	131	62 (4.5%)	100
(A)+(B)+(C)		566 (41%)	787	340 (25%)	545
<b>(D) Miscellaneous</b>					
D1	Not enough information to classify	248 (18%)	303	116 (8.4%)	158
D2	Not significant similarities to known proteins	569 (41%)	616	0 (0%)	0
Subtotal		817 (59%)	919	116 (8.4%)	158
<b>Total</b>		<b>1383 (100%)</b>	<b>1706</b>	<b>456 (33%)</b>	<b>703</b>



**Fig. 2. Clustering of ESTs and classification of their sequences and expression patterns. (A)** Result of clustering. One thousand seven hundred six clones from the ANE cDNA library formed 1,383 clusters (ANE clusters). Of them, 1,264 (91%) are ANE-specific clusters, which do not mingle with previously reported AEM clones (Shibata et al., 2001). When ESTs from the ANE and AEM libraries were combined, total 2,963 ESTs formed 2,179 clusters. The number of clones in each single cluster was counted from total 2,179 clusters or 1,264 ANE-specific clusters. Ordinate, the number of cluster; abscissa, the number of clones in a single cluster; green bar, the number of clusters based on total clusters; red bar, the number of clusters based on ANE-specific clusters. **(B)** Classification of predicted proteins and expression patterns of ANE clusters. **(a)** ANE clusters (1,383) were classified according to sequence similarity to known genes or proteins of *Xenopus* or other species in public databases. Clusters were divided into three classes as indicated. **(b)** ANE clusters (1,383) were also classified according to functions of genes (see Table 1). **(c)** Out of 968 clusters which are functionally unclassified or possible developmental regulators (Classes B, C and D), 757 clones from 729 clusters were analyzed for expression patterns by WISH and classified into three categories as indicated. Values besides each column are given as the number of clusters and percentages in parentheses.

candidates for novel *Xenopus* genes.

To classify the 1,383 clusters according to the functions of the proteins, we used a previously described classification method (Lee et al., 1999). Class A proteins are associated with functions used in many kinds of cells, Class B proteins are associated with cell-cell communication and Class C proteins function as transcription factors and other gene regulatory proteins. As summarized in Table 1 and Fig. 2B, 566 clusters were categorized into Classes A (415 clusters), B (61 clusters) and C (90 clusters), which were further categorized into subclasses (Table 1). Among these, the most redundant genes in our ESTs were HMG-14 (high mobility group protein 14; the same as "high mobility group nucleosomal binding domain 1, Hmgn1") (0.8%; 14 of 1,706 total ESTs; in one cluster), the stage-specific elongation factor 42Sp48 protein (0.8%; 15/1,706; in two clusters), variant histone H2A.Z12 (0.7%; 12/1,706; in one cluster), high mobility box 3 (Hmgb3; 1.1%; 19/1,706; in 4 cluster) and mitochondrial DNA (4.2%; 72/1,706; in 14 clusters).

Besides the 566 clusters with significant similarities, 248 clusters (303 clones) were classified into Class D1 in which sequences are similar to those of reported proteins with unknown functions or to hypothetical proteins. Another 569 clusters (616 clones) were categorized into Class D2, which includes sequences with no significant similarities to other known proteins (Table 1). Thus, 817 clusters (D1 + D2) require characterization in the future.

#### Expression pattern analysis of ANE clones

The expression patterns of 757 genes (729 clusters) regarded as candidate novel genes were analyzed by WISH. As shown in

Fig. 2B, 530 genes (70%) were expressed in some restricted area. Of these, 308 genes (41%) were classified into "Patterned (early stage)", which signifies patterned expression in the neural plate before stage 19 (that is, before neural-tube closure), whereas 222 genes (29%) were classified into "Patterned (late stage)", which signifies patterned expression in the neural tube after stage 19 (Table 2). The remaining 163 (22%) and 64 (8.4%) genes were classified into "Ubiquitous" and "Undetected", respectively (Fig. 2B and Table 2). "Patterned (early stage)" genes were further classified into five groups according to their specific expression domains: (1) anterior neural plate/anterior neural ridge, (2) eye

TABLE 2

#### SUMMARY OF EXPRESSION PATTERN SCREENING

Category	No. of Clones (%)
Patterned	530 (70%)
early stage	308 (41%)
(1) anterior neural plate/anterior neural ridge	46 (6.1%)
(2) eye anlagen/neural ridge	23 (3.0%)
(3) neural plate/neural ridge	102 (13%)
(4) neural and other tissues	129 (17%)
(5) non-neural tissues	8 (1.1%)
late stage	222 (29%)
(i) CNS	217 (29%)
(ii) other tissues	5 (0.7%)
Other	227 (30%)
ubiquitous	163 (22%)
undetected*	64 (8.4%)
<b>Total</b>	<b>757 (100%)</b>

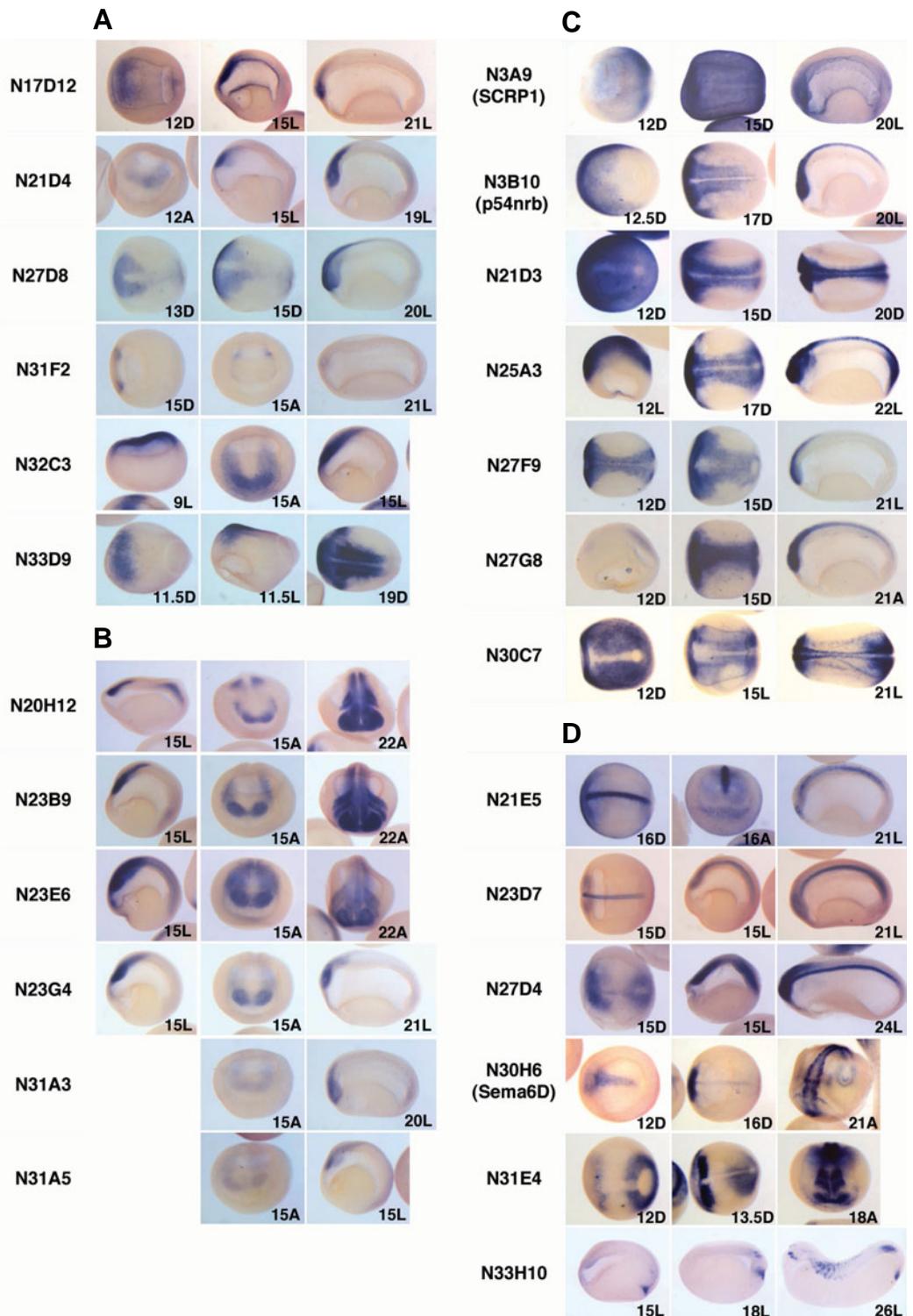
\*undetected, signals were not detected throughout the stages even after staining for 2 days.

anlagen/neural ridge, (3) neural plate/neural ridge, (4) neural and other tissues and (5) nonneural tissues (Table 2). "Patterned (late stage)" genes were simply classified into two categories: (i) CNS and (ii) Other tissues. Of the 757 genes examined for their expression patterns, a total of 69 genes in groups 1 and 2 were specifically expressed in the ANE region. These genes, which might be involved in early development of the brain, are what our screening initially aimed.

To judge the comprehensiveness of our screen, as many known *Xenopus* genes as possible that are region-specifically expressed in the ANE area were chosen and their occurrence among the ANE 5' ESTs was counted. Table 3 shows that 31 (50%) of 61 reported genes appeared among the ANE ESTs, implying that just 1,706 ESTs could include roughly half the genes that are specifically expressed in the ANE. This high frequency is probably due to the use of a regional cDNA library and negative selection by plaque hybridization.

**Expression patterns of 25 selected ANE genes**

We performed further sequence analysis of the 69 genes that are expressed in the neural ridge, the anterior neural plate, or both, some of which were identified as known genes. From the remaining uncharacterized genes, we selected the 12 genes with clear expression patterns. In addition, we also selected 7 genes which are mainly expressed in the neural plate (Table 2; groups 3 and 4) and 6 genes with unique expression patterns (group i and ii), as listed in Table 4 for further analysis (the name of the representative clone from each cluster is used for the name of the gene hereafter). As shown in Fig. 3A, *N17D12*, *N21D4*, *N27D8*, *N31F2*, *N32C3* and *N33D9* are specifically expressed in the ANE region at the late-gastrula/early-neurula to mid-neurula stages (stages 11.5–15). The ex-



**Fig. 3. Expression patterns of possible novel genes.** The left column shows cluster (gene) names and the right panels, show expression patterns visualized by WISH. Developmental stages according to Nieuwkoop and Faber (1967) (Nieuwkoop and Faber, 1967) and orientations of embryos are indicated on the bottom of each panel: D, dorsal; L, lateral; A, anterior. (A) Genes which were expressed in the anterior neural plate or anterior neural ridge, or both. (B) Genes which were expressed in eye anlagen together with or without the neural ridge. (C) Genes which were widely expressed in the neural plate or neural ridge, or both. (D) Genes which were expressed in the neural tissue as well as other tissues.

pression of *N32C3* was detected in the animal pole region at the late-blastula stage and is restricted to the anterior neural ridge at the mid-neurula stage (stage 15, Fig. 3A), whereas *N31F2* is probably expressed in the trigeminal placodes at the mid-neurula stage (Fig. 3A). *N20H12*, *N23B9*, *N23E6*, *N23G4*, *N31A3* and *N31A5* were specifically expressed in both the lateral neural ridge and the eye anlagen at the mid-neurula stage (stage 15, Fig 3B).

The genes shown in Fig. 3C are ubiquitously expressed in the neural plate. *N3A9* (*SCR1*; see below), *N3B10* (*p54nr1*; see

below), *N21D3*, *N25A3*, *N27F9*, *N27G8* and *N30C7* are largely expressed in the entire neural plate region at the mid-neurula stage (Fig. 3C) and these expression patterns are maintained in the tailbud stages.

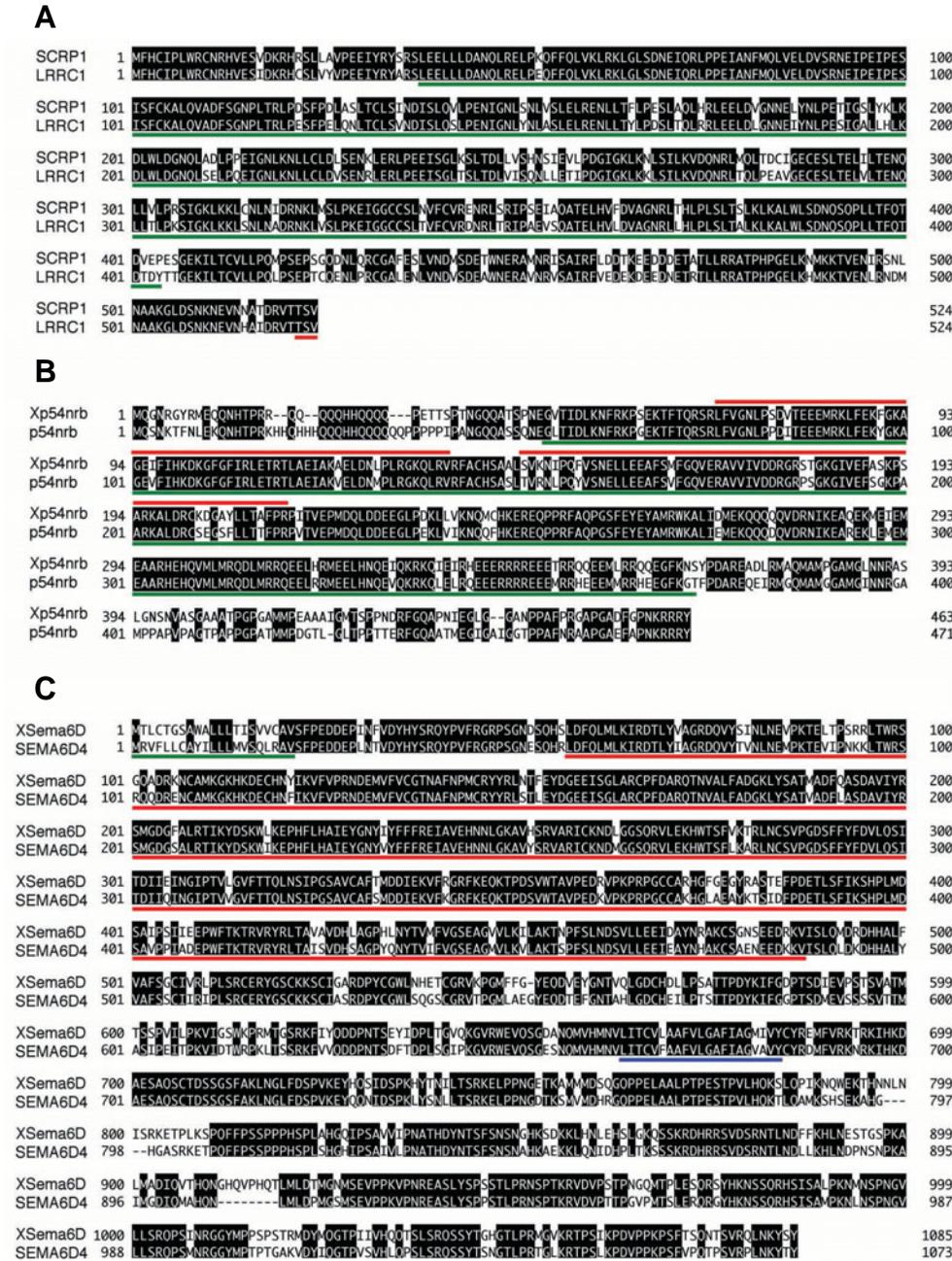
The genes that show unique expression patterns, but are not grouped in Fig. 3A through 3C are presented in Fig. 3D. *N21E5* and *N23D7* are expressed in the floor plate and the notochord, respectively, at the mid-neurula to tailbud stages (Fig. 3D). The expression of *N21E5* was also observed in the cement gland anlage and

is maintained in the cement gland at later stages (stage 28, data not shown). *N27D4* and *N31E4* are expressed in both the anterior neural plate and the region around the yolk plug at the mid-neurula stage (Fig. 3C). *N27D4* is also expressed in the notochord region at the tailbud stage (stage 24). *N30H6* (*Sema6D*; see below) is specifically expressed in the neural plate (Fig. 3D; see also Fig. 6) and this expression is localized in the ANE region at the mid-neurula stage (see Fig. 6 for more detail). *N33H10* is expressed specifically in the ventral blastopore at mid- to late-neurula stages and, at later stages, in an anterior tip that may correspond to the stomodeal-hypophyseal anlage, anterior somites and tailbud (Fig. 3D).

When comparing the expression patterns of 69 genes that are expressed in the ANE region, we noticed that 16 of them, including *XHR1* (Shinga et al., 2001), are expressed in the eye anlagen and in the so-called patterning centers, such as the anterior and lateral neural ridge (ANR, LNR) and the MHB, from the late-gastrula to early-tailbud stages, as summarized in Table 5. Whereas the remaining 53 genes are expressed in anterior neural region, those expressions are weak and/or are presented indistinct boundary of expression. However, we found no genes with expression patterns reflecting the fine fate map of early *Xenopus* neurula embryos (Eagleson and Harris, 1990, Eagleson et al., 1995), implying that detailed regionalization has not occurred at the early neural-plate stage.

**Analysis of unidentified EST clones by sequencing the entire inserts, cDNA cloning and 5' RACE**

We used library screening and 5' RACE to determine the entire coding sequences of genes that showed



**Fig. 4. Alignments of the predicted amino acid sequences of isolated clones and their orthologous proteins in other species. (A)** *SCR1* (*N3A9*) and human *LRR1* (*AAH03193*). Green line, LRR region; red line, PDZ domain binding sequence. **(B)** *Xenopus p54nr1* (*N3B10*) and the human ortholog. Green underline, DBHS domain; red overline, RRM1 and RRM2. **(C)** *Xenopus Sema6D4* (*N30H6*) and the human ortholog. Green line, signal sequence; red line, sema domain; blue line, transmembrane domain.

specific expression patterns or similarity to known developmental regulators, as described above. We isolated the full coding sequences for *N3A9* (a novel gene named *Scribble-related protein 1*, *SCR1*), *N3B10* (*p54nr1*) and *N30H6* (*Sema6D*) and analyzed their sequence similarities and their developmental expression patterns, as follows.

### **N3A9 (SCR1)**

The predicted amino acid sequence encoded by *N3A9* is homologous to the *Drosophila* protein Scribble, which is reportedly involved in cell polarity (Bilder and Perrimon, 2000). To analyze the entire coding region of the *N3A9* gene, we isolated four cDNA clones, N3A9-2 through -5, by screening a *Xenopus* head cDNA library. One of them, N3A9-3, was considered to include a full-length cDNA, with an open reading frame (ORF) encoding a protein of 524 amino acids. This predicted protein showed the highest homology to human LRRC1 (leucine-rich repeat-containing protein 1), which is a cDNA clone isolated by the National Institutes of Health Mammalian Gene Collection (NIH-MGC) Project (<http://image.llnl.gov/>). N3A9-3 and human LRRC1 show 85% identity in the leucine-rich repeats (LRRs) and 83% identity in overall and

have a PDZ-domain-binding sequence, Thr-Ser-Val, at the C terminus (Fig. 4A). Thus, *N3A9* is likely to be the *Xenopus* ortholog of human *LRRC1*. However, those proteins show homology only to the N-terminal LRRs of *Drosophila* Scribble and do not have PDZ domains in the C-terminal region. Therefore, we designated the *Xenopus* protein "scribble-related protein 1 (SCR1)".

We examined the developmental expression patterns of *SCR1* by northern blot analysis with RNA preparations from staged embryos, which showed the presence of 2.7-kb and 3.8-kb transcripts of *SCR1* (Fig. 5A). In unfertilized eggs, the expression of the 2.7-kb mRNA was weaker than that of the 3.8-kb mRNA. From the cleavage stage to the blastula stage, the levels of both mRNAs declined gradually, but increased at the early-gastrula stage (stage 10.5) as zygotic expression and became almost equal after the late-gastrula stage (stage 12.5). Thus, *SCR1* is expressed throughout all the developmental stages examined. The presence of two mRNAs may reflect splicing variants or a related gene.

### **N3B10 (p54nr1)**

Because *N3B10* is expressed in the anterior neural tissue of the early neurula, it was characterized by cDNA cloning and expres-

TABLE 3

#### SUMMARY OF ISOLATED KNOWN *XENOPUS* cDNAs WHICH WERE REPORTED TO SHOW REGIONALIZED EXPRESSION PATTERNS

No.	Genes	Gene products	Characterization	No. of times isolated	mRNA localization*
1	<i>eIF-4AIII</i>	DEAD-box protein	translation initiation factor (RNA helicase/DEAD-box helicase)	2	A
2	<i>EphA2</i>	receptor tyrosine kinase	Eph family receptor	2	A
3	<i>FGFR4b</i>	transmembrane protein	FGF receptor	1	A
4	<i>Otx1</i>	homeodomain protein	transcription factor	2	A
5	<i>Otx2</i>	homeodomain protein	transcription factor	1	A
6	<i>Otx5b</i>	homeodomain protein	transcription factor	3	A
7	<i>Anf1</i>	homeodomain protein	transcription factor	3	A
8	<i>ERB1</i>	Ets protein	Ets-type transcription factor	2	A
9	<i>lim5</i>	LIM-homeodomain protein	transcription factor	1	A
10	<i>MAN1</i>	LEM domain-containing protein	inner nuclear membrane protein	1	A
11	<i>pax2a</i>	paired/pairial homeo- domain protein	transcription factor	1	A
12	<i>Rx1A</i>	paired-like homeo/homeo- domain protein	transcription factor	2	A
13	<i>Six3</i>	six/homeodomain protein	transcription factor	1	A
14	<i>CIRP1</i>	RNP protein	RNA-binding protein (cold-inducible RNA-binding protein)	2	N
15	<i>FoxA4**</i>	winged-helix protein	transcription factor	2	N
16	<i>frizzled7</i>	transmembrane protein	Wnt receptor	1	N
17	<i>FRL-1</i>	secreted protein	signal-transducing protein	3	N
18	<i>Geminin H/L</i>	C-terminal coiled-coil domain protein	negative regulator for cell cycle	2	N
19	<i>hairy2b</i>	bHLH protein	transcriptional repressor	1	N
20	<i>LS13B/Sox11</i>	HMG domain protein	transcription factor	3	N
21	<i>nrp1A/B</i>	RNP protein	RNA-binding protein	2	N
22	<i>presenilin-a</i>	transmembrane protein	signal-transduction protein	1	N
23	<i>prickle</i>	LIM protein	intracellular Wnt signal antagonist	1	N
24	<i>Ran</i>	small GTPase	small GTP-binding protein (nuclear GTPase)	6	N
25	<i>RhoA</i>	small GTPase	signal-transduction protein	1	N
26	<i>Sox2</i>	HMG domain protein (Sry-related HMG factor)	transcriptional activator	1	N
27	<i>Sox3</i>	HMG domain protein	transcriptional activator	2	N
28	<i>Su(H)1</i>	basic helix-loop-helix protein	transcription factor	4	N
29	<i>TCF3b</i>	HMG-domain protein	transcription factor	2	N
30	<i>Zic-1 (Zic-related-1)</i>	zinc finger domain protein	transcription factor	1	N
31	<i>Zic3</i>	zinc finger domain protein	transcription factor	1	N
<b>Total</b>				<b>58</b>	

Genes which are reportedly expressed mainly in the neural tissue but did not appear in the ANE 5'-ESTs are: AP1, Ash3, Cpl1, CYP26, Delta1, eIF4a2, Emx2, En2, Fgf3, Fgf8, FoxB1, FoxD1, FoxD3, FoxG1, Gbx2, HoxB9, HoxD1, Iro1, Iro2, Iro3, Isthmin, Lim3, NCAM, NeuroD, Ngnr1, Notch1, SoxD, Synapobrevin, Tiarin, Zic5

\*, A: mainly anterior neural tissue; N: including neural tissue.

\*\* , FoxA4 corresponds to XFKH1 and XFD1

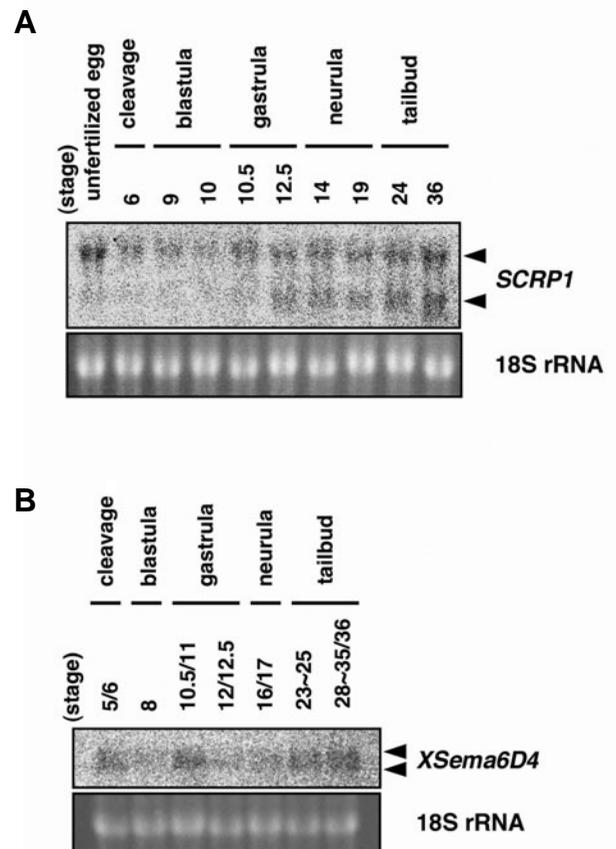
sion pattern analysis. Five cDNA clones, N3B10-2 through -6, were isolated by screening a *Xenopus* head cDNA library. Of these, N3A10-2 contained a long ORF, encoding a predicted protein of 463 amino acids (Fig. 4B). A database search revealed that the predicted product of the *N3B10* gene has 76% identity to human 54-kD nuclear RNA-binding protein (p54nrb) (Dong *et al.*, 1993) and is therefore referred to as Xp54nrb. RNA recognition motifs (RRMs) 1 and 2 occur in the N-terminal regions of both proteins and share 91% and 80% identities, respectively. Besides RRM1 and RRM2, there is a DBHS (*Drosophila* behavior/human splicing) domain, which is based on the consensus sequence found in p54nrb, human splicing factor PSF and *Drosophila* puff-specific protein BJ6 (Dong *et al.*, 1993). BJ6 is one of three products encoded by the alternatively spliced no-on-transient A gene (nonA) (von Besser *et al.*, 1990).

### N30H6 (Sema6D)

The deduced partial amino acid sequence of N30H6 has homology to Semaphorin 6 family members, which are thought to be involved in axon guidance (Pasterkamp and Kolodkin, 2003). However, the expression of *N30H6* in the anterior neural tissue and notochord (Fig. 3; see below) implies another role for Semaphorin in early development, which led us to analyze *N30H6* in detail. cDNA cloning of *N30H6* was performed using a gastrula cDNA library and the 5' end of the cDNA was isolated by 5' RACE. One of the seven cDNA clones isolated, N30H6-14, was considered to include the entire coding region. The deduced amino acid sequence of N30H6-14 showed high homology to a new member of the Semaphorin 6 family, referred to as Semaphorin 6D (Qu *et al.*, 2002). Whereas Semaphorin 6D has four types of splicing variants, SEMA6D1–4, which differ in the C-terminal region, *N30H6-14* encodes a protein corresponding to SEMA6D4, which is therefore designated XSema6D4 (Fig. 4C).

Northern blot analysis revealed that two *XSema6D* mRNAs of about 6.2 and 5.8 kb are present at the cleavage stage (stages 5/6), the levels of which decline at the blastula stage (stage 8), increase temporarily at the early- to mid-gastrula stages (stages 10.5/11) and are elevated again at later stages (tailbud stages 23–25) (Fig. 5B). These results are consistent with mRNA profiles analyzed by WISH, as described below.

At the late blastula stage (stage 9), *XSema6D* expression was detected in the animal hemisphere (Fig. 6A). Examination of sections indicated that this expression was localized mainly to the blastocoel roof (Fig. 6B). During gastrulation (stage 11.5), transcripts of *XSema6D* were also detected in the dorsal mesoderm, a region known as the Spemann organizer (Fig. 6C). In sagittal sections of embryos at this stage, *XSema6D* expression was seen in the organizer region and the sensorial layer of the neuroectoderm (Fig. 6D). At the mid-neurula (stage 15), *XSema6D* expression was mainly detected in the notochord and anterior neural plate (Fig. 6E). Sagittal hemisectioned embryos at stage 15 showed that its expression is seen in the sensorial layer of the anterior neuroectoderm and in the notochord (Fig. 6F). At the late neurula (stage 17), its expression was also detected in the presumptive otic vesicle (indicated by arrow in Fig. 6G). At tailbud stages, its expression was observed in the dorsal region of the otic vesicles, the trigeminal ganglia and the brain (Fig. 6H) and later in the head region, including the brain, branchial arches and otic vesicle and in the spinal cord, somites and pronephros (Fig. 6I).



**Fig. 5. Temporal expression profiles of *Xenopus* SCR1 and XSema6D4 by Northern blot analysis. (A)** SCR1. The upper panel shows 3.8- and 2.7-kb transcripts of SCR1 (arrow heads). **(B)** XSema6D4. The upper panel shows about 6.2 and 5.8-kb transcripts of XSema6D4 (arrow heads). Numbers, Nieuwkoop and Faber developmental stages; 18S rRNA stained with ethidium bromide, loading control.

Thus, the later expression of *XSema6D* might be associated with axon guidance in those regions, whereas its early expression in the ectoderm and mesoderm occurs in a region-specific manner. This implies that XSema6D has some other roles in development, such as in regionalization and morphogenetic movements through cell–cell communication.

## Discussion

To understand the molecular mechanisms of early neural patterning, we collected data from the systematic screening of an ANE cDNA library derived from *Xenopus* late-gastrula embryos. We isolated 817 possible novel genes (D1 plus D2 in Table 1) and identified 530 genes with localized expression patterns in the neural plate (Fig. 2B and Table 2). Based on the expression patterns of 69 genes that are expressed in the ANE region, we found that 12 of these genes, including *XHR1* (Shinga *et al.*, 2001), are expressed in the so-called patterning centers for the neuroectoderm, such as the ANR, LNR and MHB regions (Table 5). We focused on 25 possible developmental regulatory genes, including those 12 genes except for *XHR1*, based on their homology to known genes or proteins and on their developmental

expression patterns. Three of those genes were identified as novel *Xenopus* genes, *SCRP1*, *Xp54nrb* and *XSema6D*, by analysis of their entire coding regions.

### Regional specification of the neuroectoderm compared with the fate map of the neural plate

We initially expected that some genes might be expressed in restricted areas of the presumptive brain region, reflecting the fine fate map of early *Xenopus* neurula embryos (Eagleson and Harris, 1990, Eagleson *et al.*, 1995). However, we found no such expression patterns in the genes we isolated in this study, even though our screening probably included about half the genes expressed in the anterior neural plate region in the late-gastrula to mid-neurula stages (Table 3). Interestingly, a certain number of genes were expressed in the neural patterning centers: four genes in the ANR, nine genes in the LNR, one gene already reported as *XHR1* in the MHB region and two genes in the MLT, as well as two genes in the notochord. Furthermore, 13 genes were expressed in the eye anlagen (Table 5). Therefore, early regionalization of the anterior neuroectoderm appears to occur in the neural patterning centers and the eye anlagen, as well as in the AP subregions delineated by *Six3*, *Pax2*, *Rx1*, *XANF1/Hesx/Rpx*, *BF1/FoxG1*, *BF2/FoxD1* (in the forebrain) and so on (Bally-Cuif and Hammerschmidt, 2003). However, early regionalization of the posterior neuroectoderm appears to occur in mediolateral subregions, as described below. We speculate that subdivision of the neural plate into respective brain regions is not yet established during the early- to mid-neurula stages, unlike that in the posterior region.

Recently, it has been noted that three stripes of proneuronal

regions, which correspond to future motor, inter and sensory neurons, are defined by nonneuronal regions including the interstripes in the posterior neural plate, as well as by the MLT, the presumptive MHB region and the ANP. The interstripes and MLT differentially express the *Zic2* and *Xiro3* genes, which are considered to be pre-patterning genes (Bally-Cuif and Hammerschmidt, 2003), from the mid-gastrula stage, whereas the presumptive MHB region expresses *XHR1* (*her5* in zebrafish) at the early-gastrula stage. It has been suggested that *XANF1* and *Xrx1*, which are expressed in the ANP, prevent neurogenesis and promote cell division. Therefore, it is likely that the neural plate is initially divided into two regions: the proneuronal region and the nonneuronal and cell-dividing region (Bally-Cuif and Hammerschmidt, 2003). More detailed regionalization of the neural plate or neural tube corresponding to the reported fate map (Eagleson and Harris, 1990) is likely to be induced by the neural patterning centers and may also be defined by interactions between the roughly defined pre-patterned regions.

### Efficacy of regional cDNA library screening for isolating novel developmental regulatory genes

In general, developmental regulatory genes start to be expressed in a specific region where new differentiation and morphogenesis will occur during development. This expectation allowed us to rationalize our strategy to isolate genes with expression patterns localized in a restricted area. On the other hand, a certain number of developmental regulatory genes that encode intracellular components of signal transduction pathways are broadly or ubiquitously distributed in embryos, compared with ligands and tissue-specific transcription factors. Previously, sev-

TABLE 4  
SUMMARY OF INFORMATION FOR ISOLATED cDNA CLONES

Clone name	No. of times isolated	Insert (kbp)	Homology search method	Results of homology search					
				Identity	Score	E. value			
<i>N17D12</i>	1	2,7	BLASTX	O14524	<i>Homo sapiens</i>	Hypothetical protein KIAA0286 (HA6800)	140	6E-33	
<i>N21D4</i>	1	2,4	-	-	-	no-hit	-	-	
<i>N27D8</i>	1	0,3	-	-	-	no-hit	-	-	
<i>N31F2</i>	1	1,4	-	-	-	no-hit	-	-	
<i>N32C3</i>	1	0,8	BLASTX	A48830	<i>Mus musculus</i>	probable transcription regulator NT fin12	66	1E-11	
<i>N33D9</i>	1	1,3	BLASTN	BC044687	<i>Xenopus laevis</i>	similar to secreted frizzled-related sequence protein 2	404	1E-110	
<i>N20H12</i>	1	0,8	BLASTN	BC044048	<i>Xenopus laevis</i>	cDNA clone MGC:53515 IMAGE:5572607	878	0	
<i>N23B9</i>	1	1	BLASTX	Q9BYC8	<i>Homo sapiens</i>	60S ribosomal protein L32, mitochondrial precursor	169	1E-41	
<i>N23E6</i>	1	1,8	-	-	-	no-hit	-	-	
<i>N23G4</i>	1	0,8	-	-	-	no-hit	-	-	
<i>N31A3</i>	1	0,8	BLASTN	BC044102	<i>Xenopus laevis</i>	similar to collaborates/cooperates with ARF protein	567	1E-159	
<i>N31A5</i>	1	0,6	-	-	-	no-hit	-	-	
<i>N3A9/SCRP1</i>	2	2,2	BLASTX	AAH03193	<i>Homo sapiens</i>	Similar to scribbled	857	0	
<i>N3B10/Xp54nrb</i>	2	2,1	BLASTX	NP_031389	<i>Homo sapiens</i>	Nuclear RNA-binding protein, 54-kD	598	1E-169	
<i>N21D3</i>	1	1,1	-	-	-	no-hit	-	-	
<i>N25A3</i>	1	0,9	-	-	-	no-hit	-	-	
<i>N27D4</i>	1	0,8	-	-	-	no-hit	-	-	
<i>N27F9</i>	1	1,2	BLASTN	BC041501	<i>Xenopus laevis</i>	similar to programmed cell death 10	866	0	
<i>N27G8</i>	1	0,7	-	-	-	no-hit	-	-	
<i>N30C7</i>	1	1,6	-	-	-	no-hit	-	-	
<i>N21E5</i>	1	1,7	-	-	-	no-hit	-	-	
<i>N23D7</i>	1	1,6	-	-	-	no-hit	-	-	
<i>N30H6/XSema6D</i>	1	5,6	BLASTX	AAM69452	<i>Homo sapiens</i>	semaphorin 6D isoform 4	1786	0	
<i>N31E4</i>	1	0,4	-	-	-	no-hit	-	-	
<i>N33H10</i>	1	3,0	-	-	-	no-hit	-	-	

eral studies using systematic analyses of ESTs and expression patterns successfully isolated candidate genes of developmental regulators (Gawantka *et al.*, 1998, Lee *et al.*, 1999, Kudoh *et al.*, 2001, Satou *et al.*, 2001, Shibata *et al.*, 2001, Satou *et al.*, 2002, Pollet *et al.*, 2005).

In terms of the isolation efficiency of genes showing patterned expression, we isolated 530 (70%) of 757 genes that were categorized as having "patterned expression" (Table 2). This is higher than the 24% to 26% efficiency achieved with cDNA libraries from whole neurulae (stage 13) and LiCl-treated embryos (stage 13) (Gawantka *et al.*, 1998, Pollet *et al.*, 2005). This difference is probably due to our screening strategy in which a regional cDNA library and negative selection by plaque hybridization were used to remove clones that are highly expressed in the trunk region of tailbud-stage embryos. In our previous screening of an AEM cDNA library (Shibata *et al.*, 2001), using basically the same strategy as in this study, genes with patterned expression were isolated with 49% efficiency. This difference (70% versus 49%) may be attributable, at least partly, to the negative selection in the ANE library of the six genes that frequently appeared in the AEM ESTs (see Materials and Methods) and also to tissue differences between the ANE and AEM.

We isolated 69 ANE-specific genes (9.1%) from the 757 genes examined by WISH, that are expressed in either the anterior neural plate (46 genes) or the anterior neural ridge (23 genes) (Table 2). Twenty-five of them are presented in Fig. 3. In our previous screening of the AEM cDNA library, nine (6%) of 198 clusters analyzed showed expression patterns in organizer or organizer-derived tissues (Shibata *et al.*, 2001). We expected ANE-specific genes to be concentrated in the ANE cDNA library by roughly 20-fold, because the average yield of RNA from the ANE region is about 5% of that from whole embryos. As shown in Table 3, 50% of known genes with expression mainly restricted to the ANE were included in our ESTs, implying that more than half

the novel ANE-specific genes were recovered in our relatively small-scale screening of 1,706 clones. This corresponds roughly to 40,000 (1,706 multiplied by 20) informative clones from a whole-embryo library.

#### Gene profiles of ESTs from the ANE library

In our ANE ESTs, clusters that were classified into category A6 (intermediary synthesis and catabolism enzymes) and B2 (intracellular signal transduction pathway molecules including kinases and signal intermediates) were less frequent in our screening (4.5% and 2.4%, respectively, Table 1) than in the analysis of ESTs from *Ciona intestinalis*. In that analysis, A6 and B2 comprised 10.2% and 5.6%, respectively, of the 1,066 ESTs from cleavage-stage embryos, or 9.6% and 5.0%, respectively, of the 1,213 ESTs from tailbud-stage embryos (Satou *et al.*, 2002). These differences may reflect the expectation that most genes that belong to categories A6 and B2 are expressed ubiquitously, as is exemplified by the intracellular components of Wnt, FGF, BMP and activin/nodal signaling in early embryogenesis (Schohl and Fagotto, 2002). This means that ANE ESTs might consist of less redundant and more variable genes compared with the *Ciona* ESTs, because our ESTs were derived from a regional cDNA library and were depleted of housekeeping and ubiquitous genes by negative screening.

In our previous screening of a *Xenopus* AEM cDNA library, the most redundant clones encoded Xcirp (cold-inducible RNA binding protein; 3.8%; 39/1,039 clones), glutamine synthetase (1.3%; 13/1,039) and PR264 (1.1%; 11/1,039) (Shibata *et al.*, 2001). In this study, because those genes were removed from ANE clones at the beginning of the screening process, the most redundant clones encoded HMG-14, 42Sp48 protein, variant histone H2A.Z12, Hmgb3 or have mitochondrial DNA sequences. These results may provide evidence that cell division is more pronounced in the neural plate than in nonneural ectodermal and axial mesodermal cells that stop dividing after involution (Saka and Smith, 2001).

TABLE 5

#### EXPRESSION OF EYE ANLAGEN AND KNOWN SIGNALING CENTERS

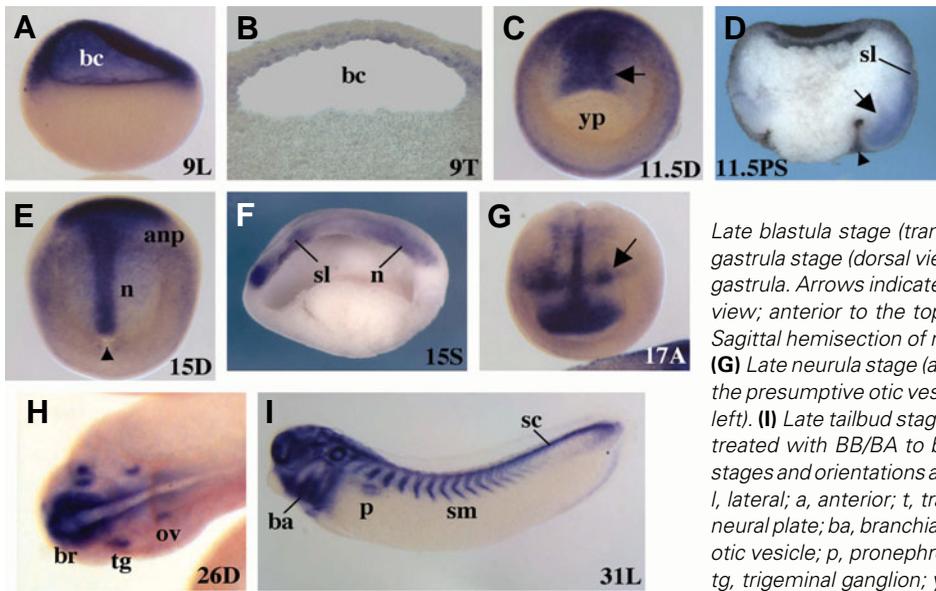
Clone name	EA	ANR	LNR	MHB	MLT	TB	Reference
N17D12	■						Fig. 3A
N21D4	■	■	■				
N27D8	■	■	■				
N31F2	■	■	■				
N32C3	■		■			■	Fig. 3B
N33D9	■		■			■	
N20H12	■		■			■	
N23B9	■		■			■	
N23E6	■		■			■	
N23G4	■		■			■	Fig. 3D
N31A3	■		■			■	
N31A5	■		■			■	
N21E5					■	■	
N23D7	■						Shinga <i>et al.</i> , 2001
N27D4	■				■	■	
N30H6/XSema6D	■		■		■	■	
N33H10						■	
N9G8/XHR1				■			
<b>Total numbers</b>	<b>13</b>	<b>4</b>	<b>9</b>	<b>1</b>	<b>2</b>	<b>5</b>	

EA, eye anlage; ANR, anterior neural ridge; LNR, lateral neural ridge; MHB, midbrain-hindbrain boundary region; MLT, midline tissues; TB, tail bud.

#### Features of novel *Xenopus* genes, SCR1, Xp54nrb and XSema6D

We have identified three novel *Xenopus* genes, SCR1, Xp54nrb and XSema6D, the structures and expression patterns of which are interesting in terms of neural development. SCR1 is a new gene related to *Drosophila scribble* (Bildler and Perrimon, 2000) and mouse *scribble/circletail* (*Scrb1*) (Murdoch *et al.*, 2003), members of the LAP family. *Drosophila scribble* is reportedly involved in apical-basal cell polarity, whereas mouse *Scrb1* was identified as the gene responsible for the *circletail* mutant, which shows defects in neural-tube closure, referred to as craniorachischisis. Because SCR1 has 16 LRRs in the N-terminal region but no PDZ domain in the C-terminal region, unlike Scribble and Scrb1, it may play a role as a modulator of Scribble in neural-tube closure, apical-basal cell polarity, or both.

Human p54nrb (rat nonO) was isolated as a protein with two tandemly repeated RNP-type RRM (Dong *et al.*, 1993, Yang *et al.*, 1993). It has



**Fig. 6. WISH analysis of *XSema6D* (N30H6) during early development. (A)** Late blastula stage (lateral view; animal pole to the top). **(B)**

Late blastula stage (transversal section; animal pole to the top). **(C)** Mid gastrula stage (dorsal view; anterior to the top). **(D)** Sagittal section of mid gastrula. Arrows indicate the organizer (C,D). **(E)** Mid neurula stage (dorsal view; anterior to the top). Arrowheads indicate the blastopore (D,E). **(F)** Sagittal hemisection of mid neurula (anterior view; dorsal to the top). An arrow indicates the presumptive otic vesicle. **(H)** Tailbud stage (dorsal view; anterior to the left). **(I)** Late tailbud stage (lateral view; anterior to the left). Embryos were treated with BB/BA to become transparent (A,D,E,G-I). Developmental stages and orientations are indicated on the bottom of each panel: d, dorsal; l, lateral; a, anterior; t, transversal; s, sagittal. Abbreviations: anp, anterior neural plate; ba, branchial arches; bc, blastocoel; br, brain; n, notochord; ov, otic vesicle; p, pronephros; sc, spinal cord; sl, sensorial layer; sm, somites; tg, trigeminal ganglion; yp, yolk plug.

been suggested that p54nrb and the splicing factor, PSF, form a heterodimer that stimulates topoisomerase I activity (Straub *et al.*, 1998, Straub *et al.*, 2000). A recent study has shown that a complex containing p54nrb, PSF and the nuclear matrix structural protein matrin 3 (mat3) binds specifically to inosine-containing RNAs and is involved in nuclear retention of unwanted hyperedited RNAs (Zhang and Carmichael, 2001). Because p54nrb is specifically expressed in neural tissue (Fig. 3; right panel), it is reasonable to speculate that Xp54nrb is involved in neural cell differentiation or patterning by regulating the expression of edited transcripts.

Semaphorins form a large family of axon-guidance molecules and are categorized into eight classes (1–7 and V) according to the characteristics of their structures. Insect *Sema1a* and mammalian *Sema3A* and *Sema4D* have been extensively investigated and several functions of semaphorins have been implicated in axon pathfinding (Pasterkamp and Kolodkin, 2003). Recently, a new paralogous gene of the *Sema6* genes, *semaphorin 6D* (*SEMA6D*), was isolated in humans. *SEMA6D* has five isoforms, *SEMA6D1–4* and *SEMA6Ds* (short isoform) derived from alternative splicing. *SEMA6D* is expressed abundantly in various tissues and acts in inducing growth-cone collapse and inhibiting axonal extension (Qu *et al.*, 2002). Whereas human *SEMA6D* seems to be expressed ubiquitously, as assayed by northern blotting of various adult tissues (Qu *et al.*, 2002), *XSema6D* is expressed in the animal hemisphere of the blastula and in the organizer and neuroectoderm of the gastrula (Fig. 6). Therefore, *XSema6D* might have a function distinct from axon guidance in early embryogenesis.

### Conclusions and perspective

In this study, we used systematic screening of an ANE cDNA library by constructing a relational database, XEXTDB, in which ESTs and WISH data are connected to each other, to identify many candidate novel developmental regulatory genes that are expressed in restricted patterns in the neural plate or neural ridge. Because the number of ESTs from various organisms is increas-

ing dramatically day by day, ESTs are very useful in predicting the structures of full-length cDNAs by the generation of contigs. They can also be used to predict novel proteins by comparing the ORFs between distantly divergent organisms. Using these approaches, we identified the novel gene *SCRPF1* and described for the first time the developmental expression patterns of *SCRPF1*, *Xp54nrb* and *XSema6D*. Furthermore, when our 757 ESTs that are linked to developmental expression patterns in XEXTDB are available to the public (<http://xextdb.nibb.ac.jp/>) and incorporated into contigs with other publicly available ESTs, the expression patterns of any genes corresponding to those contigs will be immediately accessible in the XEXTDB. The structures of predicted proteins and their developmental expression patterns will facilitate functional analysis of novel genes during the development of various tissues and organs, including the brain. All *Xenopus* transcripts will be identified in the future by repeated systematic screening and expression pattern analysis including our studies (this paper and the previous one (Shibata *et al.*, 2001)) as well as studies by other groups (Gawantka *et al.*, 1998, Baldessari *et al.*, 2005, Pollet *et al.*, 2005, Shin *et al.*, 2005, Voigt *et al.*, 2005) and substantial accumulation of ESTs and expression patterns will facilitate advances in functional genomics.

### Materials and Methods

#### Manipulation of *Xenopus* embryos

Unfertilized wild-type or albino *Xenopus* embryos were ovulated by injection of gonadotropin (Teikokuzoki Co.). Embryos were obtained by artificial fertilization with macerated testis, dejellied in 2% L-cysteine-HCl (pH 7.9) and reared in 0.1X Steinberg's solution at 14–21°C. Embryos were staged according to the criteria of Nieuwkoop and Faber (Nieuwkoop and Faber, 1967).

#### Screening of a regional cDNA library by plaque hybridization and making ESTs

A unidirectional cDNA library was constructed from the anterior neuroectoderm of stage 12–12.5 *Xenopus* embryos as previously reported (Shinga *et al.*, 2001). Plaques were transferred onto nylon mem-

brane filters (Nytran, Schleicher and Schuell) and filters were hybridized with [ $\alpha$ - $^{32}$ P]dCTP labeled cDNA prepared with total RNA from the trunk region of tailbud embryos (stages 26-28), then washed with 2X SSPE/0.1% SDS at 65°C and exposed to a X-ray film (BioMax MR, Kodak). The same filters were also probed with mixed 6 different cDNAs encoding *Xc1p*, adenosine triphosphatase sodium-potassium pump  $\alpha$ 1 subunit, glutamine synthetase, HMG-X protein, ribosomal protein L7a and *prp1* (acidic ribosomal protein 1) to remove genes which had appeared frequently in AEM cDNA library screening (Shibata *et al.*, 2001). Inserts of cDNA clones were amplified by PCR using the SK (5'-TCTAGAACTAGTGGATC-3') and T7 (5'-CCGGATCCTAATACGACTCACTATA-3') primers in 96-well plates. PCR products were purified using MultiScreen 96-Well Filtration Plates (MILLIPORE) according to the manufacturer's protocol. Briefly, PCR products were mixed with saturated KI solution in advance, absorbed to MultiScreen FB Plates (MILLIPORE) by centrifugation, washed with 70% ethanol/30% 0.1 M NaCl and extracted with TE buffer. To remove remaining KI, extracts were passed through MultiScreen HV Plates filling with Sephadex G-50 Superfine (Pharmacia). Purified PCR products were sequenced from the 5' end with the T3 primer (5'-AATTAACCCTCACTAAA-3') using an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems) for making ESTs. The average of informative EST data was about 450 nucleotides long.

#### Sequencing and analysis of plasmid cDNA clones

Phage cDNA clones were converted into plasmids by *in vivo* excision of the pBluescript SK(-) phagemid from the  $\lambda$ ZAP II Express vector according to the ZAP-OUT method (Stratagene). The missing 5' regions of cDNA clones were obtained by 5' RACE (rapid amplification of cDNA ends) with a 5'-Full RACE Core Set (Takara) and PCR products were inserted into multiple cloning site of the pCR2.1 vector with a TA Cloning Kit (Invitrogen). To obtain longer cDNA clones, cDNA libraries of stage 11-12 *Xenopus* embryos (a gift from B. Blumberg) and the head region of stages 28-30 *Xenopus* embryos (a gift from R. Harland) were used. To determine entire cDNA sequences, deleted plasmid constructs were made by means of restriction enzyme treatment, or plasmid constructs were randomly inserted by *in vitro* transposon with a sequencing primer site for GeneJumper sequencing primers A and B using the GeneJumper Primer Insertion Kit for Sequencing (Invitrogen). Plasmid constructs were sequenced using ABI PRISM 310 or 3100 Genetic Analyzer (PE Applied Biosystems), or Li-COR LIC-4200L-2 (Aloka) with T3/T7 primers for pBluescript SK (-) plasmid clones, M13 (5'-CAGGAAACAGCTATGAC-3')/T7 primers for pCR2.1 plasmid clones, or GeneJumper Sequencing Primers A (5'-ATCAGCGCCGCGATCC-3') and B (5'-TTATTCGGTGCAAAAGGATCC-3') for GeneJumper clones. Alignments of deduced amino acid sequences were performed using CLASTALW (<http://www.ddbj.nig.ac.jp/>) and motif searches were performed using SMART 3.3 (<http://smart.embl-heidelberg.de/>).

#### Northern hybridization

Total RNA from embryos of different developmental stages was extracted by the acid guanidium isothiocyanate-phenol-chloroform method (Chomczynski and Sacchi, 1987). Total RNA (10  $\mu$ g) was electrophoresed on a 1% agarose/formaldehyde gel, blotted onto a nylon membrane (Nytran, Schleicher and Schell) and cross-linked by UV irradiation. Blot hybridization was performed with a [ $\alpha$ - $^{32}$ P]dCTP labeled cDNA probe, washed with 2X SSPE/0.1% SDS at 65°C, exposed to a BAS-IP MN 2041 film (Fuji Film) and analyzed with BAS-2500 (Fuji Film).

#### Whole-mount in situ hybridization (WISH) and sectioning

Embryos were fixed in MEMFA (0.1 M MOPS, pH 7.4, 2 mM EGTA, 1 mM MgSO<sub>4</sub>, 3.7% formaldehyde). DIG-labeled anti-sense RNA was synthesized with T7 RNA polymerase (Boehringer Mannheim) and with purified PCR products or linearized pBluescript SK(-) plasmid clones possessing the T7 promoter at the 3' end as templates. WISH was performed manually or with an automated system, AIH101B (Aloka),

according to Harland (Harland, 1991). BM Purple (Boehringer Mannheim) was used for chromogenic reactions (24-48 hours). When necessary, stained embryos were cleared with benzyl benzoate/benzyl alcohol (BB/BA; 2:1). Some stained embryos were embedded in paraffin and sectioned at 15  $\mu$ m.

#### Construction of a relational database, XEXTDB

XEXTDB (<http://xextdb.nibb.ac.jp/8080>) was modeled after Nematode Expression Pattern Database (NEXTDB; constructed by Y. Kohara and T. Shin-i, National Institute of Genetics; <http://nematode.lab.nig.ac.jp/>). Partial 5' sequences of ANE clones were registered in XEXTDB as ESTs and were subjected to clustering and BLAST searches using XEXTDB against GenBank database installed from NCBI (National Center for Biotechnology Information), which includes gbinv (invertebrate sequences) and gbvrt (other vertebrate sequences), excluding human, primate, rodent and other mammalian sequences), PIR and SWISSPROT protein databases installed from DDBJ (DNA Data Bank of Japan) and EBI (European Bioinformatics Institute), respectively. Expression profiles were also registered in XEXTDB with linkage to the corresponding ESTs.

#### Acknowledgements

We thank Yuji Kohara for allowing us to convert database system of Nematode Expression Pattern Database (NEXTDB) into Xenopus Expression Pattern Database (XEXTDB) and Tadasu Shir-i, Masaki Humoto, Sigeru Saito (Infocom Co.) for construction and maintenance of XEXTDB. We also thank Richard M. Harland for a Xenopus stages 28-30 head cDNA library, Bruce Blumberg for a Xenopus st. 11-12 cDNA library, Sumiko Taira for sectioning and Mikihiro Shibata for discussions. This work was supported in part by Grant-in-Aid for Science Research from the Ministry of Education, Science, Sports and Culture of Japan and by Toray Science Foundation, Japan.

#### References

- BALDESSARI, D., SHIN, Y., KREBS, O., KONIG, R., KOIDE, T., VINAYAGAM, A., FENGER, U., MOCHII, M., TERASAKA, C., KITAYAMA, A. *et al.* (2005). Global gene expression profiling and cluster analysis in *Xenopus laevis*. *Mech. Dev.* 122: 441-475.
- 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.
- BILDER, D. and PERRIMON, N. (2000). Localization of apical epithelial determinants by the basolateral PDZ protein Scribble. *Nature* 403: 676-680.
- CHOMCZYNSKI, P. and SACCHI, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162: 156-159.
- DE ROBERTIS, E.M. and KURODA, H. (2004). Dorsal-ventral patterning and neural induction in *Xenopus* embryos. *Annu. Rev. Cell Dev. Biol.* 20: 285-308.
- DE ROBERTIS, E.M., LARRAIN, J., OELGESCHLAGER, M. and WESSELY, O. (2000). The establishment of Spemann's organizer and patterning of the vertebrate embryo. *Nat. Rev. Genet.* 1: 171-181.
- DONG, B., HOROWITZ, D.S., KOBAYASHI, R. and KRAINER, A.R. (1993). Purification and cDNA cloning of HeLa cell p54nrb, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and *Drosophila* NONA/BJ6. *Nucleic Acids Res.* 21: 4085-4092.
- EAGLESON, G., FERREIRO, B. and HARRIS, W.A. (1995). Fate of the anterior neural ridge and the morphogenesis of the *Xenopus* forebrain. *J. Neurobiol.* 28: 146-158.
- EAGLESON, G.W. and HARRIS, W.A. (1990). Mapping of the presumptive brain regions in the neural plate of *Xenopus laevis*. *J. Neurobiol.* 21: 427-440.
- GAMSE, J. and SIVE, H. (2000). Vertebrate anteroposterior patterning: the *Xenopus* neurectoderm as a paradigm. *Bioessays* 22: 976-986.
- GAWANTKA, V., POLLET, N., DELIUS, H., VINGRON, M., PFISTER, R., NITSCH, R., BLUMENSTOCK, C. and NIEHRS, C. (1998). Gene expression screening

- in *Xenopus* identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mech. Dev.* 77: 95-141.
- HARLAND, R. and GERHART, J. (1997). Formation and function of Spemann's organizer. *Annu. Rev. Cell Dev. Biol.* 13: 611-667.
- HARLAND, R.M. (1991). *In situ* hybridization: an improved whole-mount method for *Xenopus* embryos. In *Methods Cell Biol.*, vol. 36 (ed. KAY, B. K. and PENG, H. B.). Academic Press, San Diego, CA, pp.685-695.
- HOLLEMANN, T., CHEN, Y., GRUNZ, H. and PIELER, T. (1998). Regionalized metabolic activity establishes boundaries of retinoic acid signalling. *EMBO J.* 17: 7361-7372.
- KUDOH, T., TSANG, M., HUKRIEDE, N.A., CHEN, X., DEDEKIAN, M., CLARKE, C.J., KIANG, A., SCHULTZ, S., EPSTEIN, J.A., TOYAMA, R. *et al.* (2001). A gene expression screen in zebrafish embryogenesis. *Genome Res.* 11: 1979-1987.
- KUDOH, T., WILSON, S.W. and DAWID, I.B. (2002). Distinct roles for Fgf, Wnt and retinoic acid in posteriorizing the neural ectoderm. *Development* 129: 4335-4346.
- LEE, Y.H., HUANG, G.M., CAMERON, R.A., GRAHAM, G., DAVIDSON, E.H., HOOD, L. and BRITTEN, R.J. (1999). EST analysis of gene expression in early cleavage-stage sea urchin embryos. *Development* 126: 3857-3867.
- LUMSDEN, A. and KRUMLAUF, R. (1996). Patterning the vertebrate neuraxis. *Science* 274: 1109-1115.
- MANGOLD, O. (1933). Über die Induktionsfähigkeit der verschiedenen Bezirke der Neurula von Urodelen. *Naturwissenschaften* 43: 761-766.
- MUNOZ-SANJUAN, I. and BRIVANLOU, A.H. (2002). Neural induction, the default model and embryonic stem cells. *Nat. Rev. Neurosci.* 3: 271-280.
- MURDOCH, J.N., HENDERSON, D.J., DOUDNEY, K., GASTON-MASSUET, C., PHILLIPS, H.M., PATERNOTTE, C., ARKELL, R., STANIER, P. and COPP, A.J. (2003). Disruption of scribble (*Scrb1*) causes severe neural tube defects in the circletail mouse. *Hum. Mol. Genet.* 12: 87-98.
- NIEHRS, C. (1999). Head in the WNT: the molecular nature of Spemann's head organizer. *Trends Genet.* 15: 314-319.
- NIEUWKOOP, P.D. and FABER, J. (1967). *Normal Table of Xenopus laevis (Daudin)*. North Holland, Amsterdam.
- PASTERKAMP, R.J. and KOLODKIN, A.L. (2003). Semaphorin junction: making tracks toward neural connectivity. *Curr. Opin. Neurobiol.* 13: 79-89.
- POLLET, N., MUNCKE, N., VERBEEK, B., LI, Y., FENGER, U., DELIUS, H. and NIEHRS, C. (2005). An atlas of differential gene expression during early *Xenopus* embryogenesis. *Mech. Dev.* 122: 365-439.
- PUELLES, L. and RUBENSTEIN, J.L. (2003). Forebrain gene expression domains and the evolving prosomeric model. *Trends Neurosci.* 26: 469-476.
- QU, X., WEI, H., ZHAI, Y., QUE, H., CHEN, Q., TANG, F., WU, Y., XING, G., ZHU, Y., LIU, S. *et al.* (2002). Identification, characterization and functional study of the two novel human members of the semaphorin gene family. *J. Biol. Chem.* 277: 35574-35585.
- SAKA, Y. and SMITH, J.C. (2001). Spatial and temporal patterns of cell division during early *Xenopus* embryogenesis. *Dev. Biol.* 229: 307-318.
- SATOU, Y., TAKATORI, N., FUJIWARA, S., NISHIKATA, T., SAIGA, H., KUSAKABE, T., SHIN-I, T., KOHARA, Y. and SATOH, N. (2002). *Ciona intestinalis* cDNA projects: expressed sequence tag analyses and gene expression profiles during embryogenesis. *Gene* 287: 83-96.
- SATOU, Y., TAKATORI, N., YAMADA, L., MOCHIZUKI, Y., HAMAGUCHI, M., ISHIKAWA, H., CHIBA, S., IMAI, K., KANO, S., MURAKAMI, S.D. *et al.* (2001). Gene expression profiles in *Ciona intestinalis* tailbud embryos. *Development* 128: 2893-2904.
- SCHOHL, A. and FAGOTTO, F. (2002). Beta-catenin, MAPK and Smad signaling during early *Xenopus* development. *Development* 129: 37-52.
- SHIBATA, M., ITOH, M., OHMORI, S.Y., SHINGA, J. and TAIRA, M. (2001). Systematic screening and expression analysis of the head organizer genes in *Xenopus* embryos. *Dev. Biol.* 239: 241-256.
- SHIN, Y., KITAYAMA, A., KOIDE, T., PEIFFER, D.A., MOCHII, M., LIAO, A., UENO, N. and CHO, K.W. (2005). Identification of neural genes using *Xenopus* DNA microarrays. *Dev. Dyn.* 232: 432-444.
- SHINGA, J., ITOH, M., SHIOKAWA, K., TAIRA, S. and TAIRA, M. (2001). Early patterning of the prospective midbrain-hindbrain boundary by the *HES*-related gene *XHR1* in *Xenopus* embryos. *Mech. Dev.* 109: 225-239.
- STRAUB, T., GRUE, P., UHSE, A., LISBY, M., KNUDSEN, B.R., TANGE, T.O., WESTERGAARD, O. and BOEGE, F. (1998). The RNA-splicing factor PSF/p54 controls DNA-topoisomerase I activity by a direct interaction. *J. Biol. Chem.* 273: 26261-26264.
- STRAUB, T., KNUDSEN, B.R. and BOEGE, F. (2000). PSF/p54(nrb) stimulates "jumping" of DNA topoisomerase I between separate DNA helices. *Biochemistry* 39: 7552-7558.
- TAIRA, M., SAINT-JEANNET, J.-P. and DAWID, I.B. (1997). Role of the *Xlim-1* and *Xbra* genes in anteroposterior patterning of neural tissue by the head and trunk organizer. *Proc. Natl. Acad. Sci. USA* 94: 895-900.
- TANABE, Y. and JESSELL, T.M. (1996). Diversity and pattern in the developing spinal cord. *Science* 274: 1115-1123.
- VOIGT, J., CHEN, J.A., GILCHRIST, M., AMAYA, E. and PAPALOPULU, N. (2005). Expression cloning screening of a unique and full-length set of cDNA clones is an efficient method for identifying genes involved in *Xenopus* neurogenesis. *Mech. Dev.* 122: 289-306.
- VON BESSER, H., SCHNABEL, P., WIELAND, C., FRITZ, E., STANEWSKY, R. and SAUMWEBER, H. (1990). The puff-specific *Drosophila* protein Bj6, encoded by the gene *no-on transient A*, shows homology to RNA-binding proteins. *Chromosoma* 100: 37-47.
- YANG, Y.S., HANKE, J.H., CARAYANNOPOULOS, L., CRAFT, C.M., CAPRA, J.D. and TUCKER, P.W. (1993). NonO, a non-POU-domain-containing, octamer-binding protein, is the mammalian homolog of *Drosophila* nonAdiss. *Mol. Cell. Biol.* 13: 5593-5603.
- ZHANG, Z. and CARMICHAEL, G.G. (2001). The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* 106: 465-475.

Received: June 2005

Reviewed by Referees: July 2005

Modified by Authors and Accepted for Publication: August 2005

Edited by: Makoto Asashima