



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

A gene family of HMG box factors with homology to TCF-1

Castrop, J.; Norren, K. van; Clevers, J.C.

published in

Nucleic Acids Research
1992

document version

Publisher's PDF, also known as Version of record

document license

CC BY-NC

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Castrop, J., Norren, K. V., & Clevers, J. C. (1992). A gene family of HMG box factors with homology to TCF-1. *Nucleic Acids Research*, 611.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl

A gene family of HMG-box transcription factors with homology to TCF-1

Jan Castrop, Klaske van Norren and Hans Clevers*

Department of Immunology, University Hospital, P.O.Box 85500, 3508 GA, Utrecht, The Netherlands

Submitted December 18, 1991

EMBL accession nos X62870–X62872 (incl.)

We and others have recently cloned two related human lymphoid transcription factors, TCF-1 and TCF-1 α (1, 2). Both factors contain a virtually identical version of a novel type of DNA-binding domain, the High Mobility Group 1 (HMG) box (3). TCF-1 and TCF-1 α bind to functional motifs in the T cell-specific enhancers of the TCR- α , - β , δ and CD3- ϵ genes (4). The HMG box of TCF-1 is encoded on two exons separated by a 130 bp intron (M. van de Wetering and H. Clevers, submitted). By low stringency Southern blotting using a cDNA probe encoding the HMG box of human TCF-1, we observed the presence of several related sequences in the human and murine genomes (not shown). To characterize these putative TCF-like genes we have adopted a guessmer-based Polymerase Chain Reaction (PCR)-strategy. Degenerate PCR-primers were designed based on sequences encoding the TCF-1 and TCF-1 α HMG boxes (GGGAATTCAATGC(GATC)TT(TC)ATG(TC)T(GATC)TA(TC)ATGAA and GGAAGCTTAC(GA)TA(GA)TT(GA)TC(ATC)CG(GATC)GC(ATC)G(AT)CCA). A PCR-reaction on human genomic DNA (HeLa) was performed for 35 cycles (Taq Polymerase, Promega; annealing 2', 44°C; extension 1', 72°C; melting 1', 95°C) according to the manufacturer's instructions. Resulting fragments were digested with HindIII and EcoRI and cloned in pBluescriptSK (Stratagene). Sequencing of thirty cloned fragments revealed the existence of two putative TCF-like genes, which were termed TCF-3 and TCF-4. PCR-fragments encoding the TCF homologues were then used to screen a human genomic EMBL-3 library. Restriction maps of the inserts of positive bacteriophages were determined for both candidate genes. Subsequently, fragments hybridizing to the TCF-1 HMG box probe were sequenced. Sequence comparison of TCF-3 and TCF-4 with the two HMG box exons of the TCF-1 gene revealed complete conservation of the sizes of all exons and a striking homology at the predicted amino acid level (TCF-3: EMBL Data Library Accession number X62870; TCF-4: Accession Numbers X62871 and X62872; see Figure 1). Consensus splice donor and splice acceptor sites were present at the appropriate positions. It is to be expected that the proteins encoded by the TCF-3 and TCF-4 genes will display DNA-binding characteristics that are

very similar to those of TCF-1 and TCF-1 α . The TCF-3 and -4 genes are not expressed in cells of the lymphoid lineage (not shown). Further characterization of these genes awaits the identification of their sites of expression.

Previous alignments have revealed 20–30% identity at protein level between members of the HMG box family (1, 2, 3). The observations presented here imply the existence in the mammalian genome of a sub-family of TCF-like HMG-box transcription factors with amino acid identity of >90% in the HMG-box region.

ACKNOWLEDGEMENTS

The authors wish to thank Dr P. Warmerdam for the human genomic EMBL3 library and F. Wien for secretarial help.

REFERENCES

1. van de Wetering, M., Oosterwegel, M., Dooijes, D. and Clevers, H. (1991) *EMBO J.* **10**, 123–132.
2. Waterman, M.L., Fischer, W.H. and Jones, K.A. (1991) *Genes Dev.* **5**, 656–668.
3. Jantzen, H.-M., Admon, A., Bell, S.P. and Tjian, R. (1990) *Nature* **344**, 830–836.
4. Oosterwegel, M.A., van de Wetering, M.L., Holstege, F.C.P., Prosser, H.M., Owen, M.J. and Clevers, H.C. (1991) *Int. Immunol.* **3**, 1189–1192.

TCF-1	TQAESKAEKEAKKPTIKKPLNAFMYMKEMRAKVI AECTLKESAAINQILGRR
TCF-1 α	P-H-QRK-Q-P-R-H-----V-----
TCF-3	KSPVTVKKE-E---HV-----V-----K
TCF-4	KHQD--K-E-K---H-----V-----
TCF-1	WHALSREEQAKYYEFARKERQLHMQLYPGWSARDNY
TCF-1 α	-----L-----
TCF-3	--N-----L-----T-----
TCF-4	-----L-----

Figure 1. Alignment of the predicted HMG-boxes of the four TCF genes. Top: 5' HMG box exon. Bottom: 3' HMG box exon. A dash indicates identity to the corresponding residue in TCF-1. Significant sequence divergence only occurs in the first eleven amino acid residues.

* To whom correspondence should be addressed