MOLECULAR AND CELLULAR BIOLOGY, Oct. 1996, p. 5865–5875 Vol. 16, No. 10
0270-7306/96/$04.00+0

Two New Members of the Murine *Sim* Gene Family Are Transcriptional Repressors and Show Different Expression Patterns during Mouse Embryogenesis

MASATSUGU EMA,[1] MASANOBU MORITA,[1] SHUNTARO IKAWA,[2] MASAHIRO TANAKA,[1] YOUICHI MATSUDA,[3] OSAMU GOTOH,[4] YUKIO SAIJOH,[5] HIDETA FUJII,[5] HIROSHI HAMADA,[5] YASUO KIKUCHI,[1] AND YOSHIAKI FUJII-KURIYAMA[1]*

Department of Chemistry, Graduate School of Science,[1] and Department of Molecular Genetics, Institute of Development, Aging and Cancer,[2] Tohoku University, Sendai 980-77, Genome Research Group, National Institute of Radiological Sciences, Inage, Chiba,[3] Department of Biochemistry, Saitama Cancer Research Institute, Inamachi, Saitama 362,[4] and Department of Developmental Biology and Cancer Prevention, Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo 113,[5] Japan

Received 22 March 1996/Returned for modification 17 May 1996/Accepted 9 July 1996

From a cDNA library of mouse skeletal muscle, we have isolated mouse *Sim1* (*mSim1*) cDNA encoding a polypeptide of 765 amino acids with striking amino acid identity in basic helix-loop-helix (89% identity) and PAS (89% identity) domains to previously identified mSim2, although the carboxy-terminal third of the molecule did not show any similarity to mSim2 or *Drosophila Sim* (dSim). Yeast two-hybrid analysis and coimmunoprecipitation experiments demonstrated that both of the mSim gene products interacted with Arnt even more efficiently than AhR, a natural partner of Arnt, suggesting a functional cooperativity with Arnt. In sharp contrast with dSim having transcription-enhancing activity in the carboxy-terminal region, the two mSims possessed a repressive activity toward Arnt in the heterodimer complex. This is the first example of bHLH-PAS proteins with transrepressor activity, although some genetic data suggest that dSim plays a repressive role in gene expression (Z. Chang, D. Price, S. Bockheim, M. J. Boedigheimer, R. Smith, and A. Laughon, Dev. Biol. 160:315–332, 1993; D. M. Mellerick and M. Nirenberg, Dev. Biol. 171:306–316, 1995). Whole-mount in situ hybridization showed restricted and characteristic expression patterns of the two *mSim* mRNAs in various tissues and organs during embryogenesis, such as those for the somite, the nephrogenic cord, and the mesencephalon (for *mSim1*) and those for the diencephalon, branchial arches, and limbs (for *mSim2*). From sequence similarity and chromosomal localization, it is concluded that *mSim2* is an ortholog of *hSIM2*, which is proposed to be a candidate gene responsible for Down's syndrome. The sites of *mSim2* expression showed an overlap with the affected regions of the syndrome, further strengthening involvement of mSim2 in Down's syndrome.

Powerful genetics and embryonic manipulation have allowed detailed anatomical studies of the midline development of the central nervous system (CNS) in *Drosophila* species (6). Several mutations affecting the developmental processes of ventral midline *Drosophila* cells have been identified in *single-minded* (*Sim*), *slit*, *otd*, *spi*, *rho*, and *star* genes (10, 29, 46, 47, 48, 50, 67, 75). Among these genes, the *Sim* gene, which is expressed in all midline precursors, is thought to play a pivotal role in midline development, for its defectiveness leads to the absence of midline lineage cells, including three pairs of midline glia and neurons, and anterior and posterior axon commissures, which create longitudinal axon fusion in the midline (29, 67). Although the mechanism of its action in the development and/or maintenance of the CNS midline remains unelucidated, Sim function in transcriptional activation is required for the expression of various genes in CNS midline cells, including *Sim* itself, *Toll*, and *slit* (46).

Furthermore, it is of interest that the human *Sim* homolog (*hSIM2*) was identified in the Down's syndrome critical region (DSCR), the cause of Down's syndrome, which is characterized by mental retardation, congenital heart defect, and dysmorphic face and limbs in its trisomic state (8, 44, 45). Taken together, these findings suggest that mammalian *Sim* also plays a key role in the development of the neural system and other organs.

Molecular cloning of the *Drosophila Sim* (*dSim*) gene revealed that its product possesses characteristic basic helix-loop-helix (bHLH) and PAS (designated as a conserved region among Per, AhR-Arnt, and Sim) motifs (5, 10, 13, 23, 26, 27). Recently, it has been found that these structural motifs are also shared by human HIF1α (hypoxia-inducible factor 1α) (68), the *Drosophila trachealess* gene product (26, 74), and mouse Arnt2 (mArnt2) (22). Analogously to transcriptional function of the AhR-Arnt heterodimer involved in the inducible expression of the *CYP1A1* gene by xenobiotics, it was speculated that dSim works as a transcriptional regulator in association with a *Drosophila* Arnt (dArnt) homolog and/or some other partner molecule(s) (70). Recently, dSim has been further characterized to be a nuclear protein with a transcriptional activation domain in the carboxy-terminal half (17). A cDNA for a mouse homolog (*mSim2*, previously designated *mSim* by us [14]), renamed according to Muenke et al. [44]) of *dSim* has been isolated (14). The properties of the encoded protein have been partially characterized by whole-mount in situ hybridization and biochemical methods. Interestingly, in the embryonic stage 8.5 days postcoitus (dpc), mSim2 is expressed exclusively in the ventral diencephalon and forms a heterodimer with Arnt, suggesting that mSim2 works as a transcriptional regulator in

* Corresponding author. Mailing address: Department of Chemistry, Graduate School of Science, Tohoku University, Sendai 980-77, Japan. Phone: 81-22-217-6590. Fax: 81-22-217-6594.

mouse neurogenesis as a heterodimer with Arnt (14). During the course of mSim2 analysis, we isolated another cDNA clone, designated *mSim1* cDNA, which encodes a polypeptide highly similar to but distinct from mSim2.

In the present paper, we describe the predicted amino acid sequence of mSim1 and compare the transcriptional activities of the two mSim proteins. Despite the similarity in their transrepressor activities, the two mSim mRNAs showed different expression patterns during mouse embryogenesis. Chromosomal localization of mSim1 is also assigned to the B3-B4 band of the mouse chromosome 10, which is different from the previously determined mSim2 localization.

MATERIALS AND METHODS

Isolation of *mSim1* cDNA. In the course of screening a mouse skeletal muscle cDNA library with *mSim2* cDNA (previously designated *mSim* in reference 14), we found a cDNA clone different from *mSim2* and termed it *mSim1*. *mSim1* cDNA, with a size of 1.2 kb, was used as a probe to screen 2×10^6 plaques of the same library. Two positive clones were isolated, and phage DNAs were purified by the standard protocol (60). cDNA inserts were subcloned into pBluescript SK vector (Stratagene), and both strands were sequenced by the dideoxy chain termination method (61).

FISH for chromosome mapping. The direct R-banding fluorescent in situ hybridization (FISH) method was used for chromosomal assignment of the *mSim1* gene on the mouse chromosome. R-banded chromosomes and FISH were prepared as described elsewhere (38, 39). The mouse cDNA probe, a 1.6-kb fragment in pBluescript, was labeled by nick translation with biotin 16-dUTP (Boehringer Mannheim) according to the manufacturer's protocol. Excitations at wavelengths of 450 to 490 nm (Nikon filter set B-2A) and near 365 nm (UV-2A) were used for observation. Kodak Ektachrome ASA 100 films were used for microphotography.

Recombinant plasmids. pEFBOS (42), pSG424 (58), and pG5EC (34) were described previously. pEFBOS-GALDBD (amino acids [aa] 1 to 147) and pEFBOS-GALDBD (aa 1 to 147)-Arnt will be described elsewhere (49). pGBT-AhR, pGBT-hArnt, pGBT-Arnt2, pGAD-AhR, pGBT-Arnt, and pGBT-Arnt2 were described previously (22). pGBT-mSim2 (aa 1 to 324) and pGAD-mSim2 (aa 1 to 324) were constructed by inserting the cDNA fragment corresponding to aa 1 to 324 into the *Sma*I sites of pGBT9 and pGAD424, respectively. Similarly, pGBT-mSim1 and pGAD-mSim1 were constructed by inserting the cDNA fragment corresponding to aa 1 to 322 into the *Sma*I sites of pGBT9 and pGAD424, respectively.

pEF-BOS-mSim1 and pEF-BOS-mSim2 were constructed by subcloning *mSim1* cDNA (nucleotides [nt] 25 to 2547) and *mSim2* cDNA (nt 577 to 2531) into the *Xba*I site of pEF-BOS, which had been blunt ended by T4 DNA polymerase, respectively. pEF-BOS-mSim1 (aa 1 to 322) and -mSim2 (aa 1 to 324) were constructed by inserting *mSim1* cDNA (nt 25 to 1005) and *mSim2* cDNA (nt 557 to 1638) into the *Xba*I site of pEF-BOS, which had been blunt ended. pSG-dSim (aa 398 to 676), pSG-mSim1 (aa 348 to 765), and pSG-mSim2 (aa 362 to 657) were constructed by subcloning cDNA fragments corresponding to nt 1834 to 2741, nt 1080 to 2547, and nt 1638 to 2908 of *dSim*, *mSim1*, and *mSim2*, respectively, into the blunt-ended multicloning site of pSG424. Then, the cloned *Bgl*II-*Sac*I fragment harboring GALDBD (aa 1 to 147) and Sim carboxy-terminal halves were cut off from the pSG derivatives and ligated with the *Xba*I site of pEF-BOS, which had been already blunt ended to generate highly efficient expression vectors under the control of the elongation factor 1α promoter. These constructions were confirmed by sequencing analysis.

Coimmunoprecipitation assay in vitro. Full-length *mSim* cDNAs and human p53 cDNA cloned into the pBluescript vector were used to generate ^{35}S-labeled mSim proteins and hp53 protein in vitro by the TNT-coupled reticulocyte lysate system (Promega) and T3 (for mSim1) or T7 (for mSim2 and hp53) RNA polymerase. The protein interaction assay was performed as follows. Approximately 10 μg of baculovirus-produced Arnt protein (64) was incubated with ^{35}S-labeled mSim protein or hp53 in 30 μl of NETN buffer (150 mM NaCl, 1 mM EDTA, 50 mM Tris [pH 7.5], 0.1% Nonidet P-40) for 1 h at 23°C and mixed with 200 μl of NETN buffer containing 1% blocking reagent (Boehringer Mannheim) and either anti-Arnt antiserum (64) or control normal serum. After incubation for 1 h at 4°C, protein A-Sepharose was added, and the mixture was further incubated for 1 h at 4°C with rocking and was then centrifuged. The precipitates were washed three times with 1 ml of NETN buffer, solubilized with Laemmli buffer, and finally analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Cell lines and transfection. Both COS-7 (monkey kidney cell line) and 293T (human embryonic kidney cell line) cells were maintained as monolayers in Dulbecco's modified Eagle's medium supplemented with 10% (vol/vol) fetal calf serum. The calcium phosphate method was used to introduce the expression plasmids into the cells (19). At the same time, a *lacZ*-expressing vector, pENL (a kind gift from Y. Nabeshima, National Institute of Neuroscience, Tokyo, Japan), was cotransfected as a standard for normalization. After 48 h of transfection, the cells were harvested and the cell lysates were subjected to assays of chloramphenicol acetyltransferase (CAT) activities by using [^{14}C]chloramphenicol (ICN) and of β-galactosidase activities by using *o*-nitrophenyl-β-D-galactopyranoside (ONPG).

Immunoblot analysis. Immunoblot analysis was performed with an anti-GAL4 DNA-binding domain (DBD) antibody (65) as described previously (12).

Yeast two-hybrid system. Protein-protein interaction was investigated by the two-hybrid system developed by P. L. Bartel (4).

β-Galactosidase activities were determined as follows. SFY526 (*MAT***a** *ura3 his3 ade2 lys2 trp1 leu2 can*^{r} *gal4 gal80 URA3*::*GAL1-lacZ*) was transformed with each pair of the fusion plasmids containing the GAL4 DBD and activation domain. Three independent colonies were inoculated in selecting medium for both leucine and tryptophan requirements and were shaken for 10 h (in the presence of 3-methylcholanthrene in the case of AhR). Cell extracts from 1.5 ml of culture were prepared and incubated with ONPG used as a substrate for 1 h, in parallel with the control reaction with extracts of nontransfected cells. Three independent experiments were performed, and β-galactosidase activities were determined on the basis of the following formula: $1{,}000 \times OD_{420} \div T \div V \div OD_{600}$, where T is time of reaction, V is volume of culture used in the assay (in milliliters), and OD_{420} and OD_{600} are optical densities at 420 and 600 nm, respectively.

RNA blot analysis. Preparation of total RNAs from various tissues was performed according to a previously published method (9). RNA blot analysis was performed as described previously (60). *mSim1* cDNA (nt 1572 to 2174) was used as a probe. The *mSim2* probe has already been described (14).

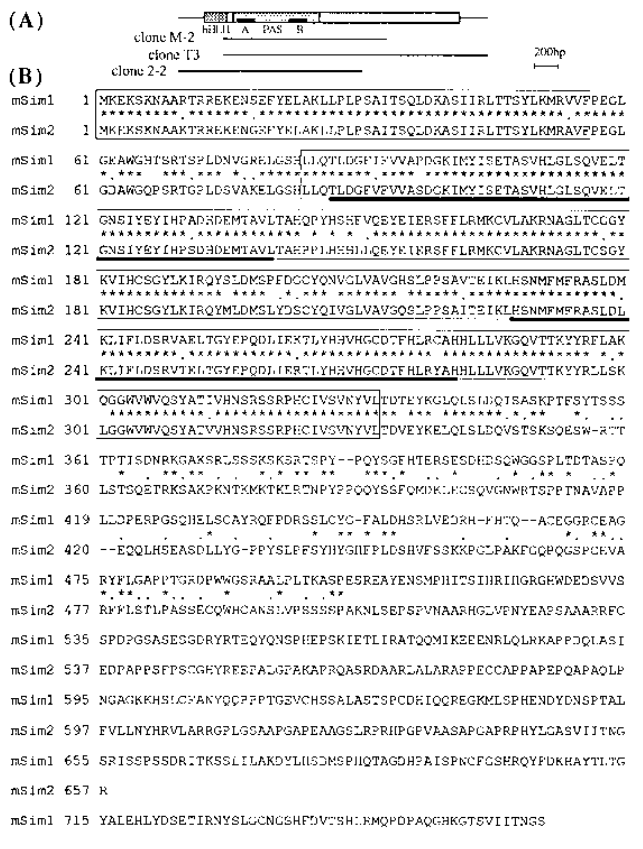
Whole-mount in situ hybridization. Embryos from 7.5 to 10.5 dpc were analyzed for *mSim1*, *mSim2*, and *Sonic hedgehog* (*Shh*) expression by whole-mount in situ hybridization with digoxigenin-labeled RNA probes as previously described (75). The *mSim2* probe (nt 2064 to 2908) and *mSim1* probe (nt 1511 to 1870) in pBluescript were transcribed in the sense and the antisense directions by T3 and T7 RNA polymerases, respectively. The mouse *Shh* probe was described previously (11).

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper are registered in the DDBJ, EMBL, and GenBank nucleotide sequence databases under accession number D79209.

RESULTS

Cloning of mSim1, a novel member of the murine *Sim* gene family. In the process of screening the murine skeletal muscle cDNA library with *mSim2* cDNA (previously designated *mSim*) as a labeled probe, we isolated another cDNA clone, designated *mSim1* cDNA, which is distinct from but closely related to *mSim2*. Two additional independent clones were isolated by screening approximately 2×10^6 plaques of the above-described library, altogether encompassing an entire coding sequence (Fig. 1A). The compiled cDNA sequence consisted of a 5′ untranslated region (UTR) sequence of 222 bp and a subsequent long open reading frame of 2,298 bp which was followed by a 3′-UTR sequence of 212 bp. It encoded a polypeptide of 765 aa with a calculated molecular mass of 87 kDa. The sequence (ACGATGA) surrounding the putative initiation codon showed reasonable agreement with the Kozak consensus sequence (5 of 7) (30). The amino acid sequences of mSim1 and mSim2 revealed a striking identity in their amino-terminal halves (aa 1 to 332), including the bHLH (89% identity) and PAS (89% identity) domains (Fig. 1B), which were followed by a sequence with intermediate identity (33% [aa 333 to 504]). However, we could not detect any obvious similarity in the further carboxy-terminal region [aa 505 to 765]).

Comparison of the amino acid sequences of the two genes with the partial sequences of *hSIM2*, which were previously identified in the DSCR by exon trapping, suggests that the *mSim2* gene is an ortholog of the *hSIM2* gene. Although availability of the *hSIM2* sequence is limited to the bHLH region and some part of the PAS region, the amino acid sequence of the bHLH region of *hSIM2* (8) completely matched that of *mSim2* but not that of *mSim1* (data not shown). When the amino acid sequence in the bHLH-PAS region of *mSim1* was compared with those of other members of the bHLH-PAS family genes, *mSim1* was most closely related to *mSim2* and, together with *mSim2*, was classified into the AhR group of the



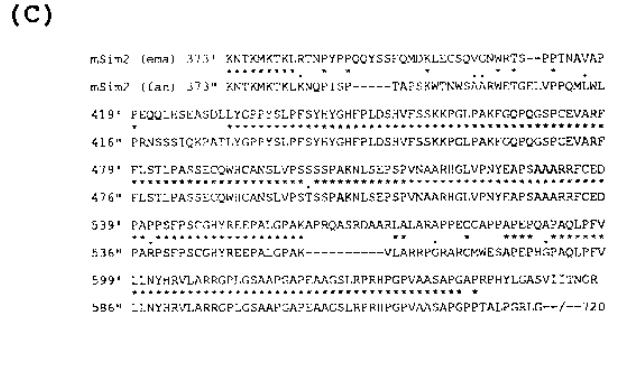


FIG. 1. Comparison of primary structures of *mSim* gene products. (A) cDNAs encoding mSim1 are schematically represented. Box, amino acid coding sequences; thin lines, untranslated sequences; A and B, direct repeat within the PAS domain. (B) Identical amino acids (∗) and conservative substitutions (·) are indicated. Dashes, deletions to maximize the sequence similarity; boxes, the bHLH region and PAS domain; solid bars, direct repeats of the PAS domain. (C) Comparison of our mSim2 amino acid sequence with that of Fan et al. (16). Only variable parts of the two sequences are presented. (D) Nucleotide sequences of mSim2 around those causing the amino acid differences between the two mSim2s as shown in panel C. C_{1697}TG codes for Arg-382 in our sequence, C_{1718}CA codes for Pro-388, C_{1844}TT codes for Leu-430, G_{2236}CG codes for Ala-561, G_{2258}AC codes for Asp-568, and G_{2473}GC codes for Gly-640.

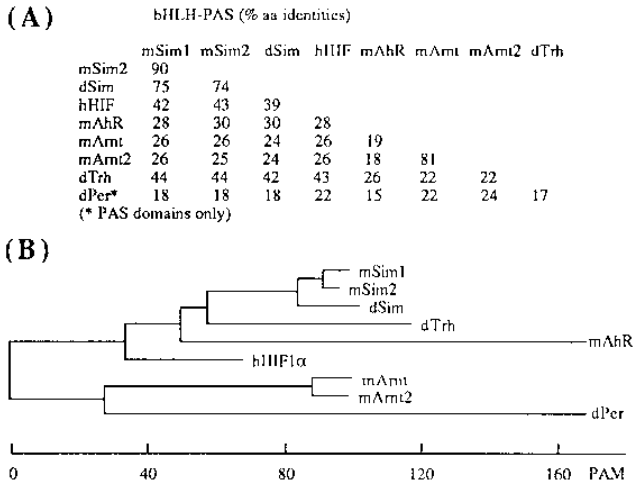


FIG. 2. Comparison of primary structure of mSim1 with those of the previously isolated bHLH-PAS proteins. (A) Homology matrix among bHLH-PAS proteins. Sequence identities were calculated on the basis of the alignment of the sequences by the method described by Gotoh (20). (B) Evolutionary relation of PAS domains. The phylogenetic tree was constructed by the method described by Saitou and Nei (59). Abbreviations: m, mouse; h, human; d, *Drosophila*; PAM, accepted point mutations per 100 sites.

two major groups of the bHLH-PAS family (Fig. 2). Very recently, the sequences of the two mSim homologs were reported by Fan et al. The two sequences are very similar to but could not be said to be essentially consistent with ours. Concerning the mSim1 sequence, our mSim1 sequence is longer than that of Fan et al. by 70 aa. This is due to the nucleotide differences in the sequence around the termination codon. The sequence T_{2332}GAAAAGGC is replaced by TGTTTTGGC in ours, resulting in an elongation of 70 aa. A preliminary experiment shows that the mouse genomic sequence containing the termination codon of the *mSim1* gene matches that of our cDNA (24). The following five additional amino acid replacements were found (first, from our sequence; second, from Fan et al. [16]): His-133 (CAC) for Leu (CTC), Thr-176 (ACC) for a deletion, Pro-322 (CCG) for Arg (CGG), Ala-480 (GCA) for Pro (CCA), and Asp-537 (GAC) for Ser (AGC). The sequence differences are presumably partly due to sequencing errors, and partly polymorphism. The in vitro translation system gave a product of 90 kDa, in good agreement with the predicted molecular mass (87 kDa) from our amino acid sequence (see Fig. 4A).

Concerning mSim2, our mSim2 amino acid sequence is identical to that of Fan et al. as far as position 381, except for amino acid replacement of Arg-263 (AGG) (ours) with Lys (AAG) (Fan et al. [16]) and of Glu-336 (GAA) with Gly (GGA). From position 382 to the C terminus, the two sequences have many deletions and amino acid replacements with some stretches of identical amino acids. These amino acids and nucleotide replacements and deletions are presented in Fig. 1C and D. The two amino acid sequences are predicted to consist of 657 (ours) and 720 (Fan et al.) aa, respectively. The molecular mass of the in vitro-translated product (75 kDa) of our *mSim2* mRNA is in good agreement with the one predicted (73 kDa) from the amino acid sequence (see Fig. 4A). The sequence difference between the two could not be explained simply by sequencing errors or by polymorphism, but most probably could be explained by alternative splicing. This will be clarified by cloning and sequencing of genomic DNA of *mSim2.*

Chromosomal localization of *mSim1*. To confirm orthology

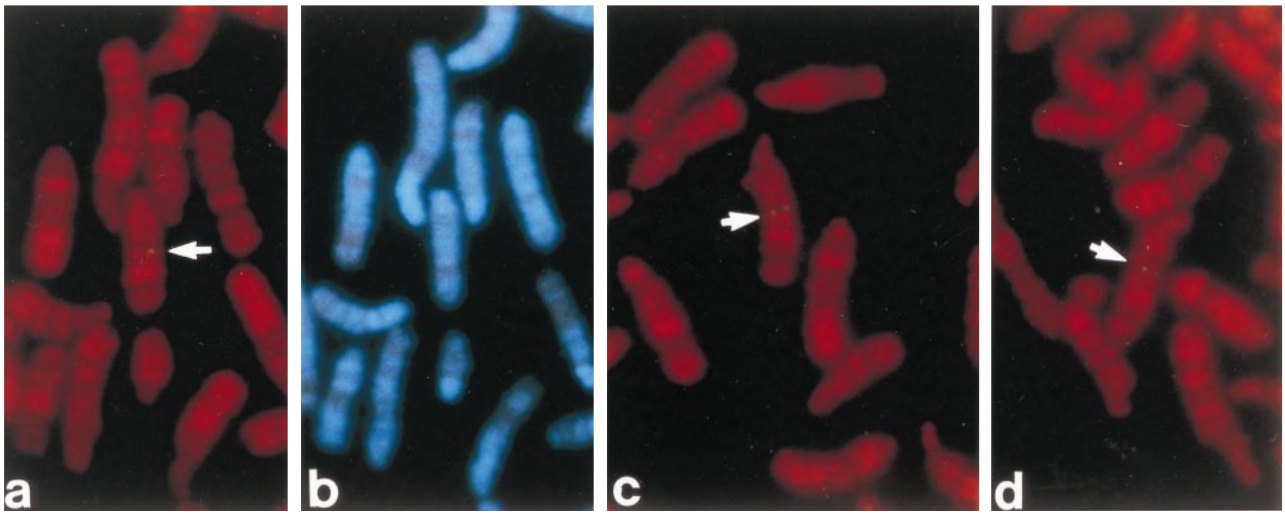


FIG. 3. Chromosomal localization of the *mSim1* gene on mouse R-banded chromosomes. An mSim1 cDNA fragment (1.6 kb) was used as a biotinylated probe. The hybridization signals are indicated by arrows. The metaphase spreads were photographed with Nikon B-2A (a, c, and d) and UV-2A (b) filters. R-banded (a, c, and d) and G-banded (b) patterns are demonstrated.

between the two *Sim*-related genes and the *hSIM2* gene, we determined chromosomal localization of *mSim2* by FISH. In agreement with the recent localization of this gene by using an interspecific back-crossing panel (16), the *mSim1* gene was assigned to the B3-B4 band of mouse chromosome 10, which is syntenic with human chromosome 6q22-31 (Fig. 3). This localization is in contrast with the previously determined localization of the *mSim2* gene at the C3.3-C4 band of the mouse chromosome 16, which is syntenic with human chromosome 21q22.2, where *hSIM2* gene was mapped (14), thus further strengthening the interpretation that the *mSim2* gene is orthologous to the *hSIM2*.

Dimerization of mSim proteins with other PAS family proteins. In general, the bHLH-PAS proteins work as a dimer complex with other members of the bHLH-PAS family or sometimes with themselves (40, 53, 68, 71). The mode of dimerization of the two mSim proteins with other bHLH-PAS proteins was investigated with the yeast two-hybrid system (Table 1) and by coimmunoprecipitation assaying (Fig. 4B). As shown in Table 1, both mSim1 and mSim2 interacted with Arnt or Arnt2 even more efficiently than AhR heterodimerized with Arnt or Arnt2. In confirmation of the previous results, Arnt and Arnt2 also interacted with themselves, albeit with less affinity (22, 64), while none of the AhR and Sim group, including HIF1α, interacted with a member of the same group. In agreement with these results, mSim1 and mSim2 were coimmunoprecipitated with Arnt by anti-Arnt antibody when they were incubated with Arnt (Fig. 4B). The dimerization properties of the bHLH-PAS proteins are reflected in their grouping on the basis of structural relatedness in the bHLH-PAS domain (Fig. 2). The bHLH-PAS proteins are structurally divided into two major groups, as shown in Fig. 2B. A member of the AhR group forms a dimer only with a member of the other group, while a member of the Arnt group forms a dimer with a member of either group, although the homodimer formation is generally weak. Since the PAS as well as the HLH domain is involved in the dimerization interaction, information about the three-dimensional structure of the PAS domain is needed to explain the dimerization properties.

Transcriptional activities of the two mSim proteins. We investigated the transcriptional activities of mSim1 and mSim2 in mammalian and yeast expression systems with a reporter plasmid containing the GAL4-binding site (upstream activating sequence) and an expression plasmid of fusion proteins containing the GAL4 DBD (Fig. 5).

Since the bHLH-PAS transcriptional factors are known to have a transcriptional activation domain in the carboxy-terminal half of the molecule (17, 22, 28, 32, 65, 72), we constructed expression plasmids to assay the transcriptional activities of mSim1 and mSim2 by fusing the cDNA sequences encoding the carboxy-terminal halves of mSim1 and mSim2 (aa 348 to 765 and 362 to 657, respectively) with the GAL4 DBD. An

TABLE 1. Interaction between mSim family and other PAS proteins as revealed by the yeast two-hybrid system

DBD fusion	β-Galactosidase activity (Miller units) for GAL4 activation domain fusions[a]						
	GAD424	mSim1	mSim2	Arnt	Arnt2	AHR	HIF1α
GBT9	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02
mSim1	<0.02	<0.02	<0.02	10.1 ± 0.5	18.0 ± 1.1	<0.02	<0.02
mSim2	<0.02	<0.02	<0.02	17.6 ± 5.7	10.4 ± 0.5	<0.02	<0.02
Arnt	<0.02	0.970 ± 0.448	0.671 ± 0.072	0.066 ± 0.016	0.094 ± 0.087	0.193 ± 0.170	1.71 ± 0.24
Arnt2	<0.02	1.43 ± 0.30	ND	0.078 ± 0.025	0.065 ± 0.062	0.166 ± 0.086	3.49 ± 0.24
AHR	<0.02	<0.02	<0.02	0.849 ± 0.449	ND	<0.02	<0.02
HIF1α	<0.02	<0.02	<0.02	14.5 ± 5.9	21.1 ± 1.0	<0.02	<0.02

[a] Yeast strain SFY526 containing *lacZ* under the control of the *GAL1* promoter was cotransformed with the indicated plasmids, and β-galactosidase activities were determined as described in Materials and Methods. ND, not determined.

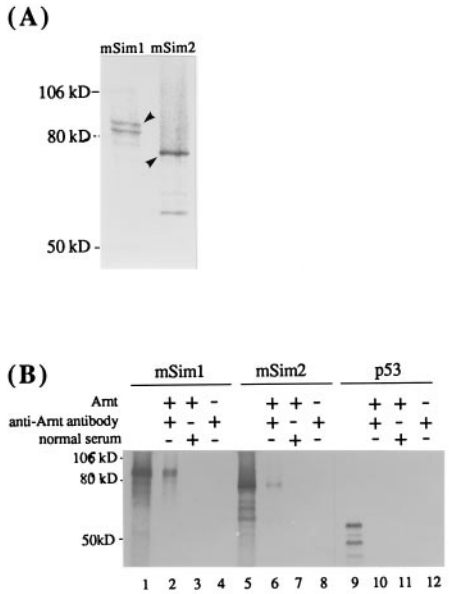


FIG. 4. Interaction of mSim1 and mSim2 with Arnt as revealed by coimmunoprecipitation assay. (A) SDS-PAGE of mSim1 and mSim2 produced in an in vitro translation system. Arrowheads indicate the locations of the complete translation products. (B) Interaction of the two mSim proteins with Arnt was investigated by coimmunoprecipitation assaying with extracts from Sf9 transfected with baculovirus carrying Arnt cDNA and ^{35}S-labeled mSim proteins synthesized in the in vitro reticulocyte translation system. ^{35}S-labeled proteins and antiserum used in the experiment are indicated at the top. Lanes 1, 5, and 9, input proteins. Incubation of Arnt (10 μg of cell extracts) with in vitro-synthesized mSim1 or mSim2 was performed at 23°C for 1 h, and then the reaction mixture was immunoprecipitated with anti-Arnt antibody (20 μg) or nonimmunized serum (20 μg). ^{35}S-labeled p53 was used as a negative control (lanes 9 to 12). Protein A-Sepharose was added to the reaction mixture, and the precipitates were dissolved in the sample buffer and subjected to electrophoresis in a 10% polyacrylamide gel containing 1% SDS. The gel was analyzed by a Fuji BAS1000 PhosphorImager analyzer.

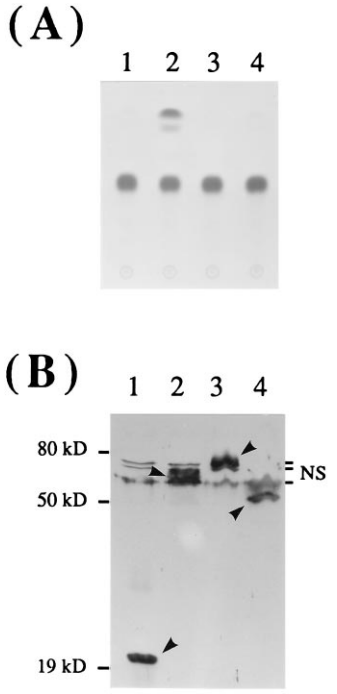


FIG. 5. Transcriptional activity of the carboxy-terminal halves of mSim1 and mSim2 in fusion proteins with GAL4 DBD. (A) Transcriptional activity of the carboxy-terminal half of mSim1 and mSim2. A reporter plasmid (pG5EC [4 μg]) and a chimeric plasmid encoding GAL DBD (aa 1 to 147) and the carboxy-terminal halves of mSim1 and mSim2 (4 μg) were cotransfected into 293T cells in 6-cm dishes by the calcium phosphate method. After 48 h of transfection, cell extracts were prepared from the cells, and their aliquots (5 μg on protein basis) were used to estimate expressed CAT activities. Lanes: 1, pEFBOS-GAL DBD (aa 1 to 147); 2, pEFBOS-GAL DBD (aa 1 to 147)-dSim (aa 398 to 676); 3, pEFBOS-GAL DBD (aa 1 to 147)-mSim1 (aa 348 to 765); 4, pEFBOS-GAL DBD (aa 1 to 147)-mSim2 (aa 362 to 657). (B) Immunoblot analysis of fusion proteins consisting of GAL DBD and the carboxy-terminal half of mSim1 or mSim2 expressed in 293T cells. The cell extracts (100 μg) were used for gel electrophoresis (10%) and were subjected to immunoblot analysis with an anti-GAL DBD antibody as described in Materials and Methods. Lanes: 1, pEFBOS-GAL DBD (aa 1 to 147); 2, pEFBOS-GAL DBD (aa 1 to 147)-dSim (aa 398 to 676); 3, pEFBOS-GAL DBD (aa 1 to 147)-mSim1 (aa 348 to 765); 4, pEFBOS-GAL DBD (aa 1 to 147)-mSim2 (aa 362 to 657). Arrowheads indicate the locations of the synthesized fusion proteins. NS, nonspecific signals.

expression plasmid consisting of the carboxy-terminal half of dSim (aa 398 to 673) fused with the GAL4 DBD was also produced as a control, because the dSim carboxy-terminal half is known to have transcriptional activation activity as described previously (17). The individual expression plasmids were cotransfected into the 293 human embryonic kidney cell line with a reporter plasmid (pG5EC) carrying the CAT gene under the control of a tandem array of GAL4 DNA-binding sites (Fig. 5A). Unexpectedly, neither the mSim1 nor the mSim2 plasmid exhibited any detectable enhancement on the background transcriptional activity (Fig. 5A, lanes 3 and 4), whereas in the control experiment the dSim plasmid displayed a potent transcriptional activation (lane 2) as previously described (17). Similar results were obtained with COS-7 cells (data not shown). Several bands of the dSim fusion protein were observed. These bands were due to posttranscriptional modifications, because the molecular mass (60 to 70 kDa) of synthesized dSim fusion protein was estimated to be higher than that predicted (50 kDa). The molecular mass (75 kDa) of the synthesized mSim1 fusion protein was also higher than that predicted (66 kDa). Since comparable amounts of these proteins were expressed (Fig. 5B), the CAT activities determined were thought to be true reflections of the transcriptional activities of their carboxy-terminal regions.

Since both of the proteins are considered to work as a heterodimer with Arnt or Arnt2 as described above, we examined how mSim1 or mSim2 affects the transcriptional activity of Arnt in a heterodimer complex. We constructed an expression plasmid of a fusion protein composed of the GAL4 DBD and Arnt. Expression plasmids of dSim, mSim1, and mSim2 were also constructed under the control of the elongation factor 1α promoter in the BOS expression vector (42). When the fusion protein of the GAL4 DBD and Arnt was transfected into COS cells along with the pG5EC reporter gene, CAT expression was enhanced because of the potent transcriptional activation activity of Arnt as reported previously (28, 32, 65, 72). Additional introduction of the dSim expression plasmid resulted in a maximal threefold increment over the CAT activity expressed by the plasmid of GAL DBD-Arnt (Fig. 6B and C, lanes 3 to 7). This result suggests that the dSim protein interacts with Arnt via the HLH-PAS domain to further enhance transcriptional activity, because dSim expression showed no transcriptional activation with the pG5EC reporter plasmid without expression of GAL DBD-Arnt (data not shown). Instead of the dSim plasmid, however, addition of the mSim1 or mSim2 expression plasmid resulted in marked reduction in CAT expression in a dose-dependent manner (Fig. 6B and C, lanes 9 to 13 and lanes 15 to 19, respectively). This reduction in CAT activity was lost by truncation of the carboxy-terminal half of mSim1 or mSim2 (Fig. 6B and C, lanes 20 and 21). Despite the diversity of their

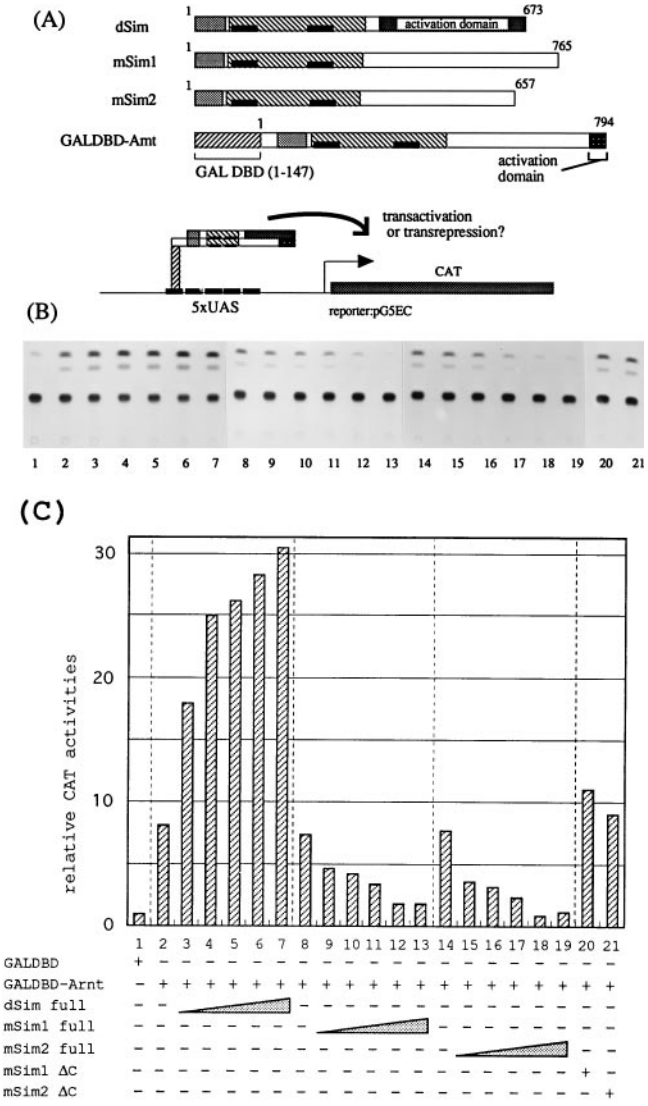


FIG. 6. Transcriptional activity of the two mSim proteins in a mammalian two-hybrid system with Arnt as a partner molecule in COS-7 cells. (A) Schematic representation of several chimeric plasmids used to estimate transcriptional activity. A fusion gene of GAL4 DBD-Arnt and *dSim*, *mSim1*, or *mSim2* cDNAs was inserted into vector pEF-BOS as described in Materials and Methods. Stippled and hatched boxes indicate bHLH domains and PAS domains, respectively. Solid bars, internal direct repeats. UAS, upstream activating sequence. (B) A reporter plasmid, pG5EC (4 μg), GAL DBD (100 ng), GAL DBD-Arnt (100 ng), and indicated expression plasmids (50, 100, and 200 ng and 0.5 and 2 μg) were cotransfected into 6-cm dishes of COS-7 cells. Cell extracts were prepared, and aliquots (50 μg) of the extracts were incubated for 1 h in the presence of [^{14}C]chloramphenicol for the CAT assay. (C) CAT activities shown in panel B were normalized by LacZ activities used as a control of transfection efficiency. pENL (2 μg) was contained in all transfection experiments. The CAT activity of pGAL DBD was defined as 1 and was used as a standard to calculate the relative CAT activities.

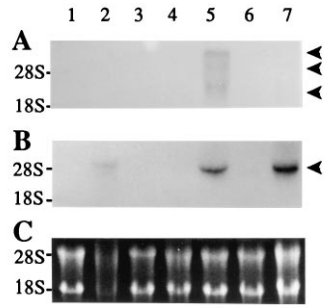


FIG. 7. RNA blot hybridization analysis with *mSim1* and *mSim2* cDNA probes. Total RNAs (20 μg) which had been prepared from various tissues of mice as described in Materials and Methods were electrophoresed in a 0.8% agarose gel containing 2.2 M formaldehyde and were then transferred to a nylon membrane. The membrane was probed with ^{32}P-labeled *mSim1* cDNA (A) and *mSim2* cDNA (B), and the electrophoresed RNAs were ethidium bromide stained (C). Lanes: 1, brain; 2, lung; 3, heart; 4; liver; 5, kidney; 6, stomach; and 7, muscle.

carboxy-terminal sequences, these results indicate that the carboxy-terminal halves of the two proteins have a transcriptional repression activity.

Spatiotemporal expression of the *mSim1* and *mSim2* genes. To study the functional roles of the two mSims, tissue distribution of *mSim1* and *mSim2* mRNAs was determined. Total RNAs were extracted from various tissues of adult mice and were subjected to RNA blot analysis to determine the expression levels of the *mSim1* and *mSim2* transcripts (Fig. 7). The two mRNAs were expressed in restricted and different tissues of the adult animals. In contrast to relatively abundant expression of mSim2 in skeletal muscle, the lungs, and the kidneys, a low level of mSim1 expression was detected in the kidney and skeletal muscle as multiple forms in length, in agreement with a low frequency of isolation of *mSim1* cDNA in the screening procedures.

To gain an insight into possible involvement of these transcription factors in embryogenesis, the spatiotemporal distribution of *mSim1* and *mSim2* mRNAs was investigated by whole-mount in situ hybridization at various stages of mouse embryogenesis from 7.5 to 10.5 dpc (Fig. 8 and 9). Unique cDNA sequences of the two mRNAs were used as labeling probes corresponding to the carboxy-terminal regions of mSim1 and mSim2 to avoid cross-hybridization.

Despite their similar transcriptional activities (see above), distinct patterns of expression of the two mRNAs were observed in mouse embryos. Expression of *mSim1* was first detected in the somites of 8.0-dpc embryos (Fig. 8B). Before differentiation of the somite, *mSim1* mRNA displayed a uniform expression level in the presomitic mesoderm (Fig. 8D). As the somite differentiated, *mSim1* mRNA expression was restricted to the dermatome but was observed neither in the sclerotome nor in the myotome (Fig. 8G and J). Its expression in the brain first became detectable in the ventrolateral regions of the rostral mesencephalon in 9.5-dpc embryos (Fig. 8E and F). In 10.5-dpc embryos, *mSim1* expression was extended rostrally to the forebrain and became detectable in the ventrolateral domain of the pretectum and the zona limitans intrathalamica of the diencephalon (Fig. 8H and I; also see Fig. 9A). It was also expressed in the ventrolateral region of the caudal neural tube, the nephrogenic cords, and the hindgut in 9.5- and 10.5-dpc embryos (Fig. 8E, G, and I). In contrast, the *mSim2* expression started by approximately 8.5 dpc in the ventrolateral regions of the diencephalon and continued until 9.5 dpc (14). Then, *mSim2* expression became evident in the zona limitans intrathalamica and was accompanied by expression in several other limited regions, such as the mesenchyme of the branchial arches and that of the forelimbs in 11.0-dpc embryos (14). Later, *mSim2* mRNA also became detectable in the rib primordia in 11.5-dpc embryos (43).

Expression of the two *mSim* genes and *Shh* in the embryonal brain. To determine the precise regions of expression of the two *mSim* mRNAs, we used *Shh* cDNA as a reference probe, because *Shh* mRNA expression has been investigated in detail

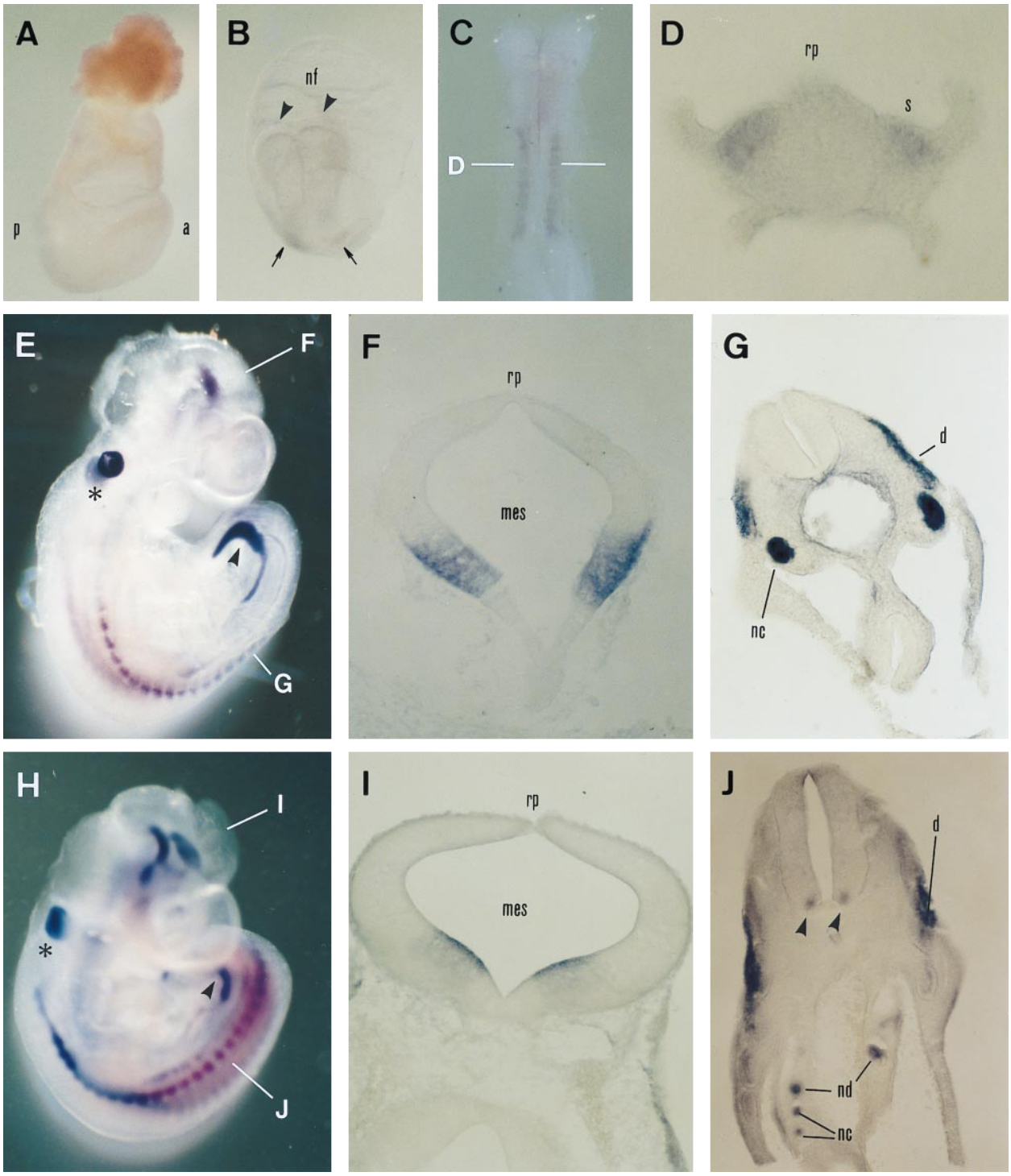


FIG. 8. Whole-mount in situ hybridization on mouse embryos using an *mSim1* cRNA probe. (A) Lateral view of a 7.5-dpc embryo. No significant signal is seen. (B) Lateral view of an 8.0-dpc embryo. Arrows indicate hybridization signals in the presomitic mesoderm. (C) Dorsal view of an 8.5-dpc embryo. D, position of the cross-section shown in panel D. (D) Transverse section of an 8.5-dpc embryo. Uniform expression of *mSim1* mRNA is observed in the presomitic mesoderm. (E) Lateral view of a 9.5-dpc embryo. F and G, positions of cross-sections at the level of the rostral mesencephalon and middle portion between the forelimb and hindlimb. Note that *mSim1* is expressed in the rostral mesencephalon. Arrowhead, hybridization signal in the hindgut; *, nonspecific signal in the optic vesicle. (F) Transverse section at the level of the rostral mesencephalon. Note that *mSim1* is expressed in the ventrolateral region but not in the ventral midline of the mesencephalon. (G) Transverse section at the middle portion between the forelimb and hindlimb. Note that *mSim1* is expressed in the dermatome and nephrogenic cord. (H) Lateral view of a 10.5-dpc embryo. I and J, positions of transverse sections; arrowhead, hybridization signal in the hindgut; *, nonspecific signal in the optic vesicle. (I) Transverse section of the rostral mesencephalon. Note that *mSim1* is intensely expressed in the neuroephithelium of the ventrolateral portion of the mesencephalon. (J) Transverse section at the middle portion between forelimbs and hindlimbs. Note hybridization signals in the dermatome. Arrowheads, signals in the ventrolateral domain of caudal neural tube between the floor plate and motor neurons. a, anterior; d, dermatome; mes, mesencephalon; nc, nephrogenic cord; nd, nephrogenic duct; nf, neural fold; p, posterior; rp, roof plate; s, somite.

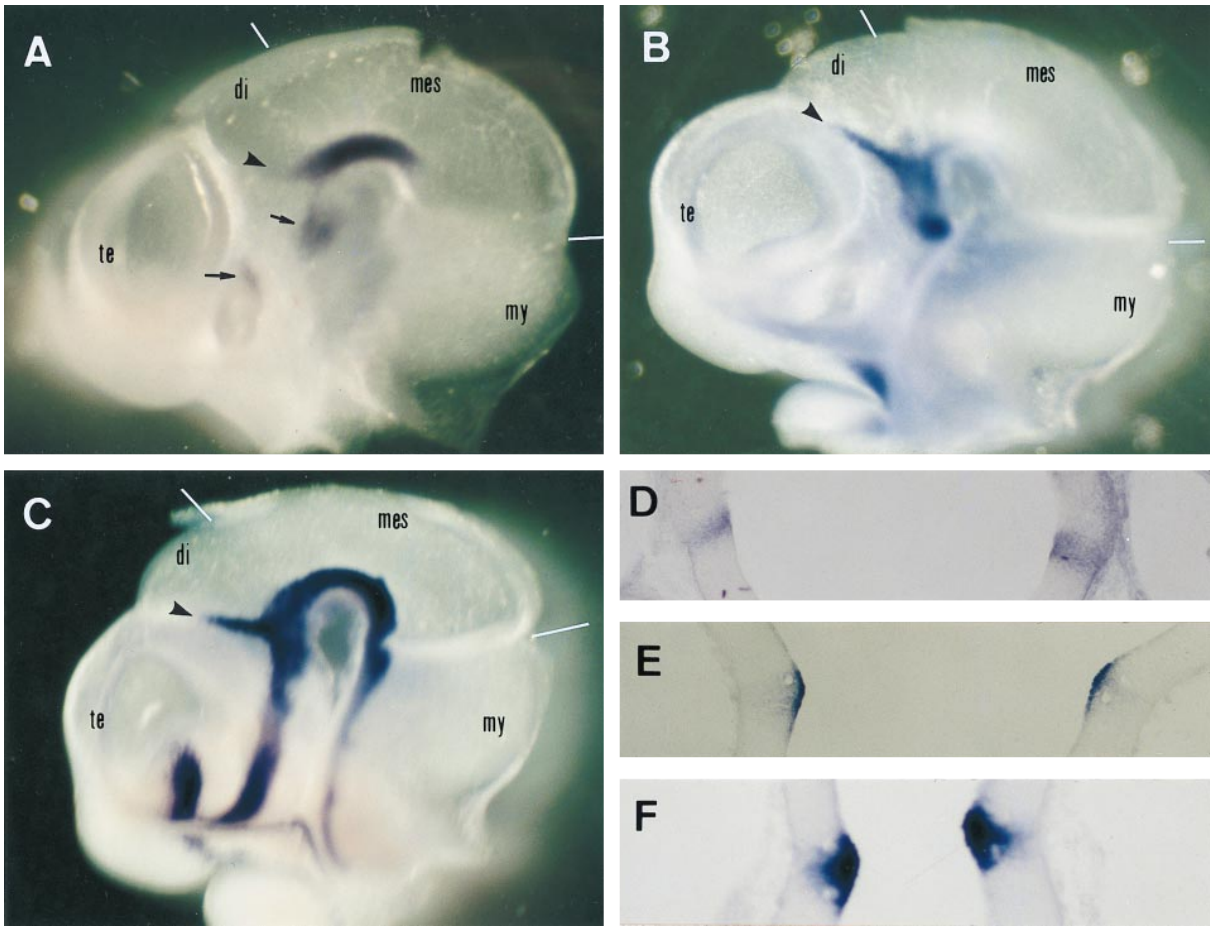


FIG. 9. Expression patterns of *mSim1*, *mSim2*, and *Shh* in the embryonal brain. Lateral views of 10.5-dpc embryos hybridized with an *mSim1* cRNA probe (A), *mSim2* (B), or *Shh* (C); coronal sections across the zona limitans intrathalamica of diencephalon of embryos hybridized with cRNAs of *mSim1* (D), *mSim2* (E), and *Shh* (F). Arrowheads, expression in the zona limitans intrathalamica; arrows, weak but significant hybridization signals of mSim1 seen in the ventral thalamus; white lines, diencephalon-mesencephalon and mesencephalon-myelencephalon boundaries; di, diencephalon; mes, mesencephalon, my, myelencephalon; tel, telencephalon.

elsewhere (57), and its distribution especially in the brain seemed to be closely associated with those of the two *mSim* mRNAs. Thus, we performed whole-mount in situ hybridization with the 10.5-dpc embryos using *Shh* cRNA as a probe and dissected the stained embryos (Fig. 9). In the 10.5-dpc embryo brain, *Shh* was expressed ventrolaterally in the diencephalon with the dorsal extension to the zona limitans intrathalamica and at the ventral midline of the mesencephalon, myelencephalon, and hindbrain (11, 57) (Fig. 9C). On the other hand, at the same stage, *mSim1* expression was observed in the ventrolateral side of the rostral mesencephalon and the caudal diencephalon and extended to the zona limitans intrathalamica (Fig. 9B). *mSim2* was also expressed at the ventrolateral domain of the diencephalon with dorsal extension to the zona limitans intrathalamica (14) (Fig. 9B). Basically, the distribution of expression of the two *mSim* mRNAs was within the regions in which *Shh* was expressed. Since accumulating evidence shows that *Shh* mediates signal functions governing pattern formation of vertebrate CNS (11, 15, 25, 55), limbs (54), and fins (31), it is of much interest to investigate whether expression of the two *mSim* mRNAs is under the control of *Shh.*

DISCUSSION

Isolation of *mSim1*, a new member of the mouse *sim* gene family. We have identified a cDNA for a novel member of mouse *Sim* homologs, termed *mSim1*, in addition to the previously identified *mSim2* (14). Comparison of amino acid sequences among mSim1 and other bHLH-PAS proteins thus far reported revealed that the bHLH-PAS domains participating in DNA sequence recognition and protein-protein interaction are highly similar among mSim1, mSim2, and dSim. In particular, the basic sequences immediately preceding the N terminus of the HLH domains which are thought to be directly involved in recognizing the nucleotide sequence show perfect identity, except for one conservative Arg-to-Lys alteration. In contrast, the sequences of the carboxy-terminal halves are variable and do not have any obviously conserved region in common. Recently, *mSim1* and *mSim2* cDNAs have been reported by another group (16). Although there are several nucleotide differences between their and our sequences, the two *mSim1* cDNAs are considered to derive partly from polymorphism and partly from sequencing errors. Concerning *mSim2*, it is suggested from the arrangement of variable and common sequences that the two cDNA sequences are most probably de-

rived from a single gene by alternative splicing, although this remains unclarified experimentally.

As reported previously (22), the bHLH-PAS proteins can be classified into two groups by degree of sequence similarity, namely the Arnt and AhR groups (see also Fig. 2). The Sim group clearly belongs to AhR (Fig. 2). The dimerization properties of the two mSim proteins also support this classification. The two mSim proteins were shown to interact only with Arnts, the other group of the bHLH-PAS family, to form a dimer, an activated form for a transcription factor, but not with themselves or a member within the same group (Table 1). This is in contrast to the Arnt group proteins, which dimerize with a member of the same group as well as with one of the other (22, 64) (Table 1). It would be interesting to see the molecular basis for compatibility in the dimer formation of the bHLH-PAS proteins.

Transcriptional activity of the two mSims. All of the bHLH-PAS transcription factors so far identified act as potent transcriptional activators. These factors are Arnt, Arnt2, AhR, HIF1α, and dSim, and they have their transcription activation domains in their carboxy-terminal halves (17, 22, 28, 32, 65, 72). In the cases of mSim1 and mSim2, however, no significant transcription-enhancing activity was observed with the transient DNA transfection system (Fig. 5). This is in contrast to the cases of Arnt, Arnt2, and AhR, which show marked transcription-enhancing activity in the DNA transfection system (22, 65). dSim was also found to function as a transcriptional activator in this system (Fig. 5). When cotransfected with Arnt, dSim further enhanced the CAT expression activated by Arnt. On the other hand, both mSim1 and mSim2 exhibited a dominant repressive effect on the transcription activated by Arnt (Fig. 6). Since each of the two mSim proteins forms a heterodimer with Arnt, it is suggested that the mSim proteins suppress Arnt transcription activity by dimer formation. When their carboxy-terminal halves were deleted, the truncated mSims, which still retained their dimer-forming ability, lost their repressor activity, suggesting that the repressor activity is localized in the carboxy-terminal halves of the two mSim proteins. Since all of the bHLH-PAS proteins reported to date, except for Per, show transcriptional activation activity as described elsewhere (17, 22, 28, 32, 65, 72), the two mSim proteins are the first examples of bHLH-PAS factors with a transrepressor activity, although it was suggested from genetic data that dSim can function as a transcriptional repressor. It remains to be seen whether dSim directly works on the genes as a transcriptional repressor (7, 41). Per has only the PAS domain and is considered another type of inhibitor of the bHLH-PAS proteins. Because of lack of the bHLH domain, it was suggested that dimer formation with the Per protein results in a complex without DNA-binding activity and sequesters the partner molecule with transcription-enhancing activities (35, 66). Analogous types of inhibition were already reported with IκB for NF-κB (3) and Id for MyoD (69).

This transcriptional property of the two mSims is reminiscent of the role of Mad and Mxi1, which are bHLH-Zip proteins involved in the regulation of Myc function, although it remains to be seen whether an antagonistic form of mSim really exists in vivo. Myc acts as a positive transcriptional regulator in a complex with Max and is antagonized by Mad or Mxi1 for dimerization with Max. Thereby, Mad and Mxi1 interfere with Myc function either by sequestering Max or by direct competition for the DNA target (1, 76). Recently, corepressors have been found to mediate the repressive effect of Mad (2, 62). It is interesting to investigate how mSims occlude the transcriptional activity of Arnt and whether there is a corepressor mediating their repressive effect.

Although we have not been able to detect any similarity in the two mSims with previously identified transrepressors, characteristic regions rich in alanine and proline residues and serine and threonine residues were found in mSim2 (21% alanine and 18% proline in aa 514 to 657) and mSim1 (15% serine in aa 332 to 765). Several transrepressors have been reported to have such regions rich in proline (36, 37), alanine (33, 63), and serine/threonine residues (18).

Expression of the two *mSim* mRNAs and their possible functions. In sharp contrast to the ubiquitous expression of Arnt mRNA (14, 22), RNA blot analysis of various tissues of adult mice revealed that the two *mSim* mRNAs were expressed in restricted tissues such as muscle, kidney, and lung tissues for *mSim2* and only in kidney and muscle tissues for *mSim1* (Fig. 7). It is interesting to investigate how the two mSims are expressed in mammalian embryogenesis, because dSim is known to work as an essential factor in neurogenesis of *Drosophila* species. In whole-mount in situ hybridization experiments, *mSim1* expression started in presomitic mesoderm at 8.0 dpc, a very early stage of somitogenesis, and was propagated to all somites as the somite differentiated. As the somites differentiated, *mSim1* expression was restricted to the dermatomes. By 9.5 dpc, the *mSim1* mRNA became detectable in the hindgut and an anterior region of the mesencephalon. Transverse sections revealed that *mSim1* expression was limited in the ventrolateral region of the rostral mesencephalon and was abundantly detectable in nephrogenic cords and dermatomes. While *mSim1* continued to be expressed in these organs at least up to 10.5 dpc, *mSim1* expression became prominent in the ventrolateral region of the caudal neural tube, the nephrogenic cords, and the mesonephric ducts at 10.5 dpc. Continuous expression of *mSim1* through somitogenesis, nephrogenesis, and neurogenesis from their very beginnings may suggest its importance in these processes. On the other hand, *mSim2* expression started in the ventrolateral region of the diencephalon in 8.5-dpc embryos (14). While its increased level of expression was observed in that region of the diencephalon, expression of *mSim2* became detectable in the forelimbs and the first and second branchial arches of 10.5-dpc embryos. In addition to the ventrolateral region of the diencephalon, *mSim2* expression became evident in the zona limitans intrathalamica of 10.5-dpc embryos, in which *mSim1* mRNA started to coexpress (14) (Fig. 9). The region of expression of the two *mSim*s in the midbrain was confirmed by using *Shh* mRNA expression as a standard. Interestingly, *Shh* expression always included those regions in which the two *mSim* mRNAs were expressed. Although expression of *Shh* preceded expression of the *mSim*s, the latter seem to be expressed only in portions of the regions in which *Shh* protein existed, such as the ventrolateral domain of the diencephalon and mesencephalon and the ventrolateral region of the spinal cord (15, 55). However, we could not detect any *mSim* gene expression in the floor plate, in sharp contrast to uniform expression of *dSim* in the mesectoderm throughout an anterior-posterior axis formed between the mesoderm and neuroectoderm (47). It would be interesting to investigate whether another mouse homolog of *dSim* works as a transcriptional activator in the floor plate and competes with these mSims for regulation of gene expression in these organs. The *Shh* product is known to be a secretory protein and is considered to play a role in signaling centers that regulate the polarity of the CNS and limbs (11, 31, 54). It would also be of interest to investigate whether there is any regulatory link between the expression of *Shh* and that of the two *mSim*s. It is suggested that during somitogenesis, a diffusible factor, BMP4, is required for acti-

vation of chicken *Sim* expression in the specification of the somite (51).

From the chromosomal localizations of the two mSims and the degree of sequence similarity, it is concluded that mSim2 is an ortholog of hSIM2, whose gene is mapped in the DSCR of chromosome 21 and is considered a candidate for the gene responsible for the congenital disease. *mSim2* is expressed in the hypothalamus, ventral thalamus, the branchial arches, and the limbs in the course of embryogenesis (see references 14, 16, and 45 and this paper). These regions for *mSim2* expression in mouse embryos are apparently related to those in which the major symptoms of Down's syndrome, namely, mental retardation and dysmorphic face and limbs, are observed. More work is expected to clarify whether and how an overdose of the *Sim2* gene is involved in the incidence of Down's syndrome.

ACKNOWLEDGMENTS

We thank H. Nakafuku (NAIST, Nara, Japan), H. Nakamura (Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan), and N. Yokotani (Osaka City University, Osaka, Japan) for helpful discussions; S. Fields (State University of New York) for pGBT9 and pGAD424 plasmids; S. Noji (Tokushima University, Tokushima, Japan) and A. P. McMahon (Harvard University) for the *Shh* plasmid; and H. Abe for help in transfection experiments.

This work was supported in part by grants-in-aid for scientific research on priority areas and for scientific research, by the International Scientific Research Program from the Ministry of Education, Culture, Sport, and Science of Japan, and by funds from Sankyo Co. M.E. is a research fellow of the Japan Society for the Promotion of Science.

REFERENCES

1. **Ayer, D. E., L. Kretzner, and R. N. Eisenman.** 1993. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. Cell **72:**211–222.
2. **Ayer, D. E., Q. E. Lawrence, and R. N. Eisenman.** 1995. Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. Cell **80:**767–776.
3. **Baeuerle, P. A., and T. Henkel.** 1994. Function and activation of NF-kappa B in the immune system. Annu. Rev. Immnol. **12:**141–179.
4. **Bartel, P. L., C.-T. Chen, R. Sternglanz, and S. Fields.** 1994. Using the two-hybrid system to detect protein-protein interactions, p. 153–179. *In* D. A. Hartley (ed.), Cellular interactions in development: a practical approach. Oxford University Press, New York.
5. **Burbach, K. M., A. Poland, and C. A. Bradfield.** 1992. Cloning of the AH-receptor cDNA reveals a novel ligand activated transcription factor. Proc. Natl. Acad. Sci. USA **89:**8185–8189.
6. **Campos-Ortega, J. A., and V. Hartenstein.** 1985. The embryonic development of Drosophila melanogaster. Springer-Verlag, Heidelberg, Germany.
7. **Chang, Z., D. Price, S. Bockheim, M. J. Boedigheimer, R. Smith, and A. Laughon.** 1993. Molecular and genetic characterization of the *Drosophila tartan* gene. Dev. Biol. **160:**315–332.
8. **Chen, H., R. Chrast, C. Rossier, A. Gos, S. E. Antonarakisis, J. Kudoh, A. Yamaki, N. Shindoh, H. Maeda, S. Minoshima, and N. Shimizu.** 1995. *single-minded* and Down syndrome? Nat. Genet. **10:**9–10.
9. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5299.
10. **Crews, S. T., J. B. Thomas, and C. S. Goodman.** 1988. The *Drosophila single-minded* gene encodes a nuclear protein with sequence similarity to the per gene product. Cell **52:**143–151.
11. **Echelard, Y., D. J. Epstein, B. St-Jacques, L. Shen, J. Mohler, J. A. McMahon, and A. P. McMahon.** 1993. *sonic hedgehog*, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell **75:**1417–1430.
12. **Ema, M., N. Ohe, M. Suzuki, J. Mimura, K. Sogawa, S. Ikawa, and Y. Fujii-Kuriyama.** 1994. Dioxin-binding activities of polymorphic forms of mouse and human arylhydrocarbon receptors. J. Biol. Chem. **269:**27337–27343.
13. **Ema, M., K. Sogawa, N. Watanabe, Y. Chujoh, N. Matsushita, O. Gotoh, Y. Funae, and Y. Fujii-Kuriyama.** 1992. cDNA cloning and structure of mouse putative Ah receptor. Biochem. Biophys. Res. Commun. **184:**246–253.
14. **Ema, M., M. Suzuki, M. Morita, K. Hirose, K. Sogawa, Y. Matsuda, O. Gotoh, Y. Saijoh, H. Fujii, H. Hamada, and Y. Fujii-Kuriyama.** 1996. cDNA cloning of a murine homologue of *Drosophila single-minded*, its mRNA expression in mouse development and chromosomal localization. Biochem. Biophys. Res. Commun. **218:**588–594.
15. **Ericson, J., J. Muhr, M. Placzek, T. Lints, T. M. Jessel, and T. Edlund.** 1995. *sonic hedgehog* induces the differentiation of ventral forebrain neurons: a common signal for ventral patterning within the neural tube. Cell **81:**745–756.
16. **Fan, C.-M., E. Kuwana, A. Bulfone, C. F. Fletcher, N. G. Copeland, N. A. Jenkins, S. Crews, S. Martinez, L. Puelles, J. L. R. Rubenstein, and M. Tessier-Lavigne.** 1996. Expression patterns of two murine homologs of *Drosophila single-minded* suggest possible roles in embryonic patterning and in the pathogenesis of Down syndrome. Mol. Cell. Neurosci. **7:**1–16.
17. **Franks, R. G., and S. T. Crews.** 1994. Transcriptional activation domains of the *single-minded* bHLH protein are required for CNS midline cell development. Mech. Dev. **45:**269–277.
18. **Gashler, A. L., S. Swaminathan, and V. P. Sukhatme.** 1993. A novel repression module, an extensive activation domain, and a bipartite nuclear localization signal defined in the immediately-early transcription factor Egr-1. Mol. Cell. Biol. **13:**4556–4571.
19. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
20. **Gotoh, O.** 1995. A weighting system and algorithm for aligning many phylogenetically related sequences. Comput. Appl. Biosci. **11:**543–551.
21. **Han, K., and J. L. Manley.** 1993. Transcriptional repression by the *Drosophila* Even-skipped protein: definition of a minimal repression domain. Genes Dev. **7:**491–503.
22. **Hirose, K., M. Morita, M. Ema, J. Mimura, H. Hamada, H. Fujii, Y. Saijoh, O. Gotoh, K. Sogawa, and Y. Fujii-Kuriyama.** 1996. cDNA cloning and tissue-specific expression of a novel bHLH/PAS factor (Arnt2) with close sequence similarity to AhR nuclear translocator (Arnt). Mol. Cell. Biol. **16:**1706–1713.
23. **Hoffman, E. C., H. Reyes, F.-F. Chu, F. Sander, L. H. Conley, B. H. Brooks, and O. Hankinson.** 1991. Cloning of a factor required for activity of the Ah (dioxin) receptor. Science **252:**954–958.
24. **Hosoya, T., and Y. Fujii-Kuriyama.** Unpublished result.
25. **Hynes, M., K. Poulsen, M. Tessier-Lavigne, and A. Rosenthal.** 1995. Control of neuronal diversity by the floor plate: contact-mediated induction of midbrain dopaminergic neurons. Cell **80:**95–101.
26. **Isaac, D. D., and D. J. Andrew.** 1996. Tubulogenesis in Drosophila, a requirement for the trachealess gene product. Genes Dev. **10:**103–118.
27. **Jackson, R. R., T. A. Bargiello, S.-H. Sun, and M. W. Young.** 1986. Product of *per* locus of Drosophila shares homology with proteoglycans. Nature (London) **320:**185–188.
28. **Jain, S., K. M. Dolwick, J. V. Schmidt, and C. A. Bradfield.** 1994. Potent transactivation domains of the Ah receptor and the Ah receptor nuclear translocator map to their carboxyl termini. J. Biol. Chem. **269:**31518–31524.
29. **Klämbt, C., J. R. Jacobs, and C. S. Goodman.** 1991. The midline of the *Drosophila* central nervous system: a model for the genetic analysis of cell fate, cell migration, and growth cone guidance. Cell **64:**801–815.
30. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
31. **Krauss, S., J.-P. Conchodet, and P. W. Ingham.** 1993. A functionally conserved homolog of the Drosophila segment polarity gene *hh* is expressed in tissues with polarizing activity in zebrafish embryos. Cell **75:**1431–1444.
32. **Li, H., L. Dong, and J. P. Whitlock, Jr.** 1994. Transcriptional activation function of the mouse Ah receptor nuclear translocator. J. Biol. Chem. **269:**28098–28105.
33. **Licht, J. D., M. J. Grossel, J. Figge, and U. M. Hansen.** 1990. Drosophila Krüppel protein is a transcriptional repressor. Nature (London) **346:**76–79.
34. **Lillie, J. W., and M. R. Green.** 1989. Transcriptional activation by the adenovirus E1a protein. Nature (London) **338:**39–44.
35. **Lindebro, M. C., L. Poellinger, and M. L. Whitelaw.** 1995. Protein-protein interaction via PAS domain: role of the PAS domain in positive and negative regulation of the bHLH/PAS dioxin receptor-Arnt transcription factor complex. EMBO J. **14:**3528–3539.
36. **Madden, S. L., D. M. Cook, J. F. Morris, A. Gashler, V. P. Sukhatme, and F. J. Rausher III.** 1991. Transcriptional repression mediated by the WT1 Wilms tumor gene product. Science **253:**1550–1553.
37. **Madden, S. L., D. M. Cook, and F. J. Rausher III.** 1993. A structure-function analysis of transcriptional repression mediated by the WT1, Wilms' tumor suppressor protein. Oncogene **8:**1713–1720.
38. **Matsuda, Y., and V. M. Chapman.** 1995. Application of fluorescence in situ hybridization in genome analysis of the mouse. Electrophoresis **16:**261–272.
39. **Matsuda, Y., Y.-H. Harada, S. Natsume-Sakai, K. Lee, T. Shiomi, and V. M. Chapman.** 1992. Location of the mouse complement factor H gene (cfh) by FISH analysis and replication R-banding. Cytogenet. Cell Genet. **61:**282–285.
40. **Matsushita, N., K. Sogawa, M. Ema, A. Yoshida, and Y. Fujii-Kuriyama.** 1993. A factor binding to the xenobiotic responsive element (XRE) of P-4501A1 gene consists of at least two helix-loop-helix proteins, Ah receptor and Arnt. J. Biol. Chem. **268:**21002–21006.
41. **Mellerick, D. M., and M. Nirenberg.** 1995. Dorsal-ventral patterning genes restrict *NK-2* homeobox gene expression to the ventral half of the central

nervous system of *Drosophila* embryos. Dev. Biol. **171:**306–316.
42. **Mizushima, S., and S. Nagata.** 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. **18:**5322.
43. **Morita, M., and Y. Fujii-Kuriyama.** Unpublished results.
44. **Muenke, M., L. J. Bone, H. F. Mitchell, I. Hart, K. Walton, K. Hall-Johnson, E. F. Ippel, J. Dietz-Band, K. Kvaloy, C.-H. Fan, M. Tessier-Lavigne, and D. Patterson.** 1995. Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of SIM2 as a candidate gene for holoprosencephaly, and mapping of SIM2 to a region of chromosome 21 important for Down syndrome. Am. J. Hum. Genet. **57:**1074–1079.
45. **Nahmane, N., G. Charron, C. Lopes, M.-L. Yaspo, C. Maunoury, L. Decorte, P.-M. Sinet, M. Bloch, and J.-M. Delabar.** 1995. Down syndrome-critical region contains a gene homologous to *Drosophila sim* expressed during rat and human central nervous system development. Proc. Natl. Acad. Sci. USA **92:**9191–9195.
46. **Nambu, J. R., R. G. Franks, S. Fu, and S. T. Crews.** 1990. The *single-minded* gene of *Drosophila* is required for the expression of genes important for the development of CNS midline development. Cell **63:**63–75.
47. **Nambu, J. R., J. O. Lewis, and S. T. Crews.** 1993. The development and function of the *Drosophila* CNS midline cells. Comp. Biochem. Physiol. **104:**399–409.
48. **Nambu, J. R., J. O. Lewis, K. O. Wharton, Jr., and S. T. Crews.** 1991. The *Drosophila single-minded* gene encodes a helix-loop-helix protein which acts as a master regulator of CNS midline. Cell **67:**1157–1167.
49. **Numayama-Tsuruta, K., and Y. Fujii-Kuriyama.** Unpublished data.
50. **Nüsslein-Volhard, C., E. Wieschaus, and H. Kluding.** 1984. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. Roux's Arch. Dev. Biol. **193:**267–282.
51. **Pourquie, O., C.-M. Fan, M. Coltey, E. Hirsinger, Y, Watanabe, C. Breant, P. Francis-West, P. Brickell, M. Tessier-Lavigne, and N. M. Le Douarin.** 1996. Lateral and axial signals involved in avian somite patterning: a role for BMP4. Cell **84:**461–472.
52. **Reddy, P., A. C. Jaquier, N. Aborich, G. Petersen, and M. Roshbash.** 1986. The period clock locus of *D. melanogaster* codes for a proteoglycan. Cell **46:**53–61.
53. **Reyes, H., S. Reisz-Porszasz, and O. Hankinson.** 1992. Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. Science **256:**1193–1195.
54. **Riddle, R., R. L. Johnson, E. Laufer, and C. Tabin.** 1993. Sonic hedgehog mediates the ZPA of polarizing activity. Cell **75:**1401–1416.
55. **Roelink, H., A. Augsberger, J. Heemskerk, V. Korzh, S. Norlin, A. Ruiz i Altaba, Y. Tanabe, M. Placzek, T. T. Edlund, T. M. Jessel, and J. Dodd.** 1994. Floor plate and motor neuron induction by *vhh-1*, a vertebrate homolog of hedgehog expressed by the notochord. Cell **76:**761–775.
56. **Rothberg, J. M., J. R. Jacobs, C. S. Goodman, and S. Artavanis-Tsakonas.** 1990. *slit*: an extracellular protein necessary for development of midline glia and commisural axon pathways contains both EGF and LRR domains. Genes Dev. **4:**2169–2187.
57. **Rubenstein, J. L. R., S. Martinez, K. Shimamura, and L. Puelles.** 1994. The embryonic vertebrate forebrain: the prosomeric model. Science **266:**578–580.
58. **Sadowski, I., and M. Ptashne.** 1989. A vector for expressing GAL4 (1-147) fusions in mammalian cells. Nucleic Acids Res. **17:**7539.
59. **Saitou, N., and M. Nei.** 1987. The neighbour-joining method: a new method for reconstructing phylogenic trees. Mol. Biol. Evol. **4:**406–425.
60. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
61. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
62. **Schreiber-Agus, N., L. Chin, K. Chen, R. Torres, G. Rao, P. Guida, A. I. Skoultchi, and R. A. DePinho.** 1995. An amino terminal domain of Mxi mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor SIN3. Cell **80:**1777–1786.
63. **Shi, Y., E. Seto, L. S. Chang, and T. Shenk.** 1991. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. Cell **67:**377–388.
64. **Sogawa, K., R. Nakano, A. Kobayashi, Y. Kikuchi, N. Ohe, N. Matsushita, and Y. Fujii-Kuriyama.** 1995. Possible function of Ah receptor nuclear translocator (Arnt) homodimer in transcriptional regulation. Proc. Natl. Acad. Sci. USA **92:**1936–1940.
65. **Sogawa, K., K. Iwabuchi, H. Abe, and Y. Fujii-Kuriyama.** 1995. Transcriptional activation domain of the Ah receptor and Arnt (Ah receptor nuclear translocator). J. Cancer Res. Clin. **121:**612–620.
66. **Takahashi, J. S.** 1992. Circadian clock genes are ticking. Science **258:**238–240.
67. **Thomas, J. B., S. T. Crews, and C. S. Goodman.** 1988. Molecular genetics of the single-minded locus: a gene involved in the development of the Drosophila nervous system. Cell **52:**133–141.
68. **Wang, G. L., B.-H. Jiang, E. A. Rue, and G. L. Semenza.** 1995. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc. Natl. Acad. Sci. USA **92:**5510–5514.
69. **Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar.** 1991. The myoD gene family: nodal point during specification of the muscle cell lineage. Science **251:**761–766.
70. **Wharton, K. A., Jr., R. G. Franks, Y. Kasai, and S. T. Crews.** 1994. Control of CNS midline transcription by asymmetric E-box-like elements: similarity to xenobiotic responsive regulation. Development **120:**3563–3569.
71. **Whitelaw, M. L., I. Pongratz, A. Wilhelmsson, J.-Å. Gustafsson, and L. Poellinger.** 1993. Ligand-dependent recruitment of the Arnt coregulator determines DNA recognition by the dioxin receptor. Mol. Cell. Biol. **13:**2504–2514.
72. **Whitelaw, M. L., J.-Å. Gustafsson, and L. Poellinger.** 1994. Identification of transactivation and repression functions of the dioxin receptor and its basic helix-loop-helix/PAS partner factor Arnt: inducible versus constitutive modes of regulation. Mol. Cell. Biol. **14:**8343–8355.
73. **Wieschaus, E., C. Nüsslein-Volhard, and G. Jürgens.** 1984. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster.* III. Zygotic loci on the second chromosome. Roux's Arch. Dev. Biol. **193:**283–295.
74. **Wilk, R., I. Weizman, and B.-Z. Shilo.** 1996. *trachealess* encodes a bHLH-PAS protein that is an inducer of tracheal cell fates in Drosophila. Genes Dev. **10:**93–102.
75. **Wilkinson, D. G.** 1992. In situ hybridization: a practical approach. Oxford University Press, London.
76. **Zervos, A. S. J. Gyuris, and R. Brent.** 1993. Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. Cell **72:**223–232.