OTEX, an androgen-regulated human member of the *paired*-like class of homeobox genes

Christoph GESERICK, Bertram WEISS, Wolf-Dieter SCHLEUNING[1] and Bernard HAENDLER[2]

Research Laboratories of Schering AG, Muellerstrasse 170–178, Berlin 13342, Germany

paired genes emerged early in evolution and code for homeobox transcription factors, having fundamental roles in various biological processes. We identified a novel human member of the *paired*-like class, which we named *OTEX*. A phylogenetic analysis revealed that *OTEX* belonged to the recently defined *PEPP* subfamily of *paired*-like homeobox genes. It was organized into three introns and, like the other *PEPP* genes, it was mapped to chromosome X. Its transcripts were detected mainly in the ovary, testis and epididymis, but also in the prostate and mammary gland. In the PC-3/ARwt prostate cell line, *OTEX* expression was stimulated dramatically following androgen treatment. Immunofluorescence studies revealed an exclusively nuclear localization of the OTEX protein. Mutation of the RARCRRHQRE amino acid sequence present at the C-terminus of the OTEX homeodomain resulted in a mainly cytoplasmic localization, indicating that this motif harboured the nuclear localization signal. No inherent transactivation function was seen for OTEX using the one-hybrid assay, and no homodimer formation was observed in the two-hybrid assay, suggesting that additional partners were needed for this activity. Taken together, the data show that OTEX represents a novel, androgen-regulated, paired-like homeobox protein, with possibly an important role in human reproduction.

Key words: evolution, homeodomain, nuclear protein, ovary, testis.

INTRODUCTION

Homeobox genes code for transcription factors harbouring a characteristic 60-amino-acid-long DNA-binding region, the homeobox domain (HD), and act as 'master' switches in developmental processes [1,2]. They are associated intimately with embryogenesis, as evidenced by extensive mutational analyses in the fly and mouse [1,2]. A critical role in development and functionality of reproductive organs has furthermore been reported for several homeobox genes. Examples include *Pax2* and *Lhx9*, which are essential for gonad formation, and *Hoxa*-10 and *Hoxa*-11, which are necessary for testicular descent and implantation [3–6].

Homeobox proteins are grouped in different classes, depending on their HD sequence and on the presence of additional motifs [1,2]. The paired class is characterized by the presence of a serine residue at position 50 of the HD, and of an additional DNA-binding region, the paired domain. Absence of both these features is characteristic of the paired-like class, for which the *Drosophila* aristaless protein is the prototype [7]. The PEPP subfamily has recently emerged as a subset of the paired-like class of homeobox proteins [8]. It includes Pem, Esx1, Spx1 (a splice variant of Esx1), ESXR1, Psx1 and Psx2 (also called Gpbox) [9–16]. All the *PEPP* genes are located on chromosome X and share a similar structure, strongly suggestive of a common ancestor [8,16]. Expression of the *PEPP* genes has been demonstrated in the placenta and in the developing gonads [9–16]. In addition, *Pem* expression takes place in adult testis, epididymis, ovary and muscle [10], and *Esx1* as well as *ESXR1* expression occurs in adult testis [11,12].

The function of the various PEPP proteins is largely unknown. Mice with a disrupted *Pem* gene have remarkably few phenotypic alterations when reared under normal laboratory conditions, suggesting redundant functions with other homeobox genes [17]. The same holds true for *Psx2*-deficient mice, which have no obvious phenotype [18]. The situation is different for *Esx1*. Here, a maternally inherited inactive mutant allele leads to placental abnormalities and fetal-growth retardation, indicating an important role in maternal–fetal communication [19].

Apart from the sequence of the recently identified *ESXR1* gene, little is known about the human members of the *PEPP* subfamily. In view of the potentially important functions of these homeobox genes in reproduction, we set out to identify new human members of the family. Here, we present the identification of *OTEX*, which encodes a protein belonging to the paired-like class and maps to chromosome X. The selective expression of *OTEX* in reproductive organs, its steroid hormone regulation and the nuclear localization of the corresponding protein suggest a role as a master switch in human gametogenesis.

EXPERIMENTAL

Identification and cloning of *OTEX*

A BLAST search of public databases enabled the identification of an expressed sequence tag (EST; AI631510) sharing significant sequence identity with *PEPP* family members. This allowed the identification of a genomic region (AC005023) containing the sequence information for an open reading frame putatively encoding a novel PEPP protein. This information was used to

Abbreviations used: DAPI, 4,6-diamidino-2-phenylindole; EST, expressed sequence tag; FCS, fetal-calf serum; HD, homeobox domain; NLS, nuclear localization signal; RT, reverse transcriptase.

[1] Present address: Paion GmbH Research Center, Biotech-Park Charlottenburg, Tegeler Weg 33, 10589 Berlin, Germany.

[2] To whom correspondence should be addressed (e-mail bernard.haendler@schering.de).

The nucleotide sequence data reported will appear in DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession number AY099086.

design the primers 5′-ATGGCGCGTTCGCTCGTCCACGA-C-3′ (OT1) and 5′-TAGTCCACGACGATGTAGACACAG-3′ (OT2), which were used for PCR amplification of full-length *OTEX*, starting from human ovary first-strand cDNA. PCR conditions were: 30 s at 94 °C, 30 s at 58 °C and 2 min at 68 °C for 40 cycles using the Advantage 2 PCR-kit (Clontech, Heidelberg, Germany). Additional PCR amplifications were performed for a precise determination of the extremities, using ovary and testis Marathon-Ready cDNA (Clontech) as templates. For analysis of the 5′-end, we used the *OTEX*-specific reverse primer 5′-TTGAGTGTGTCGGAAAACACTTTCCA-G-3′ (OT4) and the universal primer 5′-CCATCCTAAT-ACGACTCACTATAGGGC-3′. For the 3′-end, we used the *OTEX*-specific forward primer 5′-CTGGGGGCAGCATCAA-GCGCAGAAGG-3′ (OT3) and the universal primer 5′-CCA-TCCTAATACGACTCACTATAGGGC-3′. The generated products were separated by agarose-gel electrophoresis (Gibco BRL Life Technologies, Karlsruhe, Germany), purified using the Silica Spin Fragment DNA kit (Biometra, Göttingen, Germany) and subcloned into the pCRII-TOPO vector (Invitrogen, Karlsruhe, Germany). Sequencing was performed with the *Taq* polymerase using the BigDye Terminator Cycle Sequencing kit (Roche Applied Biosystems, Foster City, CA, U.S.A.). The amplified products were purified from the dye terminators using Centriflex gel-filtration cartridges (MoBiTec, Göttingen, Germany) and their sequence was determined on both strands (Agowa, Berlin, Germany).

Reverse transcriptase (RT)-PCR analysis

RNAs from different human tissues were obtained from Invitrogen, Clontech and Ambion (Huntingdon, Cambs., U.K.). Total RNA from PC-3 and PC-3/ARwt cells was purified using the RNAeasy kit (Qiagen, Hilden, Germany). The RNA was reverse-transcribed using the ProSTAR kit (Stratagene, La Jolla, CA, U.S.A.). Primer pairs for amplification of full-length *OTEX* (OT1 and OT2) or of a 300 bp internal fragment (OT3 and OT4) were used in PCR. Primers specific for the gene coding for the housekeeping, 23 kDa highly basic protein or for β-actin were used for amplification of an internal control. PCR conditions were 30 s at 95 °C, 30 s at 62 °C and 1 min at 72 °C for 35 cycles, using Ready-To-Go PCR beads (Amersham Biosciences, Freiburg, Germany).

Plasmid construction

Two constructs were engineered for use in the one- and two-hybrid assays. Full-length OTEX cDNA was PCR-amplified while adding compatible flanking restriction sites, and inserted between the *Bam*HI and *Xba*I sites of the pCMV-AD and pCMV-BD plasmids (Stratagene). The resulting constructs were checked by DNA sequencing using the BigDye Terminator Cycle Sequencing kit (Roche Applied Biosystems), as described above.

For the immunofluorescence experiments, full-length OTEX cDNA was amplified by PCR while adding compatible flanking restriction sites, and inserted between the *Bam*HI and *Eco*RI sites of the pcDNA3.1/myc-his(+)A plasmid (Invitrogen). For mutation of the nuclear localization signal (NLS), we used the QuickChange kit (Stratagene) to insert compensatory frame-shift mutations before and after the candidate region in two steps. First, we inserted a base-pair at the 3′-end using the 5′-CATCAGAGAGAATTTAATGCTCGCC-3′ primer and its complementary sequence. In the following step, we deleted a base-pair 18 codons further upstream using the 5′-GGTG-TGACTGAAGACAAGTGCGGG-3′ primer and its complementary sequence. A second mutant was generated by deleting a bp 10 codons upstream of the initial insertion using the 5′-GGTTTAAGAATAAAGGGCCAGATG-3′ primer and its complementary sequence. The presence of the mutations was confirmed by DNA sequencing.

Cell culture and transfections for one- and two-hybrid assays

Minimal essential medium, Kaighn's modification of Nutrient Mixture F-12K, RPMI 1640, OPTI-MEM, streptomycin, penicillin, Geneticin and L-glutamine were obtained from Gibco BRL Life Technologies. Fetal-calf serum (FCS) was from PAA Laboratories (Cölbe, Germany).

The CV-1 cells were grown at 37 °C in a 5% CO_2 atmosphere in minimal essential medium containing 10% (v/v) FCS, 100 units/ml penicillin, 100 μg/ml streptomycin and 4 mM L-glutamine. The PC-3 cells were grown at 37 °C in a 5% CO_2 atmosphere using Kaighn's modification of Nutrient Mixture F-12K containing 10% (v/v) FCS, 100 units/ml penicillin, 100 μg/ml streptomycin and 2 mM L-glutamine. The PC-3/ARwt cells were cultured at 37 °C in a 4.5% CO_2 atmosphere in RPMI 1640 containing 10% (v/v) FCS, 100 units/ml penicillin, 100 μg/ml streptomycin, 4 mM L-glutamine and 600 μg/ml Geneticin.

For the hormone treatment, the cells were seeded in medium supplemented with charcoal-stripped FCS for 24 h, before adding 1 nM R1881 (Dupont NEN, Boston, MA, U.S.A.). The cells were harvested 24 h later for RNA purification.

For the one-hybrid and two-hybrid assays, the cells were seeded at a concentration of 10000/100 μl per well in 96-well plates. Transfection was performed 18 h later using FuGene6 (Roche Molecular Biochemicals, Mannheim, Germany) in OPTI-MEM and the pCMV-based plasmids. Measurement of luciferase reporter activity was performed 24 h later in a Lumicount luminometer (Packard, Dreieich, Germany), after adding 100 μl of LucLite Plus reagent (Packard). The activity of the pGL3 promoter vector (Promega, Mannheim, Germany) was determined on parallel samples to assess transfection efficiency. The average value of six wells treated in parallel was determined, and the transfections were repeated at least three times independently.

Immunostaining of transfected cells

The different myc fusion constructs were used for transfection of CV-1 and PC-3 cells. The pcDNA3.1-LacZ-myc plasmid (Invitrogen) was used as the control. The cells were grown as described above, except that 5% (v/v) charcoal-stripped FCS was used. They were seeded in four-well glass-chamber slides (Nalge Nunc International, Wiesbaden, Germany) at a concentration of 100000 cells/well. Transfection was allowed to proceed for 24 h, after which time the cells were washed with PBS and fixed with methanol for 5 min. After another wash with PBS, the cells were incubated for 30 min with a mouse anti-(c-Myc) IgG (clone 9E10, BD Clontech) diluted 1:200 in PBS. They were then washed twice with PBS, and incubated with a cyanine-conjugated goat anti-mouse IgG (Jackson Immunoresearch Laboratories, West Grove, PA, U.S.A.) diluted 1:50, or alternatively with a rhodamine-labelled goat anti-mouse IgG (Jackson Immunoresearch Laboratories) diluted 1:30. 4,6-Diamidino-2-phenylindole (DAPI; from Calbiochem, Bad Soden, Germany) was added for nuclear staining. After two more washes with PBS, the cells were fixed in a 1% (w/v) paraformaldehyde solution in PBS, before mounting with Vectashield (Vector Laboratories, Burlingame, CA, U.S.A.). The stained cells were visualized with an Axiophot microscope (Carl Zeiss, Göttingen,

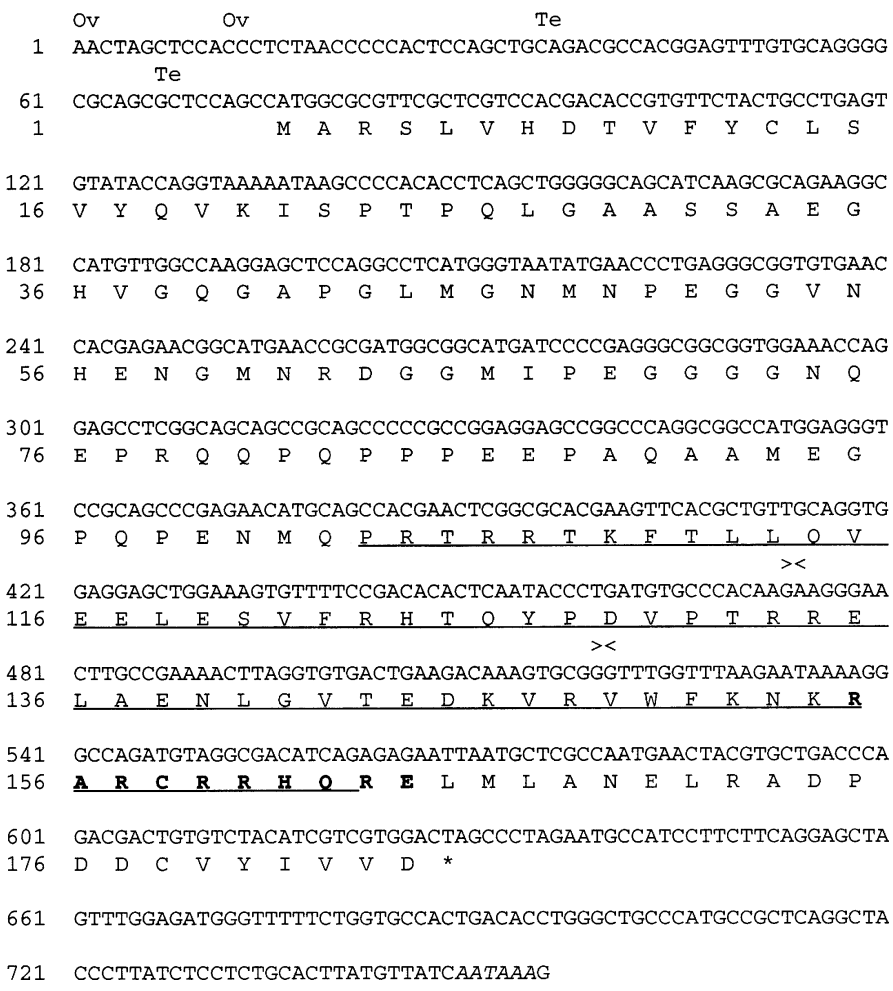


Figure 1 Sequence analysis of *OTEX*

Nucleotide and predicted amino acid sequences of *OTEX* cDNA. The 5′-ends of the longest cDNAs obtained from ovary (Ov) or testis (Te) RNA are indicated. The polyadenylation signal is shown in italics, and the exon/intron boundaries are marked with converging arrowheads. The HD is underlined and the NLS is highlighted in bold letters. The accession number of the sequence in the GenBank®/EMBL database is AY099086.

Germany) and a 3 charge-coupled-device colour camera (Sony, Cologne, Germany), using the appropriate filters.

RESULTS

Cloning of OTEX cDNA

The screening of public databases enabled the identification of an EST and of genomic sequences potentially coding for a novel paired-like homeobox protein, which was named OTEX. Specific primers were devised to amplify the corresponding cDNA from human ovary, which was then cloned and sequenced on both strands. In addition, the 5′- and 3′-extremities were re-amplified separately from testis and ovary cDNA to check for tissue-specific differences and splice variants, but co-linear sequences only differing in their lengths were obtained. The sequence of the longest cDNA clone identified is shown in Figure 1. An open reading frame coding for a 184-amino-acid-long protein was found. The translation start codon was embedded in a sequence fitting with the consensus. An in-frame stop codon was present upstream, indicating that the coding region was complete. A polyadenylation signal was observed near the end of the 3′-untranslated region.

Sequence analysis of OTEX

Analysis of the deduced OTEX protein primary structure revealed the presence of a characteristic 60-amino-acid HD (Figure 1; shown underlined) related to the paired HD (Figure 2). High sequence identities were found with PaxD from the basal cnidarian *Acropora millepora*, PAX3 and PAX7 from human, and gooseberry from *Drosophila* [20]. However, unlike these homeobox proteins, OTEX neither had a serine residue at position 50 nor did it possess an additional paired domain. This characteristic located OTEX in the paired-like family, of which the *Drosophila* protein aristaless is the prototype [7]. A phylogenetic analysis (Figure 3) indicated that OTEX was a new member of the recently defined PEPP subfamily of paired-like homeobox proteins, which comprises Pem, Esx1, Spx1 and ESXR1, as well as Psx1 and Psx2 [8]. When aligning the HDs, we calculated 53.3% sequence identity with ESXR1, 43.3–45% with Psx1 and 30–31.7% with Pem. A database search and additional phylogenetic analysis enabled us to identify another human and two more mouse members of the PEPP subfamily: THG1 (GenBank® accession number AF317219), Ehox (accession number NM_021300) and a RIKEN clone (accession

	Helix 1		Helix 2		Helix 3		Id	Pos
	10	20	30 ><	40	>< 50	60		
HsOTEX	PRTRRTKFTL	LQVEELESVF	RHTQYPDVPT	RRELAENLGV	TEDKVRVWFK	NKRARCRRHQ	100	Xq24
AmPaxD	Q-RS-----H	A-LNA--KA-	QK------Y-	-E---HR-SL	--AR-Q---S	-R---L-KKD	55	
HsESXR1	K-R---A--Q	F-LQ---NF-	DES-----VA	-ER--AR-NL	---R-Q---Q	-R--KWK-N-	53.3	Xq22
HsARX	Q-RY--T--S	Y-L----RA-	QK-H----F-	-E---MR-DL	--AR-Q---Q	-R--KW-KRE	53.3	Xp22.13
DmGsb-n	Q-RS--T--A	E-L-A--RA-	SR------Y-	-E---QTTAL	--ARIQ---S	-R---L-K-S	53.3	2 60F1
HsPAX3	Q-RS--T--A	E-L----RA-	ER-H---IY-	-E---QRAKL	--AR-Q---S	-R---W-KQA	51.7	2q35
HsPAX7	Q-RS--T--A	E-L----KA-	ER-H---IY-	-E---QRTKL	--AR-Q---S	-R---W-KQA	51.7	1p36
HsCRX	Q-RE--T--R	S-L----AL-	AK------YA	-E-V-LKINL	P-SR-Q----	-R--K--QQR	51.7	19q13
DmGsb	Q-RS--T-SN	D-IDA--RI-	AR------Y-	-E---QST-L	--AR-Q---S	-R---L-KQL	51.7	2 60F1
DmA1	Q-RY--T--S	F-L----KA-	SR-H----F-	-E---MKI-L	--ARIQ---Q	-R--KW-KQE	51.7	2 21C1
HsARIX	Q-RI--T--S	A-LK---R--	AE-H---IY-	-E---LKIDL	--AR-Q---Q	-R--KF-KQE	50	11q13
HsALX4	K-RN--T--S	Y-L----K--	QK-H----YA	-EQ--MRTDL	--AR-Q---Q	-R--KW-KRV	50	11p11.2
HsSHOX	Q-RS--N---	E-LN---RL-	DE-H---AFM	-E--SQR--L	S-AR-Q---Q	-R--K--KQE	50	Xp22.32
MmEsx1	--RY-IC--P	I-LQ---AF-	QRV----LFA	-V---RR--L	P-PR-Q---Q	-R--KW--LR	50	X 57
HsALX3	K-RN--T-ST	F-L----K--	QK-H----YA	-EQ--LRTDL	--AR-Q---Q	-R--KW-KRE	48.3	1p21-p13
DmPrd	Q-RC--T-SA	S-LD---RA-	ER-----IY-	-E---QRTNL	--ARIQ---S	-R---L-KQH	48.3	2 33C1
HsTHG1	QQPNVHA--P	--LQ---RI-	QRE-F-SEFL	--R--RSMN-	--LA-QI--E	-R--KW----	45	Xq24
MmPsx2	Q-----R--H	S-LRD-GRL-	QENRF-SLRV	--D--RWM--	D-SD-QE---	MR--LF---S	45	X 12.7
MmPsx1	L-Y---R--H	S-LHD--RL-	QE-R--SLRA	--D--RWM--	D-CD-QN--R	MR--LFQ-NR	43.3	X 12.7
MmEhox	Q-SLHYN-QW	W-LQ---RI-	QQNHFIRAEE	--H--RWI--	S-AR-MT---	KR-EHF--G-	36.7	X A2
MmRIKEN	RHGWQQS-NV	--LQ----I-	QCNH-ISTKE	ANR--RSM--	S-AT-QE--L	KR-EKY-SYK	31.7	Xq9D4Y3
MmPem	MPLQGSR-AQ	HRLR----IL	QR-NSF---	-ED-DRLMDA	CVSR-QN---	IR--AA--TR	31.7	X 12.7
RnPem	MPL--PGS-Q	RRLA---RIL	LSSGSSSG--	WE--DRWMDI	SVSR-QN---	IR--AY--NR	30	Xq35-36

Figure 2 Amino acid sequence alignment of HDs from OTEX and related proteins

Members of the PEPP subfamily are underlined. Dashes indicate identity with human OTEX sequence. The percentage identity (Id) as well as the chromosomal localization (Pos) are shown in the two furthermost right-hand columns.

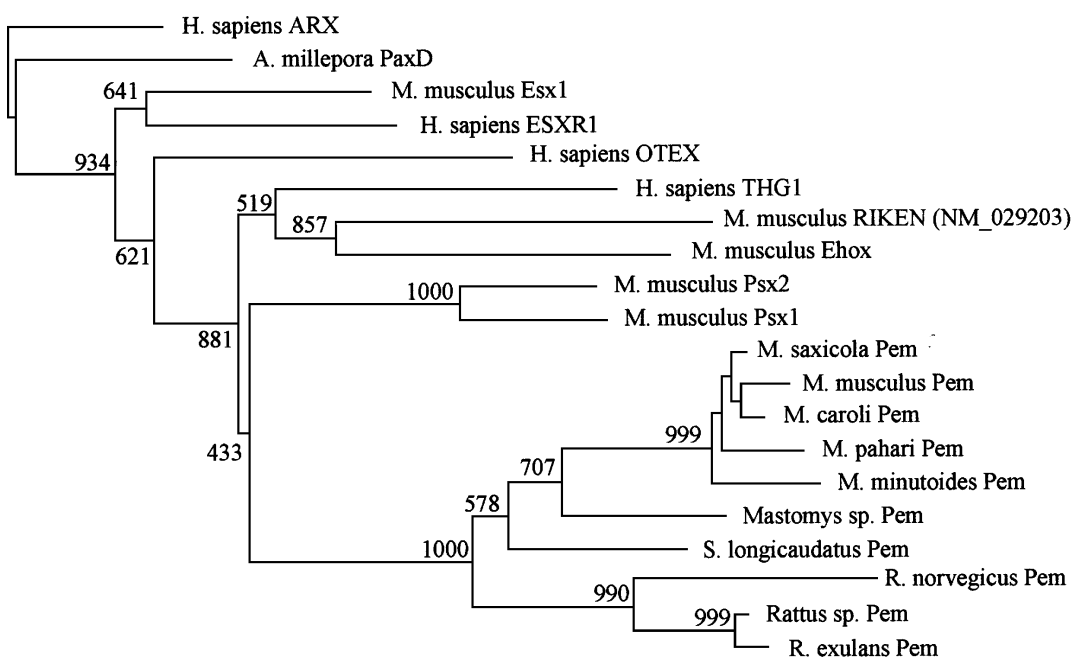
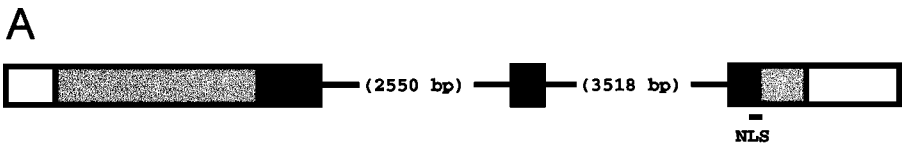


Figure 3 Phylogenetic analysis of PEPP HD sequences

The tree is shown as a rooted phylogram using human ARX as outgroup. The numbers near the branching nodes indicate the bootstrap values. For clarity, some bootstrap values were omitted.

number NM_029203) (Figures 2 and 3). The corresponding HDs exhibited 45 %, 36.7 % and 31.7 % sequence identity respectively with the OTEX HD. A comparison with rat and mouse Pem HDs revealed THG1 to be more closely related than OTEX (33.3–35 % as compared with 30–31.7 %).

A hallmark shared by the HDs of all but one PEPP subfamily members, and not observed in other HDs, was the presence of an arginine residue at position 58 (Figure 2). Another frequent feature was the presence at positions 39 and 40 of a glutamate and a valine residue respectively. Despite the low overall



B

intron	EXON	intron	Phase
AACTA..	(473 bp) ..ACAAGgtaagtgg	(2550 bp)	2
ctccacagAAGGG..	(46 bp) ..TGCGGgtcagtac	(3518 bp)	0
gattgtagGTTTG..	(236 bp) ..TAAAG		

Figure 4 Gene structure of *OTEX*

(**A**) Schematic representation of the *OTEX* gene. The 5′- and 3′-untranslated ends are shown in white; the coding region is in grey and black (for the HD). The approximate location of the NLS is indicated. (**B**) The sequence of the exon/intron boundaries is given. Phase 0 indicates that the intron is located between two codons, and phase 2 indicates that it lies after the second base of the codon.

conservation, the OTEX HD shared additional rare characteristics with the Pem HD, namely, the presence of a helix-breaking proline residue at position 29, and of a lysine at position 50, a site essential for the DNA-binding specificity of homeobox proteins. Pro[29] has already been described in the more distantly related bicoid and Hox HDs [1,2], whereas Lys[50] was found in Psx2 and Ehox, as well as in the otd/Otx/Crx, Ptx and Gsc families [7]. Other residues known to be critical for HD function were maintained in OTEX. Positions 47 and 48 of helix 3, which are essential for interaction with helix 1, were occupied by valine and tryptophan residues respectively, and position 53 was occupied by the highly conserved arginine residue, which binds to the sugar–phosphate backbone of DNA.

An alignment of the regions lying N- or C-terminal of the HD showed furthermore that, despite little overall sequence identity, there were a number of amino acids maintained between OTEX and the other PEPP proteins (results not shown). No significant homology was found with any other protein.

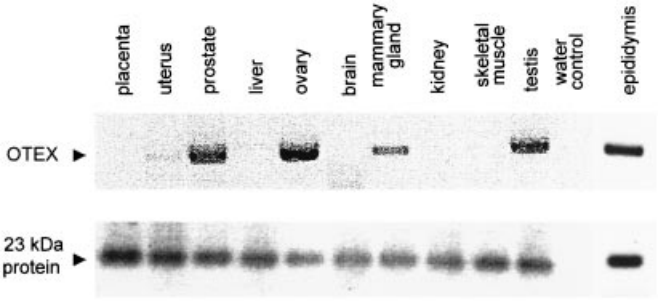


Figure 5 Tissue distribution of *OTEX*

RT-PCR was performed with OTEX-specific primers using reverse-transcribed human RNA from the indicated organs. The PCR products were separated on gels and stained with ethidium bromide. The RNA levels of the 23 kDa highly basic protein were determined as a control.

Chromosomal localization and organization of the *OTEX* gene

A search in human genome databases enabled us to map the *OTEX* gene to chromosome X, at position Xq24. All the other *PEPP* genes are also found on chromosome X. The organization of the *OTEX* gene was established by comparing the cDNA and genomic sequences. We found the *OTEX* gene to be composed of three exons and two introns (Figure 4A). The first exon comprised the 5′-untranslated region and more than half of the coding part. The second exon coded for the central part of the HD, and the third one for the remainder of the protein and the complete 3′-untranslated region. Both introns started with GT and ended with AG (Figure 4B), in accordance with the splicing rules [21]. They interrupted codons at different positions, and were therefore not in the same phase.

A comparison with other *PEPP* genes showed striking similarities. The intron located after the codon for HD position 46 is found in *Pem*, *Esx1*, *ESXR1*, *Psx2* and *THG1*, and has also been reported in other homeobox genes. The intron found in *OTEX* after the codon for HD position 31 is far more unusual, but is also present in other *PEPP* genes, as well as in *aristaless* [7]. Other *PEPP* subfamily members possess one or more additional 5′ exons, but an analysis of the genomic sequence upstream of *OTEX* using a prediction program did not reveal the presence of putative exons.

A potential Cap site was found in the genomic sequence approx. 30 bp upstream of the 5′-end of the longest clone we isolated (results not shown). There was no sequence resembling the TATA consensus at the appropriate upstream position, but several $PyPyA_{+1}NT/APyPy$ initiator consensus motifs [22] were found (results not shown). The absence of a *bona fide* TATA box has already been reported for the mouse and rat *Pem* proximal promoters [10,23].

Tissue distribution of *OTEX* transcripts

RT-PCR was performed on reverse-transcribed human RNA. Primers overlapping the initiation or stop codons were used in order to assess the expression levels of full-length *OTEX* transcripts. This also excluded the risk of amplifying contaminating genomic DNA, since two introns were present in the corresponding gene region. The strongest signal was detected in the ovary, testis and epididymis; hence the name *OTEX* (Figure 5). Specific transcripts were detected additionally in the prostate and the mammary gland. A very weak level of expression was visible in the uterus.

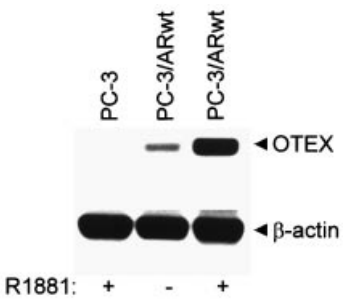


Figure 6 Androgen regulation of *OTEX* expression

RT-PCR was performed with *OTEX*-specific primers using reverse-transcribed RNA extracted from the cell lines indicated. The PCR products were separated on gels and stained with ethidium bromide. The RNA levels of β-actin were determined as a control.

Androgen regulation of *OTEX*

Expression of *Pem*, the founding member of the *PEPP* family, is strongly regulated by androgens in the testis and epididymis [10]. We tested whether this was also the case for *OTEX*, in view of the presence of specific transcripts in the testis, epididymis and prostate, three main target organs of androgens. We analysed PC-3 cells (which do not express the androgen receptor) and PC-3/ARwt cells, a derived cell line stably transfected with the human wild-type androgen receptor. RNA was extracted and analysed by RT-PCR using primers that amplified the full coding sequence of *OTEX* (Figure 6). The results showed that *OTEX* was not expressed in the parental PC-3 cell line lacking the androgen receptor, whereas the PC-3/ARwt cells showed a specific signal. This signal was greatly induced if the PC-3/ARwt cells were treated beforehand with the androgen R1881.

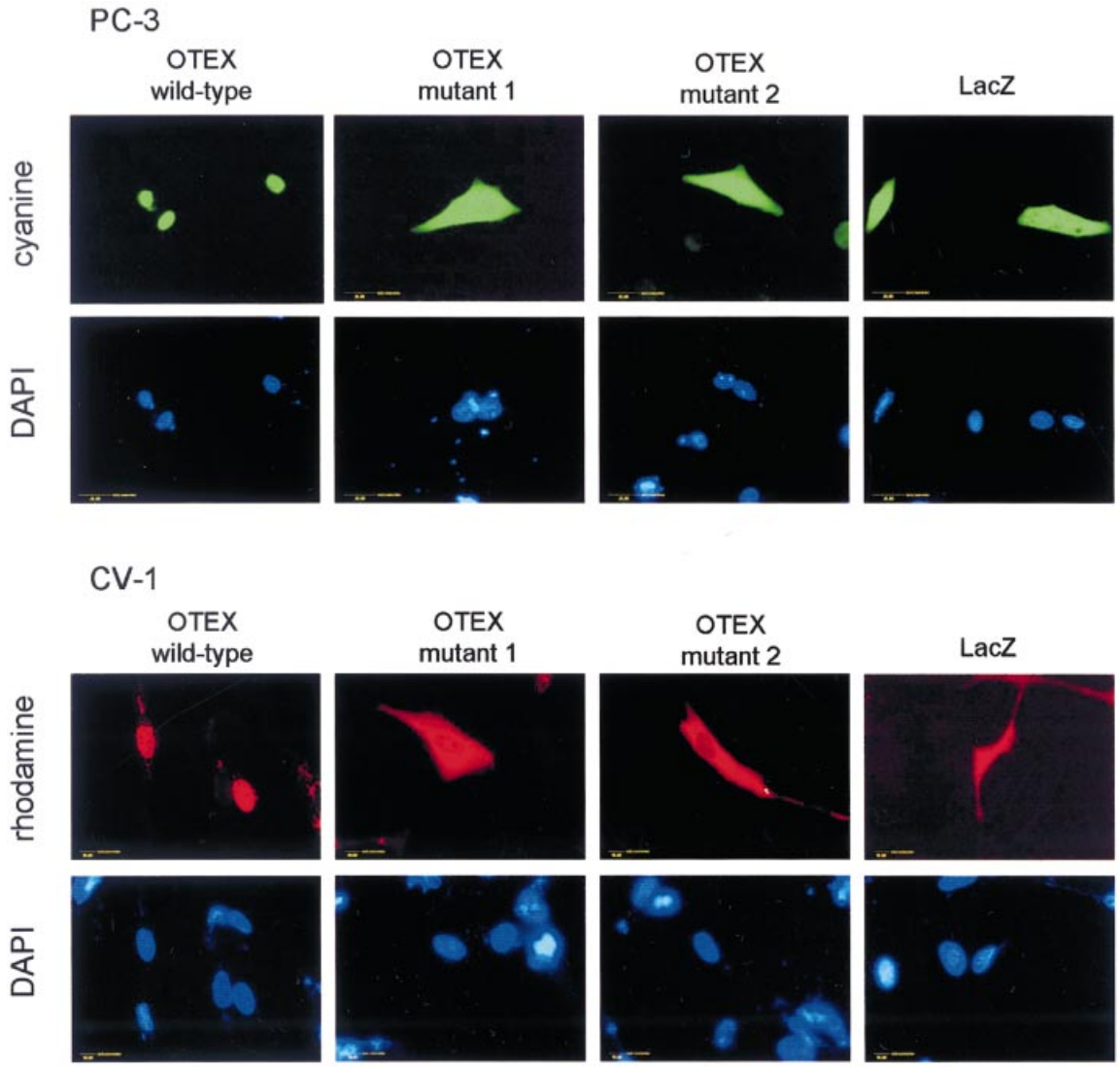


Figure 7 Nuclear localization of OTEX

Cells were transfected with expression constructs for the indicated proteins. Subcellular localization was determined using a specific primary antibody and a fluorophore-labelled secondary antibody. DAPI staining was performed to visualize the nuclei.

Table 1 Comparison of NLS sequences found in PEPP proteins

The amino acid sequence of OTEX NLS is shown written out. In the other sequences, dashes indicate identity with OTEX. The consensus was defined as amino acids maintained in at least eight out of ten sequences. Hs, *Homo sapiens*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*.

Protein	NLS
Hs OTEX	RARCRRHQRE
Hs ESXR1	--KWK-N--V
Hs THG1	--KW-----A
Mm Esx1	--KW--LR-A
Mm Psx1	--LFQ-NR-V
Mm Psx2	--LF---S-L
Mm Ehox	-EHF--G-SQ
Mm RIKEN	-EQY-SYK-L
Mm Pem	--AA--TR-R
Rn Pem	--AY--NR-R
Consensus	RAxxRRxxRx

Subcellular localization of OTEX and identification of the NLS

Because *OTEX* potentially coded for a homeobox transcription factor, we set out to determine its subcellular localization. The full-length *OTEX* coding sequence was fused to a Myc tag in the pcDNA plasmid. Transfection was performed using CV-1 or PC-3 cells, and the results were visualized by fluorescence microscopy. We used an anti-Myc antibody and a secondary rhodamine-labelled antibody to detect the OTEX–Myc fusion protein in CV-1 cells. For PC-3 cells, a cyanine-labelled secondary antibody was used. DAPI staining was performed in parallel to localize the cell nuclei. The results showed that in both CV-1 and PC-3 cells, OTEX was localized exclusively in the nucleus (Figure 7).

Active nuclear translocation usually necessitates the presence of an NLS recognized by a receptor named importin [24]. In order to identify this signal, we mutated a stretch of the OTEX protein that contained several arginine and lysine residues, since these positively charged amino acids form the core of the NLS [25]. We generated two mutants of the same region: one with a modified 18-amino-acid stretch (positions 147 to 164; Figure 1), and the other with a modified 10-amino-acid stretch (positions 155 to 164; shown in bold in Figure 1). Following transfection and immunostaining, we found that both mutants were localized mainly in the cytoplasm, with only minor staining in the nucleus (Figure 7). The residual nuclear staining was probably due to passive diffusion into the nucleus facilitated by the relatively small size of the OTEX–Myc fusion protein. A comparison between the identified NLS of OTEX and the corresponding region of other PEPP proteins enabled us to identify a consensus motif of RAxxRRxxRx (Table 1; where 'x' denotes 'any amino acid').

Activity of OTEX in one- and two-hybrid assays

In order to determine whether OTEX possessed an independent transactivation function, we performed a one-hybrid assay. The *OTEX* coding region was introduced into the pCMV-BD plasmid, in frame with the GAL4 DNA-binding domain. Following co-transfection of CV-1 cells with a vector containing GAL4 DNA elements upstream of a luciferase gene, we measured reporter activity after 24 h. No stimulation of activity was detected with any of the plasmid concentrations tested, as compared with samples in which only the reporter plasmid was transfected (Figure 8A). Comparable results were observed using PC-3 cells (results not shown).

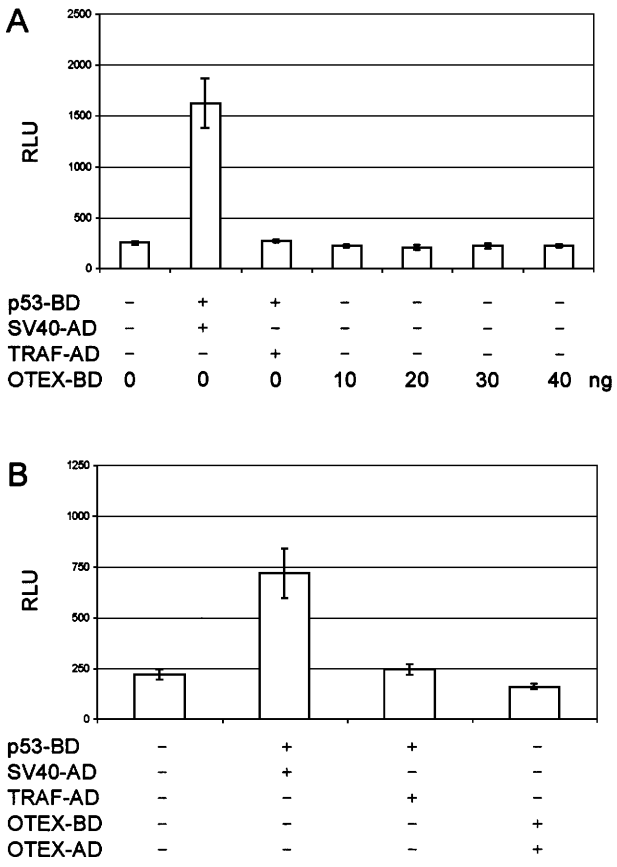


Figure 8 OTEX activity in one- and two-hybrid assays

(**A**) CV-1 cells were transfected with the indicated amounts of an expression vector for OTEX fused to the DNA-binding domain of GAL4 (BD). (**B**) CV-1 cells were co-transfected with expression vectors for OTEX fused to the DNA-binding domain of GAL4 and OTEX fused to the transactivation domain of NFκB (AD). The interaction between p53 and simian virus 40 (SV40) was determined as a positive control. Transactivation was measured after 24 h as the induction of luciferase reporter activity in relative light units (RLU). The results are representative of three separate experiments, and the bars are the means $\pm$ S.E.M. of sextuplicate values. TRAF, tumour-necrosis-factor-receptor-associated factor.

Several homeobox proteins form homodimers or heterodimers for DNA binding. In order to determine whether OTEX interacted with itself, we co-transfected plasmids coding for fusions between OTEX and the GAL4 DNA-binding domain or OTEX and the nuclear factor κB ('NF-κB') activation domain, together with the reporter vector. No significant interaction between OTEX proteins was detected in either CV-1 cells (Figure 8B) or PC-3 cells (results not shown).

DISCUSSION

Identification of a novel member of the *PEPP* subfamily

A novel human homeobox gene has been identified and named *OTEX*, on the basis of its expression in the ovary, testis and epididymis. A sequence comparison indicated that *OTEX* belonged to the recently defined *PEPP* subfamily of *paired*-like

homeobox genes. The *PEPP* genes are probably derived from a common ancestor related to the *Drosophila aristaless* gene. Several family members are known in mouse, but only one, *ESXR1*, had been characterized until now in humans. Besides *OTEX*, we identified *THG1* as another new member of the human *PEPP* subfamily, and *Ehox* and the RIKEN clone NM-029203 as two new murine members.

The existence of a common ancestor for the PEPP genes was supported by several pieces of evidence. First, all PEPP genes were located on chromosome X and, in the case of human members, were clustered in the Xq22–Xq24 region. Secondly, their HD coding region was interrupted by two introns, as found for *aristaless*. This rare intron is not found in the *PaxD* gene of *A. millepora* or in human *SHOX*, despite the comparatively high amino acid sequence identity of the respective HDs. Thirdly, an arginine residue was found at position 58 of the encoded HD (except for the RIKEN clone NM-029203) and, furthermore, OTEX contained a lysine residue at position 50, in common with Pem, Psx2 and Ehox, which is essential for DNA-binding specificity. Finally, the expression profiles of *PEPP* members partially overlapped, as they were often detected in the ovary, testis and placenta.

The *PEPP* genes are not the closest relatives of *aristaless*. Human *ARIX* [26], *ALX3* [27], *ALX4* [28], as well as the *ARX* gene (GenBank® accession number AY038071) are 72–85% identical in their deduced HD with that of *aristaless*. *ARX* is located on chromosome X, whereas the other paralogues are on chromosomes 1 and 11. *SHOX* is another human homeobox gene mapping to chromosome X. It is expressed in a variety of tissues [29], and human genetic studies have shown that its mutation leads to multiple phenotypes, including short stature and wrist deformity [29]. In the four human aristaless paralogues and in SHOX, position 58 of the HD is occupied by a lysine rather than an arginine residue. HD position 50, which is implicated in DNA-binding specificity, is a glutamine residue, which is never found in the HDs of PEPP proteins.

A phylogenetic analysis of PEPP HDs using human ARX HD as the 'outgroup' revealed the existence of five subgroups (Figure 3): (i) mouse and rat Pem; (ii) murine Psx1 and Psx2; (iii) human THG1, mouse Ehox and the mouse RIKEN clone NM-029203; (iv) human OTEX; and (v) human ESXR1 and mouse Esx1. In the phylogenic tree, OTEX is more distant from Pem than Psx1 and Psx2 are. On the other hand, OTEX shares with Pem the presence of a lysine residue at HD position 50, and has a tissue-expression profile closely resembling that of Pem. In addition, both the *OTEX* and *Pem* genes are under androgen control, and are localized on syntenic human and mouse regions of chromosome X. Taken together, this suggests that they may have comparable functions, despite the low sequence conservation. Indeed, the *PEPP* genes have a high propensity to diverge [30], which is also documented by the limited sequence identity seen outside the HD. This suggests that they may engage interactions with different protein partners to modulate biological events. A rapid evolution of proteins of the reproductive system has been demonstrated in other species [31].

Possible function of OTEX

On the basis of its sequence similarity to homeobox proteins, especially in the HD domain, OTEX is likely to act as a transcription factor. An exclusively nuclear localization was identified from our immunofluorescence experiments, and mutational analysis enabled us to identify the region responsible for nuclear transport. Several residues of this region were maintained among most PEPP proteins, suggesting that each possessed an NLS at the C-terminal end of their HD. This does not necessarily imply that all of them are nuclear proteins; Esx1, for example, was primarily detected in the cytoplasm, owing to the presence of an inhibitory motif [32]. On the other hand, experimental evidence shows that a functional NLS does exist at the C-terminal end of the HD of the human PDX-1 and mouse Crx homeobox proteins [33,34].

No inherent transactivation activity was determined for OTEX in a one-hybrid assay, suggesting that additional factors may be necessary. Similar to Esx1 [32], no evidence for homodimer formation of OTEX was obtained. We cannot, however, entirely rule out this possibility, because the presence of the DNA-binding element might be necessary for this to take place. Alternatively, a heterodimeric partner or an additional stimulus (e.g. phosphorylation event) might be essential. Another possibility is that OTEX is a transcriptional repressor, as described for other paired-like homeobox proteins [35,36]. Finally, the function of OTEX might have been impaired by fusing it to another protein domain. Additional studies are under way in order to identify the DNA response elements recognized by OTEX.

The remarkably narrow tissue distribution of *OTEX* in steroid hormone target organs points to a role in reproduction. *Pem*, *Esx1* and *ESXR1* are all expressed in the testis, as is *THG1* (C. Geserick and B. Haendler, unpublished work), and *Pem* is expressed additionally in the epididymis. Indeed, *Pem* was one of the first androgen-regulated genes coding for a transcription factor to be described, and its promoter harbours unusual, highly androgen-selective DNA response elements [23,37]. We found *OTEX* expression also to be under androgenic control in a prostate carcinoma cell line. *OTEX* and *Pem* transcripts are additionally found in the ovary and, interestingly, a form of premature ovarian failure ('POF1') maps to the chromosomal region where the three human *PEPP* genes are located [38]. The expression of *OTEX* in the ovary and mammary gland suggests furthermore that female sexual hormones may also be involved in the control of this gene.

The generation of mice deficient in *Pem* or *Psx2* did not provide further evidence as to the role of these proteins, suggesting subtle phenotypic effects that cannot be observed under usual rearing conditions. Alternatively, this might point to redundant functions, especially since some *PEPP* members have overlapping expression patterns. Redundancy has already been described for homeobox genes such as *Alx3* and *Alx4*, *Nkx2.5* and *Nkx2.6*, and *Hoxa-9* and *Hoxd-9*, for which severe phenotypic alterations have only been detected in double knock-out mice [39–41]. A detailed expression analysis in gonads using *in situ* hybridization or immunohistochemistry, and the identification of the target genes, should help to clarify the function of OTEX in human reproductive processes.

We thank Dr A. Cato (Forschungszentrum Karlsruhe, Germany) for the PC-3/ARwt cells, Dr K. Barbulescu for RNA from PC-3 and PC-3/ARwt cells, and Dr B. Kreft, A. Gunkel and J. Beschorner (Schering AG, Berlin, Germany) for help with fluorescence microscopy. The expert technical assistance of J. Wätzold and M. Preuss is much appreciated. This work was financed in part by grant 0310681B from the German Federal Ministry of Education and Research (BMBF). Responsibility for the content of this publication resides solely with the authors.

REFERENCES

1 Treisman, J., Harris, E., Wilson, D. and Desplan, C. (1992) The homeodomain: a new face for the helix-turn-helix? BioEssays **14**, 145–150
2 Gehring, W. J., Affolter, M. and Bürglin, T. (1994) Homeodomain proteins. Annu. Rev. Biochem. **63**, 487–526
3 Torres, M., Gomez-Pardo, E., Dressler, G. R. and Gruss, P. (1995) Pax-2 controls multiple steps of urogenital development. Development **121**, 4057–4065

4 Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L. et al. (2000) The LIM homeobox gene *Lhx9* is essential for mouse gonad formation. Nature (London) **403**, 909–913

5 Satokata, I., Benson, G. and Maas, R. (1995) Sexually dimorphic sterility phenotypes in *Hoxa10*-deficient mice. Nature (London) **374**, 460–463

6 Gendron, R. L., Paradis, H., Hsieh-Li, H. M., Lee, D. W., Potter, S. S. and Markoff, E. (1997) Abnormal uterine stromal and glandular function associated with maternal reproductive defects in *Hoxa11* null mice. Biol. Reprod. **56**, 1097–1105

7 Galliot, B., de Vargas, C. and Miller, D. (1999) Evolution of homeobox genes: Q_{50} paired-like genes founded the paired class. Dev. Genes Evol. **209**, 186–197

8 Maiti, S., Meistrich, M. L., Wilson, G., Shetty, G., Marcelli, M., McPhaul, M. J., Morris, P. L. and Wilkinson, M. F. (2001) Irradiation selectively inhibits expression from the androgen-dependent Pem homeobox gene promoter in Sertoli cells. Endocrinology (Baltimore) **142**, 1567–1577

9 Sasaki, A. W., Doskow, J., MacLeod, C. L., Rogers, M. B., Gudas, L. J. and Wilkinson, M. F. (1991) The oncofetal gene *Pem* encodes a homeobox and is regulated in primordial and pre-muscle stem cells. Mech. Dev. **34**, 155–164

10 Maiti, S., Doskow, J., Li, S., Nhim, R. P., Lindsey, J. S. and Wilkinson, M. F. (1996) The *Pem* homeobox gene. Androgen-dependent and -independent promoters and tissue-specific alternative RNA splicing. J. Biol. Chem. **271**, 17536–17546

11 Li, Y., Lemaire, P. and Behringer, R. R. (1997) *Esx1*, a novel X chromosome-linked homeobox gene expressed in mouse extraembryonic tissues and male germ cells. Dev. Biol. **188**, 85–95

12 Branford, W. W., Zhao, G.-Q., Valerius, M. T., Weinstein, M., Birkenmeier, E. H., Rowe, L. B. and Potter, S. S. (1997) *Spx1*, a novel X-linked homeobox gene expressed during spermatogenesis. Mech. Dev. **65**, 87–98

13 Fohn, L. E. and Behringer, R. R. (2001) *ESX1L*, a novel X chromosome-linked human homeobox gene expressed in the placenta and testis. Genomics **74**, 105–108

14 Han, Y. J., Park, A. R., Sung, D. Y. and Chun, J. Y. (1998) *Psx*, a novel murine homeobox gene expressed in placenta. Gene **207**, 159–166

15 Han, Y. J., Lee, Y. H. and Chun, J. Y. (2000) Identification and characterization of *Psx-2*, a novel member of the *Psx* (placenta-specific homeobox) family. Gene **241**, 149–155

16 Takasaki, N., McIsaac, R. and Dean, J. (2000) *Gpbox* (*Psx2*), a homeobox gene preferentially expressed in female germ cells at the onset of sexual dimorphism in mice. Dev. Biol. **223**, 181–193

17 Pitman, J. L., Lin, T. P., Kleeman, J. E., Erickson, G. F. and McLeod, C. L. (1998) Normal reproductive and macrophage function in Pem homeobox-deficient mice. Dev. Biol. **202**, 196–214

18 Takasaki, N., Rankin, T. and Dean, J. (2001) Normal gonadal development in mice lacking GPBOX, a homeobox protein expressed in germ cells at the onset of sexual dimorphism. Mol. Cell. Biol. **21**, 8197–8202

19 Li, Y. and Behringer, R. R. (1998) *Esx1* is an X-chromosome-imprinted regulator of placental development and fetal growth. Nat. Genet. **20**, 309–311

20 Miller, D. J., Hayward, D. C., Reece-Hoyes, J. S., Scholten, I., Catmull, J., Gehring, W. J., Callaerts, P., Larsen, J. E. and Ball, E. E. (2000) *Pax* gene diversity in the basal cnidarian *Acropora millepora* (Cnidaria, Anthozoa): implications for the evolution of the *Pax* gene family. Proc. Natl. Acad. Sci. U.S.A. **97**, 4475–4480

21 Breathnach, R. and Chambon, P. (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**, 349–383

22 Lo, K. and Smale, S. T. (1996) Generality of a functional initiator consensus sequence. Gene **182**, 13–22

23 Barbulescu, K., Geserick, C., Schüttke, I., Schleuning, W.-D. and Haendler, B. (2001) New androgen response elements in the murine Pem promoter mediate selective transactivation. Mol. Endocrinol. **15**, 1803–1816

24 Jans, D. A., Xiao, C. Y. and Lam, M. H. C. (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? BioEssays **22**, 532–544

25 Boulikas, T. (1994) Putative nuclear localization signals (NLS) in protein transcription factors. J. Cell. Biochem. **55**, 32–58

26 Zellmer, E., Zhang, Z., Greco, D., Rhodes, J., Cassel, S. and Lewis, E. J. (1995) A homeodomain protein selectively expressed in noradrenergic tissue regulates transcription of neurotransmitter biosynthetic genes. J. Neurosci. **15**, 8109–8120

27 Wimmer, Z., Zhu, X. X., Rouillard, J. M., Ambros, P. F., Lamb, B. J., Kuick, R., Eckart, M., Weinhausl, A., Fonatsch, C. and Hanash, S. M. (2002) Combined restriction landmark genomic scanning and virtual genome scans identify a novel human homeobox gene, ALX3, that is hypermethylated in neuroblastoma. Genes Chromosomes Cancer **33**, 285–294

28 Wuyts, W., Cleiren, E., Homfray, T., Rasore-Quartino, A., Vanhoenacker, F. and Van Hul, W. (2000) The ALX4 homeobox gene is mutated in patients with ossification defects of the skull (foramina parietalia permagna, OMIM 168500). J. Med. Genet. **37**, 916–920

29 Rao, E., Weiss, B., Fukami, M., Rump, A., Niesler, B., Mertz, A., Muroya, K., Binder, G., Kirsch, S., Winkelmann, M. et al. (1997) Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. Nat. Genet. **16**, 54–62

30 Maiti, S., Doskow, J., Sutton, K., Nhim, R. P., Lawlor, D. A., Levan, K., Lindsey, J. S. and Wilkinson, M. F. (1996) The *Pem* homeobox gene: rapid evolution of the homeodomain, X chromosomal localization, and expression in reproductive tissues. Genomics **34**, 304–316

31 Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., Wolfner, M. F. and Aquadro, C. F. (2001) Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. Proc. Natl. Acad. Sci. U.S.A. **98**, 7375–7379

32 Yan, Y.-T., Stein, S. M., Ding, J., Shen, M. M. and Abate-Shen, C. (2000) A novel PF/PN motif inhibits nuclear localization and DNA binding activity of the ESX1 homeoprotein. Mol. Cell. Biol. **20**, 661–671

33 Moede, T., Leibiger, B., Pour, H. G., Berggren, P.-O. and Leibiger, I. B. (1999) Identification of a nuclear localization signal, RRMKWKK, in the homeodomain transcription factor PDX-1. FEBS Lett. **461**, 229–234

34 Fey, Y. and Hugues, T. E. (2000) Nuclear trafficking of photoreceptor protein crx: the targeting sequence and pathologic implications. Invest. Ophthalmol. Vis. Sci. **41**, 2849–2856

35 Norris, R. A. and Kern, M. J. (2001) Identification of domains mediating transcription activation, repression, and inhibition in the paired-related homeobox protein, Prx2 (S8). DNA Cell Biol. **20**, 89–99

36 Dasen, J. S., Barbera, J. P., Herman, T. S., Connell, S. O., Olson, L., Ju, B., Tollkuhn, J., Baek, S. H., Rose, D. W. and Rosenfeld, M. G. (2001) Temporal regulation of a paired-like homeodomain repressor/TLE co-repressor complex and a related activator is required for pituitary organogenesis. Genes Dev. **15**, 3193–3207

37 Haendler, B. (2002) Androgen-selective gene regulation in the prostate. Biomed. Pharmacother. **56**, 78–83

38 Krauss, C. M., Turksoy, R. N., Atkins, L., McLaughlin, C., Brown, L. G. and Page, D. C. (1987) Familial premature ovarian failure due to an interstitial deletion of the long arm of the X chromosome. N. Engl. J. Med. **317**, 125–131

39 Beverdam, A., Brouwer, A., Reijnen, M., Korving, J. and Meijlink, F. (2001) Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice. Development **128**, 3975–3986

40 Fromental-Ramin, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C., Dierich, A., Dollé, P. and Chambon, P. (1996) Specific and redundant functions of the paralogous *Hoxa-9* and *Hoxd-9* genes in forelimb and axial skeleton patterning. Development **122**, 461–472

41 Tanaka, M., Schinke, M., Liao, H. S., Yamasaki, N. and Izumo, S. (2001) Nkx2.5 and Nkx2.6, homologs of *Drosophila tinman*, are required for development of the pharynx. Mol. Cell. Biol. **21**, 4391–4398

Received 11 March 2002/25 April 2002; accepted 30 April 2002
Published as BJ Immediate Publication 30 April 2002, DOI 10.1042/BJ20020399