SOX GENES: ARCHITECTS OF DEVELOPMENT

Heather M. Prior and Michael A. Walter
Ocular Genetics Research Group, University of Alberta, Edmonton, Alberta, Canada

Development in higher organisms involves complex genetic regulation at the molecular level. The emerging picture of development control includes several families of master regulatory genes which can affect the expression of downstream target genes in developmental cascade pathways. One new family of such development regulators is the SOX gene family. The SOX genes are named for a shared motif called the SRY box, a region homologous to the DNA-binding domain of SRY, the mammalian sex determining gene. Like SRY, SOX genes play important roles in chordate development. At least a dozen human SOX genes have been identified and partially characterized (Tables 1 and 2). Mutations in SOX9 have recently been linked to campomelic dysplasia and autosomal sex reversal, and other SOX genes may also be associated with human disease.

OVERVIEW: DEVELOPMENT CONTROL BY ARCHITECTURAL TRANSCRIPTION FACTORS

The process of development includes the increasing differentiation of pluripotent progenitor cells. Since the pattern of transcribed proteins within each cell ultimately determines its identity in the context of tissue, organ, and organism, the regulation of transcription is at the foundation of any development control pathway.

Basic gene transcription is regulated by transcription factors, proteins which bind specific DNA sequences called promoters or enhancers located within and around the gene (1,2). In addition to a DNA-binding domain, most transcription factors have a distinct activation domain which interacts with other proteins to coordinately control gene transcription. A simplified model of the regulation of downstream target genes by transcription factors is shown in Fig. 1.

Large "families" of transcription factors which are involved in the control of developmental pathways have been identified. These gene families encode proteins characterized by well-defined DNA-binding motifs, such as zinc-fingers or homeobox domains, typically highly conserved across species (3,4). Many of these genes were first studied in Drosophila. An example is the HOM/HOX (homeobox-containing) gene complexes which play a role in early pattern formation in the Drosophila larva and have also been implicated in pattern formation in early vertebrate development (5). Within the PAX (*pai*red box) family of genes (6), PAX6 is a master eye development control gene in organisms as diverse as insects, cephalopods, and humans. Interestingly, mutations in this gene cause either aniridia or autosomal dominant keratitis, two congenital eye defects in humans (7–9). Mutations in other PAX genes also cause human diseases, including coloboma and renal anomalies (PAX2), and Waardenburg syndrome (PAX3) (10,11). While SOX genes have not yet been as well characterized, they appear to also play important roles in early development and to have interesting associations to human disease.

The SOX genes belong to a large group of genes in which the DNA-binding domain is called a high mobility group (HMG) box (12). Two basic types of HMG-class proteins can be delineated. One group is characterized by proteins containing multiple HMG boxes, having a general affinity for binding DNA independent of its sequence. This group includes the HMG-1 protein, ubiquitous binding factor (UBF), and mitochondrial transcription factor 1 (MT-TF1). The second category of HMG-class proteins consists of those with a single HMG box and that bind DNA in a highly sequence-specific manner.

Address correspondence and reprint requests to: Michael A. Walter, Ocular Genetics Laboratory, 671 Heritage Medical Research Center, University of Alberta, Edmonton, Alberta, T6G 2S2, Canada.

Molecular Medicine, Volume 2, Number 4, July 1996 405-412

TABLE 1. Accession numbers of human and mouse Sox sequences

	Accession Number	Sequence Data
Human Gene		
SRY	L08063	Complete
	L10102	
	L10101	
	X53772	
SOX2	Z31560	Partial
SOX3	X71135	Complete
SOX4	X70683	Complete
SOX5	X65662	HMG box
SOX6	X65663	HMG box
SOX8	X65664	HMG box
SOX9	Z46629	Complete
	S74504–S74506	
SOX10	X65666	HMG box
SOX11	X73038	HMG box
SOX12	X73039	HMG box
Mouse Gene		
Sry	X55491	Complete
Sox1	X94126	Complete
Sox2	X94127	Complete
	U31967	
Sox3	X94125	Complete
Sox4	X70298	Complete
Sox5	X65657	Complete
	S45834	
Sox6	U32614	Complete
Sox7	X65660	HMG box
Sox8	Z18957	HMG box
Sox9	Z18958	HMG box
Sox10	Z18959	HMG box
Sox11	Z18960	HMG box
Sox12	Z18961	HMG box
Sox13	Z18962	HMG box
Sox14	Z18963	HMG box
Sox15	X70909	HMG box
Sox16	L29084	HMG box
Sox17	L29085	HMG box
Sox18	L35032	Complete

Some members of this group are proteins encoded by the yeast mating type genes *matMc* and *mat*-A1, the white cell regulatory genes T cell factor–1 (TCF-1) and the lymphocyte enhancer factor–1 (LEF-1) (13,14), as well as the SOX and SRY genes. Genes in this category share about 25% sequence identity over the 79 amino acid HMG box, with little or no similarity outside this box. It has been directly demonstrated for SRY, SOX4, and SOX5 (15–17) that these proteins bind in the minor groove of DNA at the consensus sequence A/T A/T C A A A G. The binding induces a sharp bend of 80° to 135° in the DNA template (18,19), which in turn may act to bring different regulatory regions of the target gene into close proximity. As a result, various transcription factors bound to these regions would be able to interact to activate transcription. DNA-binding proteins with this "structural" mode of action have been appropriately called architectural transcription factors (20).

THE SRY GENE AND SEX DETERMINATION

The first SOX gene to be successfully cloned and characterized was the SRY gene (Fig. 2), and it remains the defining member of the family (22). The critical region on the Y chromosome was initially established by molecular analysis of the DNA of sex reversed patients, including XX males with portions of the Y chromosome translocated to one of the X chromosomes and XY females with deletions within the Y chromosome. Although a number of candidate genes initially looked promising, SRY was shown in 1990 to be the one necessary for male sex determination since mutations within its open reading frame were found in XY sex reversed patients (23,24). All the mutations identified fell within the HMG box region of the SRY gene. Subsequent experiments showing that XX transgenic mice carrying the SRY gene were phenotypically male (25) demonstrated that this gene was sufficient for male sex determination.

In male mice, the SRY protein is expressed in the bipotential genital ridge of the developing embryo (26). Outside the genital ridge, SRY mRNA levels are regulated by an unusual translational control system whereby a unique splicing event produces non-functional circular transcripts (27). The gonad in the absence of SRY develops, by default, as an ovary. During male development, SRY acts to induce MIS (Müllerian inhibiting substance) expression, which in turn causes regression of the Müllerian duct system (female), progression of the Wolffian ducts

TABLE 2. SOX genes

Group	Gene	Chromosome (Human)	Sequence Data (Human)[a]	Disease Association	Tissue
A	SRY	Y	+	XY sex reversal	Genital ridge
B	SOX1				
	SOX2	3q26.3–27	+		Embryo CNS
	SOX3	Xq26.2–27.2	+	Borjeson-Forssman-Lehmann (?)	Embryo CNS
	SOX14				
	SOX15				
	SOX19				
C	SOX4	6p	+		Lymphocytes
	SOX11		+		Embryo CNS
	SOX12		+		
	SOX20	17p13	+		
D	SOX5		+		Adult testis
	SOX6		+		Testis, embryo CNS
	SOX13				
E	SOX8		+		
	SOX9	17q24	+	Campomelic dysplasia, sex reversal	Adult testis
	SOX10		+		
F	SOX7				
	SOX17				
	SOX18				Muscle

[a]Complete human cDNA sequence data for SRY, SOX3, SOX4, and SOX9 only.

(male), and gonadal development as testis (28). The SRY product may also act to induce male steroidogenesis, which in turn leads to male external genitalia. Affinity studies to determine the precise downstream targets of SRY have shown that it probably binds to intervening intermediate factors in both the MIS and the male steroidogenesis pathways (28).

SOX GENES

The SOX (SRY-related HMG box) genes were initially identified through their homology to the HMG box of SRY. By definition, the DNA-binding domain of SOX genes is at least 60% similar or 50% identical to the 79 amino acid HMG box of the SRY gene. At least 19 SOX genes have been identified and divided into six groups, designated A–F (Table 2), according to the similarities of their HMG box regions (Fig. 3). SOX genes have been found in Drosophila and many vertebrates, including mouse, chicken, gull, frog, turtle, zebrafish, marsupials, and humans (12,29–35).

Although the majority of work has concentrated on the mouse Sox genes, several human

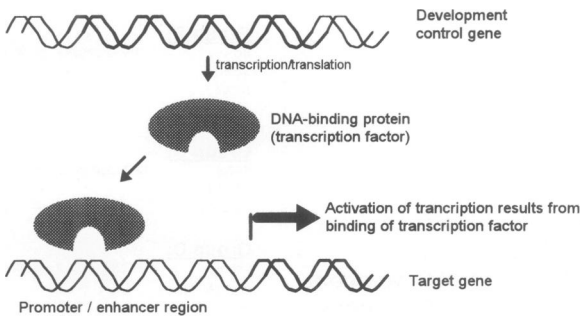

FIG. 1. Downstream regulation model

Development control genes encode transcription factors with DNA-binding domains. These factors bind to the promoter/enhancer regions of target genes and often interact with other proteins to influence transcription levels in the target genes.

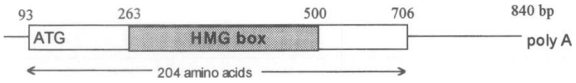

FIG. 2. Structure of the SRY gene

The SRY gene has a single open reading frame which contains a 79–amino acid HMG box (21).

genes have also been partially characterized. Apart from SRY located on the Y chromosome and SOX3 on the X chromosome, other human SOX genes are autosomal and those whose chromosomal location is known are scattered throughout the genome. Although SOX genes appear to be predominantly expressed in the developing testis and nervous system, additional studies are needed to determine whether this expression pattern has functional or evolutionary significance. Several SOX genes, including SRY, SOX3, and SOX4 (36–39), are single exon genes; SOX9, however, contains three exons (40,41). Full cDNA sequences have been reported for human SRY, SOX3, SOX4, and SOX9, and for mouse Sry, Sox1, Sox2, Sox3, Sox4, Sox6, and Sox18. Additional sequence data are primarily confined to the HMG box region (Table 2).

SRY is the sole member in its group (A) and is most similar to the group B genes. The best studied members of group B appear to function as architects of neuronal development. Sox1, Sox2, and Sox3 are expressed at high levels in the murine embryonic nervous system (42) and chicken Sox2 and Sox3 are expressed in the undifferentiated cells of the neural epithelium (34). Sox2 is expressed in developing eye tissues and a recent study implicates chicken Sox2 in the lens-specific regulation of the δ1-crystallin gene (43). Human SOX2 has been partially cloned and mapped to chromosome 3q26.3–27 (44). In addition, murine Sox2 has been found to complex with another transcription factor, Oct-3, to promote transcriptional activation of fibroblast growth factor 4 (FGF-4) in embryonic carcinoma cell lines (45). This type of complex transactivation represents an intriguing mechanism in developmental regulation.

SOX3 has the highest similarity to the HMG box of SRY, and recent work has suggested that SRY may have in fact originated as a homologue of the Sox3 gene (37,42). Since other genes on the Y chromosome are also similar to genes on the X chromosome, it suggests that the sex chromosomes may have originally been a homologous pair like the autosomes (46). In humans, SOX3 shows widespread expression in fetal tissues, including brain and spinal cord, as well as in some adult tissues (37). A patient with a deletion including the SOX3 gene has hemophilia and mental retardation. Although this patient has small testes, his male phenotype suggests that SOX3 is not necessary for testis formation. This gene is, however, a candidate for the Borjeson-Forssman-Lehmann syndrome, an X-linked condition which includes mental retardation, epilepsy, and hypogonadism, and which maps to the same region of Xq26–27 as the SOX3 gene (37).

In group C, Sox4 has been characterized more extensively than SOX20, Sox12, or Sox11 (47,48), although Sox11 has been shown to have neuronal specific expression patterns in chicken (34). Sox4 is expressed in T cells and pre–B lym-

FIG. 3. HMG box comparison of SOX genes by group

Amino acid sequnces are given for the HMG boxes of human (SOX) and mouse (Sox) genes. Amino acid identities are indicated by dashes.

```
Group A:
SRY      DRVKRPMNAF IVWSRDQRRK MALENPRMRN SEISKQLGYQ WKMLTEAEKW PFFQEAQKLQ AMHREKYPNY KYRPRRKAKM

Group B:
Sox-1    DRVKRPMNAF MVWSRGQRRK MAQENPKMHN SEISKRLGAE WKVMSEAEKR PFIDEAKRLR ALHMKEHPDY KYRPRRKTKT
SOX2     ---------- ---------- ---------- ---------- --LL--T--- ---------- ---------- ----------
SOX3     ---------- ---------- --L------- ---------D --LLTD---- ---------- -V----Y--- ----------
Sox-14             ---------- ---------- ---------- --LL------ -Y-------- -Q--E

Group C:
SOX4     GHIKRPMNAF MVWSQIERRK IMEQSPDMHN AEISKRLGKR WKLLKDSDKI PFIREAERLR LKHMADYPDY KYRPRKKVKS
SOX11              ----K----- ---------- ---------- --M----E-- ---------- ----------
Sox-12             ----H----- --D-W----- -------GR- -Q--Q--E-- --E------- ------

Group D:
Sox-5    PHIKRPMNAF MVWAKDERRK ILQAFPDMHN SNISKILGSR WKAMTNLEKQ PYYEEQARLS KQHLEKYPDY KYKPRPKRTC
Sox-6                 -R----- ---------- ---------- --S-S-Q-N- ---------- -I-----
Sox-13             ---------- ---------- -S-------- --S---Q--- ---------- R-----

Group E:
Sox-8              MVWAQAARRK LADQYPHYHN AELSKTLGKL WRLLSESEKR PFVEEAERLR VQHKKD
SOX9                  SRGQ--- M-QEN-KM-- S----R--AE -K--TDA--- --ID--K--- AV-M--Y
SOX10                 SR-Q--- M-QEN-KM-- S----R--AE -K--T----- --ID--K--- AM-N-EH

Group F:
Sox-7                 AKDERKR LAVQNPDLHN AELSKMLGKS WKALTLSQKR PYVDEAERLR LQHMQDY
Sox-17             MVW------- --Q------- ---------- -----TAE-- -F-E------ V-----
Sox-18   LRIRRPMNAF MVW------- --Q------- -V-------A --E-NTAE-- -F-E------ V--LR-HPNY KYRPRRKKQ
```

phocytes and is involved, together with TCF-1 and LEF-1, in controlling lymphocyte differentiation. Binding studies have shown that the Sox4 protein has an affinity for the DNA sequence AACAAAG, a motif found in the enhancer region of some T cell receptor genes. Sox4 also contains a serine-rich transactivation domain (separable from its DNA-binding domain), and thus it represents the first SOX gene to have the characteristic structure of a classical transactivator of transcription (49).

In group D, Sox5 and Sox6 are both expressed in adult mouse testis. Sox5 is exclusively expressed in post-meiotic round spermatids and may play a role in spermatogenesis (50). The DNA-binding domains of both Sox5 and Sox6 have affinity for the sequence AACAAT, and binding by Sox5 induces a sharp bend in the template DNA. Sox6 is also expressed in segments of the developing anterior nervous system, suggesting a possible role in CNS differentiation and growth, in addition to its proposed role in testis determination (51).

The best-characterized human SOX gene is SOX9, a member of the group E subclass. Mutations in this gene cause campomelic dysplasia (CPMD1) and autosomal sex reversal (SRA1) (40,41). CPMD1 is a rare congenital skeletal malformation syndrome characterized by bowing of the long bones and defects in cartilage formation. It is associated with autosomal sex reversal, and two-thirds of XY CPMD1 patients develop with female or ambiguous genitalia (52,53). The syndrome was localized to the distal portion of chromosome 17q by analysis of DNA from patients with chromosome translocations in this region (54). SOX9 was investigated as a candidate gene for this disease since the mouse Sox9 mapped to the homologous region and had been shown to have a primary role in skeletal formation (55,56). Mutations in SOX9 were detected in two-thirds of CPMD1 patients without chromosomal translocations, suggesting that the gene was involved in the disease phenotype; intriguingly, however, several CPMD1 translocation breakpoints mapped just outside of the SOX9 gene (40,41). The SOX9 gene codes for a protein with an HMG box DNA-binding domain and a putative activation domain containing proline and glutamine residues. The mutations in CPMD1 patients are predicted to result in loss of function alleles and an autosomal dominant mode of inheritance for the disease due to haploinsufficiency. Dosage sensitivity often plays a role in sex determination mechanisms, and further studies of SOX9 may indeed extend the current understanding into this area.

In group F, preliminary analyses have been carried out on the Sox17 and Sox18 genes in the mouse. The Sox18 protein, like Sox4, has a serine-rich region as a putative transactivation domain in addition to its HMG box. Sox18 expression is limited to smooth and striated muscle in the adult mouse (57).

FUTURE DIRECTIONS

Research into SOX genes is still in its preliminary stages. Most studies to date have focussed on isolating and sequencing various members of the family, along with studies of gene expression and DNA-binding affinities. This work has shown that members of the SOX gene family may be involved in many different aspects of development, including sex determination, testis formation, neuronal development, lymphocyte differentiation, and chondrogenesis.

Future investigations will focus on determining the molecular targets of SOX proteins in order to elucidate their mode of action more precisely. As human SOX genes are cloned, they will become positional candidates for various diseases and mutation analyses will help to correlate structural domains with function. Because SOX genes have been found in diverse organisms, model systems for studying human mutation and disease may be developed in other species. A very promising area that has just begun to be explored concerns the interactions of SOX proteins with one another and with other development control factors. Various SOX genes are thought to interact, including SRY with SOX9 and Sox5 with Sox6. Additional interactions, including Sox6 with members of the Hox gene family, chicken Sox11 with members of the Achaete-scute complex, and Sox2 with Oct-3, have also been suggested (28,37,42). These associations may allow for structurally complex DNA-binding and transactivation mechanisms, producing highly specified control of developmental pathways.

ACKNOWLEDGMENTS

We thank Kristin Hickey, Dr. Ian MacDonald, and Dr. Rachel Wevrick for their helpful comments, and acknowledge the support of the RP

Foundation—Fighting Blindness and the Alberta Heritage Foundation for Medical Research.

REFERENCES

1. Mitchell P, Tjian R. (1989) Transcription regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245:** 371–378.
2. Tjian R, Maniatis T. (1994) Transcriptional activation: A complex puzzle with few easy pieces. *Cell* **77:** 5–8.
3. Busch SJ, Sassone CP. (1990) Dimers, leucine zippers and DNA-binding domains. *Trends Genet.* **6:** 36–40.
4. Struhl K. (1990) Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem.* **14:** 137–140.
5. Krumlauf R. (1994) Hox genes in vertebrate development. *Cell* **78:** 191–201.
6. Hill RE, Hanson IM. (1992) Molecular genetics of the Pax gene family. *Curr. Opin. Cell Biol.* **4:** 967–972.
7. Jordan T, Hanson I, Zaletayev D, et al. (1992) The human PAX6 gene is mutated in two patients with aniridia. *Nat. Genet.* **1:** 328–332.
8. Halder G, Callaerts P, Gehring W. (1995) New perspectives on eye evolution. *Curr. Opin. Genet. Dev.* **5:** 602–609.
9. Mirzayans F, Pearce WG, MacDonald IM, Walter M. (1995) Mutation of the PAX6 gene in patients with autosomal dominant keratitis. *Am. J. Hum. Genet.* **57:** 539–548.
10. Tassabehji M, Read AP, Newton V, et al. (1992) Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355:** 635–636.
11. Sanyanusin P, Schimmenti LA, McNoe LA, et al. (1995) Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux. *Nat. Genet.* **9:** 358–363.
12. Laudet V, Stehelin D, Clevers H. (1993) Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res.* **21:** 2493–2501.
13. Waterman M, Jones K. (1990) Purification of TCF-1α, a T-cell–specific transcription factor that activiates the T-cell receptor cα gene ehancer in a context-dependent manner. *New Biol.* **2:** 621–636.
14. Travis A, Amsterdam A, Belanger C, Grosschedl R. (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α enhancer function. *Gene. Dev.* **5:** 880–894.
15. van de Wetering M, Clevers H. (1992) Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *EMBO J.* **11:** 3039–3044.
16. Connor F, Cary PD, Read CM, et al. (1994) DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res.* **22:** 3339–3346.
17. Harley VR, Lovell-Badge R, Goodfellow PN. (1994) Definition of a consensus DNA-binding site for SRY. *Nucleic Acids Res.* **22:** 1500–1501.
18. Ferrari S, Harley VR, Pontigiggia A, Goodfellow PN, Lovell-Badge R, Bianchi ME. (1992) SRY, like HMG 1, recognizes sharp angles in DNA. *EMBO J.* **11:** 4497–4506.
19. Giese K, Cox J, Grosschedl R. (1992) The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69:** 1–20.
20. Grosschedl R, Giese K, Pagel J. (1994) HMG domain proteins: Architectural elements in the assembely of nucleoprotein structures. *Trends Genet.* **10:** 94–100.
21. Clépet C, Schafer AJ, Sinclair AH, Palmer MS, Lovell-Badge R, Goodfellow PN. (1993) The human SRY transcript. *Hum. Mol. Genet.* **2:** 2007–2012.
22. Gubbay J, Collignon J, Koopman P, et al. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346:** 245–250.
23. Sinclair A, Berta P, Palmer MS, et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346:** 240–244.
24. Berta P, Hawkins JR, Sinclair A, Taylor A, Griffiths BL, Goodfellow PN, Fellous M. (1990) Genetic evidence equating SRY and the male sex determining gene. *Nature* **348:** 248–250.
25. Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R. (1991) Male development of chromosomally female mice transgenic for SRY. *Nature* **351:** 117–121.
26. Hacker A, Capel B, Goodfellow PN, Lovell-Badge R. (1995) Expression of Sry, the

mouse sex determining gene. *Development* **121:** 1603–1614.
27. Capel B, Swain A, Nicolis S, Hacker A, Walter MA, Koopman P, Goodfellow PN, Lovell-Badge R. (1993) Circular transcripts of the testis-determining geneSry in adult mouse testis. *Cell* **73:** 1019–1030.
28. Haqq CM, King CY, Ukiyama E, et al. (1994) Molecular basis of mammalian sexual determination: Activation of Müllerian Inhibiting Substance gene expression by SRY. *Science* **266:** 1494–1500.
29. Coriat AM, Muller U, Harry JL, Uwanogho D, Sharpe PT. (1993) PCR Amplification of Sry-related sequences reveals evolutionary conservation of the SRY-box motif. *PCR Methods Appl.* **2:** 218–222.
30. Chardard D, Chesnel A, Gozé C, Doumon C, Berta P. (1993) Pw Sox-1: The first member of the Sox gene family in Urodeles. *Nucleic Acids Res.* **21:** 3576.
31. Denny P, Swift S, Brand N, Dabhade N, Barton P, Ashworth A. (1992) A conserved family of genes related to the testis determining gene, SRY. *Nucleic Acids Res.* **20:** 2887.
32. Foster JW, Graves JA. (1994) An SRY-related sequence on the marsupial X chromosome: Implications for the evolution of the mammalian testis determining gene. *Proc. Natl. Acad. Sci. U.S.A.* **91:** 1927–1931.
33. Griffiths R. (1991) The isolation of conserved DNA sequences related to the human sex-determining region Y gene from the lesser black-backed gull (*Larus fuscus*). *Proc. R. Soc. Lond. B Biol. Sci.* **244:** 123–128.
34. Uwanogho D, Rex M, Cartwright EJ, et al. (1995) Embryonic expression of the chicken Sox2, Sox3, and Sox11 genes suggests an interactive role in neuronal development. *Mech. Dev.* **49:** 23–36.
35. Vriz S, Lovell-Badge R. (1995) The zebrafish Zf-Sox 19 protein: A novel member of the Sox family which reveals highly conserved motifs outside of the DNA-binding domain. *Gene* **153:** 275–276.
36. Behlke MA, Bogan JS, Beer-Romero P, Page DC. (1993) Evidence that the SRY protein is encoded by a single exon on the human Y chromosome. *Genomics* **17:** 736–739.
37. Stevanovic M, Lovell-Badge R, Collignon J, Goodfellow P. (1993) SOX3 is an X-linked gene related to SRY. *Hum. Mol. Genet.* **2:** 2013–2018.
38. Farr CJ, Easty DJ, Ragoussis J, Collignon J, Lovell-Badge R, Goodfellow PN. (1993) Characterization and mapping of the human SOX4 gene. *Mamm. Genome* **4:** 577–584.
39. Schilham MW, Vaneijk M, Vandewetering M, Clevers HC. (1993) The murine Sox-4 protein is encoded on a single exon. *Nucleic Acids Res.* **21:** 2009.
40. Foster JW, Dominguez-Steglich MA, Guioli S, et al. (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* **372:** 525–530.
41. Wagner T, Wirth J, Meyer J, et al. (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* **79:** 1111–1120.
42. Collignon J, Shanthini S, Hacker A, et al. (1996) A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* **122:** 509–520.
43. Kamachi Y, Sockanathan S, Liu Q, Breitman M, Lovell-Badge R, Kondoh H. (1995) Involvement of SOX proteins in lens-specific activation of crystallin genes. *EMBO J.* **14:** 3510–3519.
44. Stevanovic M, Zuffardi O, Collignon J, Lovell-Badge R, Goodfellow P. (1994) The cDNA sequence and chromosomal location of the human SOX2 gene. *Mamm. Genome* **5:** 640–642.
45. Yuan H, Corbi N, Basilico C, Dailey L. (1995) Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Gene. Dev.* **9:** 2635–2645.
46. Ohno S. (1967) *Sex Chromosomes and Sex-Linked Genes.* Springer, Berlin.
47. Goze C, Poulat F, Berta P. (1993) Partial cloning of SOX-11 and SOX-12, two new human SOX genes. *Nucleic Acids Res.* **21:** 2943.
48. Meyer J, Wirth J, Held M, Schempp W, Scherer G. (1996) SOX20, a new member of the SOX gene family, is located on chromosome 17p13. *Cytogenet. Cell Genet.* **72:** 246–249.
49. van de Wetering M, Oosterwegel M, van Norren K, Clevers H. (1993) Sox-4, a Sry-like HMG protein, is a transcriptional activator in lymphocytes. *EMBO J.* **12:** 3847–3854.
50. Denny P, Swift S, Connor F, Ashworth A. (1992) An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. *EMBO J.* **11:** 3705–3712.
51. Connor F, Wright E, Denny P, Koopman P,

Ashworth A. (1995) The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse. *Nucleic Acids Res.* **23:** 3365–3372.

52. Houston CS, Opitz JM, Spranger JW, et al. (1983) The campomelic syndrome: review, report of 17 cases, and follow-up on the currently 17-year old boy first reported by Marteaux et al. in 1971. *Am. J. Med. Genet.* **15:** 3–28.
53. McKusick VA. (1992) *Mendelian Inheritance in Man.* The Johns Hopkins Press, Baltimore, pp. 1262–1263.
54. Tommerup N, Schempp W, Meinecke P, et al. (1993) Assignment of an autosomal sex reversal locus (SRA1) and campomelic dysplasia (CMPD1) to 17q24.3–25.1. *Nat. Gene.* **4:** 170–174.
55. Wright EM, Snopek B, Koopman P. (1993) Seven new members of the Sox gene family expressed during mouse development. *Nucleic Acids Res.* **21:** 744.
56. Wright E, Hargrave MR, Christiansen J, et al. (1995) The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nat. Gene.* **9:** 15–20.
57. Dunn TL, Mynett-Johnson L, Wright EM, Hosking BM, Koopman PA, Muscat GEO. (1995) Sequence and expression of Sox-18 encoding a new HMG-box transcription factor. *Gene* **161:** 223–225.