

Sox 15, a novel member of the murine Sox family of HMG box transcription factors

Marc van de Wetering and Hans Clevers*

Immunology, University Hospital, PO Box 85500, 3508GA, Utrecht, The Netherlands

Received February 15, 1993; Accepted February 20, 1993

EMBL accession no. X70909

The recent cloning of the mammalian sex-determining gene has led to the identification of a large number of related genes (1–3). These so called *Sox* genes (for *Sry*-like HMG box genes) are defined by the presence of a novel type of DNA-binding domain, the HMG box. The HMG box is a 79 amino acid motif, responsible for sequence-specific DNA binding in *Sry* and several other proteins (4). Within the HMG box superfamily, the *Sox* genes occupy a separate branch. To clone *Sox* genes expressed in T lymphocytes, we designed PCR primers based on the published protein sequences of the *Sox 1* to *Sox 4* HMG boxes (forward primer: 5'GGGGAATTCATGGA(TC)GC(GATC)TT(TC)AT(GATC)GT(GATC)TGG and reverse primer 5'GGGAAGCTT(GATC)GG(GATC)CG(AG)TA(CT)TT(GA)TA(GA)T(TC)(GATC)GG. PCR amplification of murine T cell cDNA and subcloning of the products into *EcoRI/HindIII*-digested pBLUESCRIPT SK led to the isolation of seven different *Sox* sequences. Only one of these, *Sox 4*, was expressed in T lymphocytes (van de Wetering *et al.*, submitted). The other six sequences (including the present) were found not to be expressed in T cells, and derived from contaminating genomic DNA. The recloning of *Sox 1* to *4* in this experiment proved the validity of the primers. Two other clones were identical to *Sox 11* and *14*, described more recently (3).

One of the sequences apparently represented a novel *Sox* gene, and was termed *Sox 15*. Figure 1 presents the predicted amino acid sequence of *Sox 15*, aligned with the consensus *Sox* sequence derived by Koopman and colleagues (3). The latter authors have defined subgroups within the murine *Sox* family. Comparison of amino acid identity of *Sox 15* with all other murine *Sox* genes (see Figure 2) revealed *Sox 15* to be a distant member of this family, most closely related to the subgroup that includes *Sox 1–3* and *Sox 14*. However, homology within that subgroup is in the order of 90% between individual members, whereas the identity towards *Sox 15* is only 70–75%. As the *Sox* family has

been most extensively documented in the mouse, we have chosen to follow the nomenclature of Koopman and colleagues (3), rather than that of Denny *et al.* (2). The latter study accumulates sequences from various vertebrate species, potentially confusing orthologues with homologues existing within one species. Our observations add yet another member to the murine *Sox* gene family.

ACKNOWLEDGEMENTS

We are grateful to Peter Koopman and Robin Lovell-Badge for making available data before publication.

REFERENCES

- Gubbay, J. *et al.* (1990) *Nature* **346**, 245–250.
- Denny, P. *et al.* (1990) *Nucleic Acids Res.* **20**, 2887.
- Wright, E.M., *et al.* (1993) *Nucleic Acids Res.* in press.
- Ner, S.S. (1992) *Curr. Biol.* **2**, 208–210.

Consensus **srgerkrlk**iaq **gnpdahn**.ei **skrlgkrwkl** l.esekrpf**i** **eeerlr**.qh mkydy
Sox 15 **ssvqrqmaq** **gnpkahnsei** **skrlgaqwkl** lgdeekrpf**m** **eeakrlrarh** lrdy

Figure 1. Amino acid sequence of *Sox 15* (EMBL accession number X70909) aligned with the consensus *Sox* sequence (3). Identities appear in bold face. PCR oligomers prime directly outside the presented sequence.

<i>Sox</i> gene (3)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	<i>mSry</i>
% identity	69	75	78	60	48	46	50	55	55	51	59	59	47	71	52

Figure 2. Percentage of amino acid residues identical between *Sox 15* and the indicated members of the mouse *Sox* gene family.

* To whom correspondence should be addressed