*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 4987–4991, May 1995
Genetics

# Isolation and characterization of a *MAGE* gene family in the Xp21.3 region

**(X chromosome/dosage-sensitive sex reversal locus)**

F. MUSCATELLI*, A. P. WALKER*†, E. DE PLAEN‡, A. N. STAFFORD*, AND A. P. MONACO*§

*Imperial Cancer Research Fund Laboratories, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom; and ‡Ludwig Institute for Cancer Research, Brussels Branch, 74 avenue Hippocrate, B-1200 Brussels, Belgium

*Communicated by David J. Weatherall, University of Oxford, Oxford, United Kingdom, January 17, 1995*

**ABSTRACT** **A human gene with strong homology to the *MAGE* gene family located in Xq27–qter has been isolated by using exon-trapping of cosmids in the Xp21.3 region. We have mapped and sequenced cDNA and genomic clones corresponding to this gene, *MAGE-Xp*, and shown that the last exon contains the open reading frame and is present in a minimum of five copies in a 30-kb interval. *MAGE-Xp* is expressed only in testis and, unlike the Xq27–qter *MAGE* genes, it is not expressed in any of 12 different tumor tissues tested. However, the gene and predicted protein structure are conserved, suggesting a similar function. *MAGE-Xp* is located in the 160-kb critical interval defined for the locus involved in sex determination within Xp21 and is 50 kb distal to the *DAX-1* gene, which is responsible for X-chromosome-linked adrenal hypoplasia congenita.**

In 1991, van der Bruggen *et al.* (1) identified a human gene, *MAGE-1*, that directs the expression of an antigen, MZ2-E, recognized on a human melanoma by autologous cytolytic T lymphocytes. MZ2-E is a nonapeptide encoded by *MAGE-1* associated with a major histocompatibility class I molecule, HLA-A1 (2). *MAGE-1* is expressed in other tumors of various histological types (1, 3–5) but normally only in testes (4).

By using cross-hybridization with *MAGE-1*, 11 closely related genes have been isolated and characterized (1, 6). The 12 *MAGE*-related genes have been located in Xq27-qter and shown to have similar genomic structures: a minimum of two exons, with the last exon containing the open reading frame (ORF) (6, 7). They form the *MAGE* gene family, which has probably arisen from successive gene duplications. The coding region is highly conserved, whereas variation is observed in the 5′ noncoding exons, suggesting differential transcriptional control. However, all the *MAGE* genes in Xq27–qter are expressed in a number of tumors of different histological types, where they could encode antigenic peptides separate from MZ2-E (8). In normal tissues, their expression is limited to testis and placenta (6). The function of the putative protein is not known, but the amino acid sequence appears to be conserved and related to the mouse necdin protein (6). Necdin is a protein expressed in neurally differentiated embryonal carcinoma cells, in the developing mouse brain during the early stages of neuronal generation and differentiation, and in neurons at advanced stages of differentiation (9, 10).

We report here the identification of a gene, *MAGE-Xp*, with homology to the *MAGE* gene family significant but not strong enough to be detected by cross-hybridization using a *MAGE* probe from the Xq27–qter region. We have isolated this gene by a positional cloning approach in the Xp21.3 critical interval defined for the adrenal hypoplasia congenita (*AHC*) and the dosage-sensitive sex reversal (*DSS*) X-chromosome-linked loci (11–13). The gene responsible for X-linked AHC and hypogonadotropic hypogonadism (HH) has recently been identified (14, 15), as *DAX-1*. No candidate gene has been shown to be responsible for the male-to-female sex reversal phenotype seen with duplications of this region (13). From a cosmid contig covering the *DSS* and *AHC* loci, we have isolated potential exons by using the exon amplification method (16). One of the exons amplified was used to screen on adult testis cDNA library, and cDNA clones corresponding to the *MAGE-Xp* gene were isolated and characterized.¶

## MATERIALS AND METHODS

**Yeast Artificial Chromosomes (YACs), Cosmids, and Patients.** YAC clones 7124 (ICRFy900E0274, 640 kb), 7130 (ICRFy900F0767, 680 kb), 7128 (ICRFy900A1283, 560 kb), and 7125 (ICRFy900BO543, 410 kb) hybridized with probe 4965-5, which as previously been mapped (15). Cosmids D5 and 4965 have been used to map *MAGE-Xp*. These cosmids and patients M.C., P.J., and A.M. have been previously reported (15).

All YAC, cosmid, and patient DNAs were isolated, digested with *Eco*RI, subjected to electrophoresis in 1% agarose gels, and alkali blotted as previously reported (17). Blots were prehybridized and hybridized in a 50% formamide solution at 42°C (18) with radioactively labeled probe inserts; they were washed twice for 20 min 1× SSC and twice in 0.5× SSC at 65°C and were exposed to x-ray film at −70°C.

**Exon Amplification, cDNA Library Screening, and DNA Sequencing.** The exon amplification protocol was as previously described (16, 17). The exon amplification product 36 was hybridized to replica filters (Hybond N+, Amersham) containing $1.2 \times 10^6$ plaques of a human adult testis cDNA library (λgt11; Clontech). Resulting duplicated positive clones were purified through two rounds of dilution platings and rescreening. Inserts were amplified by using primers specific to λgt11 sequences flanking the cloning site. Inserts were subcloned in the pAMP1 vector as described by the supplier (Clone Amp system, BRL).

Sequencing was performed on double-stranded pAMP1 DNA by the dideoxynucleotide method with Sequenase Version 2.0 (United States Biochemical), using primers flanking the polylinker or primers derived from the obtained sequence.

**Amplification of the 5′ cDNA End.** The 5′ RACE System was used, following the instructions of the supplier (BRL/Life Technologies). The anchored primer is 5′-*cctttgtttacctgaacctgctcctc*-3′ (exon 2). The resulting PCR product was cloned in

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: AHC, adrenal hypoplasia congenita; DSS, dosage-sensitive sex reversal; YAC, yeast artificial chromosome; ORF, open reading frame.

†Present address: Academic Department of Medicine, Royal Free Hospital, London NW3 2QG, United Kingdom.
§To whom reprint requests should be addressed.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. X82539; *Homo sapiens* mRNA for *MAGE-Xp*).

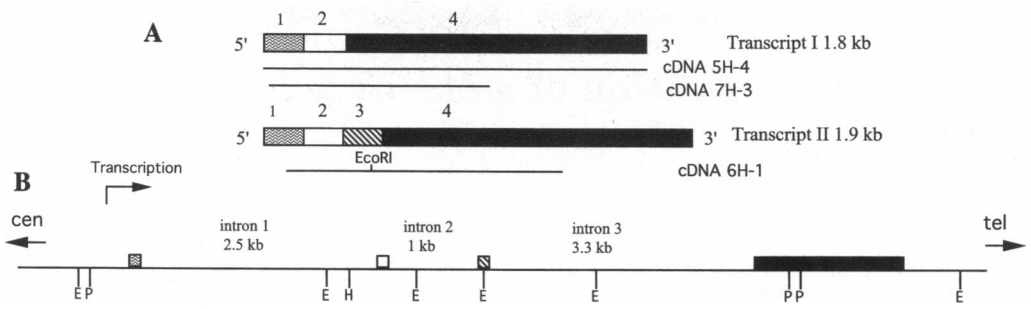


FIG. 1. Structure of *MAGE-Xp* cDNA clones and genomic locus. Two cDNAs (*A*) derived by differential splicing were characterized. The structure of the corresponding gene was established (*B*) from hybridization results on the cosmid contig and PCR products spanning the introns. E, *Eco*RI; P, *Pst* I; H, *Hin*dIII; cen, centromere; tel, telomere.

the pGEM-Tvector (Promega). Six clones have been sequenced as described above.

**Isolation of Exon- and Intron-Specific Probes.** Specific probes for each intron were produced by PCR amplification using exon primers flanking each intron as follows (in the 5′ → 3′ direction): intron 1, CAGTGTGGTGTCCAGCAGTGTCTC (sense) and CCTTTGTTTACCTGAACCTGCTCCTC (antisense); intron 2, GGACTGTGATCATATGAAGATCATCC (sense) and GAGAATTCACCCTCGGCTGACCAG (antisense); and intron 3, CTGGTCAGCCGAGGGTGAATTCTC (sense) and CAGGCAGGAAAGTGAGCAGAAAGG (antisense). Specific probes for exons were produced by using the following primers (5′ → 3′): exons 1 and 2, CAGTGTGGTGTCCAGCAGTGTCTC (sense) and CCTTTGTTTACCTGAACCTGCTCCTC (antisense); exon 2, GGACTGTGATCATATGAAGATCATCC (sense) and CCTTTGTTTACCTGAACCTGCTCCTC (antisense); exons 2 and 3, GGACTGTGATCATATGAAGATCATCC (sense) and GAGAATTCACCCTCGGCTGACCAG (antisense); and exon 4, CTGCCTTTCTGCTCACTTTCCTGC (sense) and TCTCTTCCTCATCTCTCAAAGCCTC (antisense). PCR amplification was performed in a reaction volume of 100 μl containing 1–10 ng of cosmid DNA, each primer at 100 pM, dNTPs (200 μM each), and 3% (vol/vol) dimethyl sulfoxide in the reaction buffer supplied by the manufacturer (Boehringer) with the addition of 1 unit of *Taq* DNA polymerase. Samples were denatured for 3 min before the enzyme was added and then amplified for 30 cycles of 45 sec at 95°C, 30 sec at 55°C, and 1 min at 72°C, followed by a final extension for 5 min at 72°C on a Perkin–Elmer/Cetus PCR machine.

**PCR Assays.** RNA purification and cDNA synthesis were performed as described (5). For amplification, 1/20 of the cDNA produced from 2 μg of total RNA was supplemented with 5 μl of PCR buffer (500 mM KCl/15 mM $MgCl_2$/1% Triton X-100/100 mM Tris·HCl, pH 8.8), 1 μl of 10 mM dNTPs, 25 pmol of each primer, 1.25 units of Dynazyme (Finnzymes, Helsinki), and water to a final volume of 50 μl. PCR was performed in a Trio-Thermoblock (Biometra, Göttingen, Germany) for 30 cycles (1 min at 94°C, 2 min at 63°C, and 3 min at 72°C). We used primers 5′-GTGGTGTCCAGCAGTGTCTC-3′ (sense, exon 1) and 5′-GTCAGATTCGGTACATGACACAG-3′ (antisense, in exon 4). When cDNA was amplified, we observed an abundant 446-bp fragment and a less abundant 531-bp fragment (including the alternative exon).

## RESULTS

**Isolation and Mapping of *MAGE-Xp*.** Exon amplification experiments (16) were performed on seven cosmids from contig A (15), which define the critical region of the *AHC* locus overlapping with the *DSS* locus. An exon-trap product (exon 2) was isolated and used to screen an adult human testis cDNA library. Forty positive hybridization signals were detected, 20 of which were purified and their inserts were isolated. Ten cDNA clones were partially sequenced and three clones (5H-4, 6H-1, and 7H-3) were completely sequenced. Sequence anal-

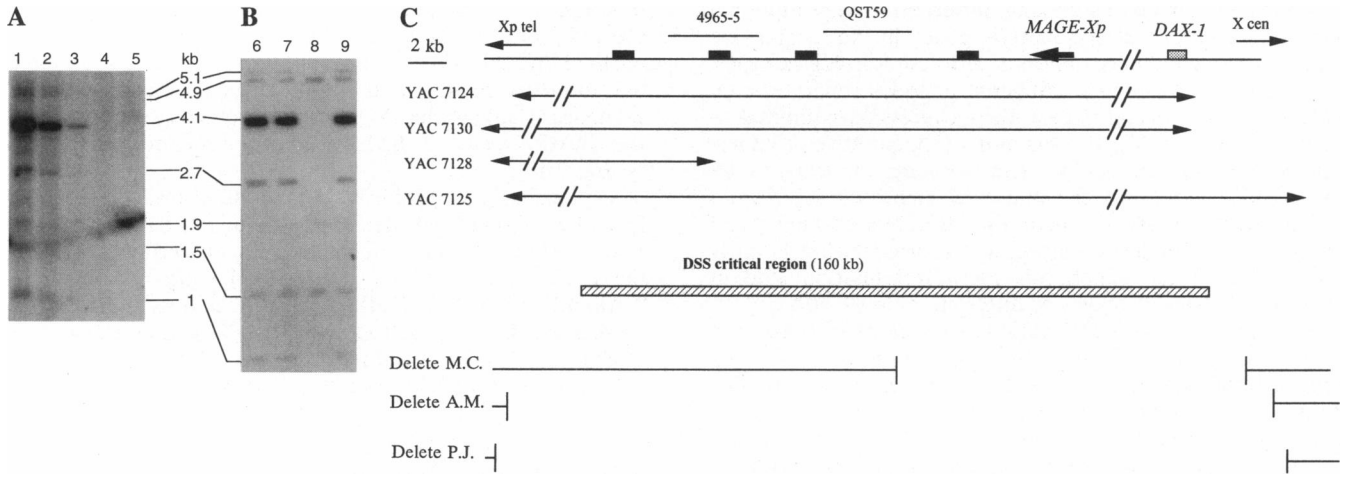


FIG. 2. (*A* and *B*) Hybridization of testis cDNA clone 7H-3 to human genomic DNA (*A*) and DNA isolated from YAC clones (*B*). (*A*) Samples digested with *Eco*RI: lane 1, normal male DNA; lane 2, normal female DNA; lanes 3–5, DNA from patients M.C. (3), A.M. (4), and P.J. (5), with deletions in Xp21.3 (15). (*B*) DNA isolated from YAC clones, covering the AHC-DSS locus in Xp21.3, digested with *Eco*RI: lane 6, YAC 7124; lane 7, YAC 7130; lane 8, YAC 7128; and lane 9, YAC 7130. (*C*) Schematic map of the genomic interval containing the *MAGE-Xp* gene (arrow) and its copies (rectangular filled boxes). Location is indicated for the probes 4965-5 (*DXS1074*), QST59 (*DXS319*), and the *DAX-1* gene, which is approximately 60 kb proximal to QST59. The DSS critical region and the deleted region in patients M.C., A.M., and P.J. are indicated.

ysis of these clones indicated the existence of two transcripts, 1.8 and 1.9 kb, the latter derived from differential splicing on an extra exon of 85 bp (Fig. 1*A*). We have defined the number and size of the introns by PCR on cloned genomic DNA. The gene contains four exons, one of which is differentially spliced; their sizes are 73, 105, 85, and 1603 bp, respectively. These exons are separated by three introns, 2.5, 1, and 3.3 kb, respectively. The size of the complete gene is 8.6 kb and a restriction enzyme site map for *Eco*RI, *Hin*dIII, and *Pst* I has been established (Fig. 1*B*). The proposed direction of transcription of this gene is from centromere to telomere, based on the map position of the 5′ and 3′ exons. This gene has been named *MAGE-Xp*, considering its sequence homology with the *MAGE* family of genes in Xq27–qter (see below).

Fig. 2 shows the hybridization pattern of cDNA 7H-3 with *Eco*RI digests of human genomic DNA samples and YAC DNA clones from the Xp21.3 region. Although the pattern of fragments is complex, we can assign the 4-, 2.7-, and 1-kb fragments to the genomic map of *MAGE-Xp*, since the *Eco*RI restriction map of *MAGE-Xp* is known from cosmids (Fig. 1*B*). The presence of additional hybridizing bands of 5.1, 4.8, 3, and 1.5 kb predicts multiple copies of *MAGE-Xp* in the Xp21.3 region. We hybridized, separately, the individual exons and introns, produced by PCR amplification, to the YAC and cosmid contigs constructed in this region (15). These hybridizations showed that the first, second, and third exons are present as single copies in this region (data not shown) and the fourth exon is present in a minimum of five copies. All the exon 4 copies are clustered in a 30-kb region containing the probes 4965-5 (*DXS1074*), QST59 (*DXS319*), and 332F (*DXS1020*). Therefore, the *MAGE-Xp* gene maps 10 kb proximal to the QST59 probe and 50 kb distal to the *DAX-1* gene, which is involved in AHC (14, 15) and is within the critical region for the *DSS* locus (13, 15) (Fig. 2*C*).

**Sequence Analysis of the *MAGE-Xp* cDNAs and the Genomic Exon–Intron Boundaries.** Three overlapping testis cDNA clones, 5H-4, 6H-1, and 7H-3, 1.7, 1.3, and 1.4 kb, respectively, have been mapped by restriction enzyme digestion and completely sequenced. cDNA clones 5H-4 and 7H-3 do not contain an *Eco*RI site and produce identical *Eco*RI hybridization patterns on human genomic DNA (as shown in Fig. 2). cDNA 6H-1 contains an *Eco*RI site and detects an additional 0.7-kb *Eco*RI fragment by hybridization on human genomic DNA and YAC DNA clones from the Xp21.3 region compared with cDNA clones 5H-4 and 7H-3 (data not shown). Sequence analysis showed that cDNA 6H-1 contains an extra exon of 85 bp between exons 2 and 4.

To obtain the full-length cDNA sequence, we used a rapid amplification of 5′ cDNA ends (5′ RACE) system (see *Materials and Methods*). The complete nucleotide sequence was assembled and is shown along with the predicted amino acid sequence in Fig. 3. It represents a total of 1866 nucleotides, including the extra 85-bp exon. A single ORF of 1041 nucleotides that encodes a predicted protein of 347 amino acids was identified. The AUG at nt position 324 was selected as the putative initiator because it is in good agreement with the Kozak consensus sequence (19) and it is the first AUG after the stop codon at nt position 198. Considering the two other reading frames, there are no significant ORFs. The structure of this gene is unusual in that the ORF is located in the fourth

```
                                                                GA GTG TTG CAA CTG GGC
CTG GCA TGT TTC AGC GTG GTG TCC AGC AGT GTC TCC CAC TCC TTG TGA AGT CTG AGG TTG
CAA AAG GAC TGT GAT CAT ATG AAG ATC ATC CAG GAG TAC AAC TCG AAA TTC TCA GAA AAC
AGG ACC TTG ATG TGA GAG GAG CAG GTT CAG GTA AAC AAA GGG CGA GGA CCC GAG CGA GCT
TAA GGC CAG TGG GGT GCA GCG TCT GGT CAG CCG AGG GTG AAT TCT CAG GAC TGG TCG GGA
GTC AAG GTG CCA CAT CTC CTG CCT TTC TGC TCA CTT TCC TGC CTG TTT TGC CTG ACC ACA
GCC ATC ATG CCT CGG GGT CAG AAG AGT AAG CTC CGT GCT CGT GAG AAA CGC CGC AAG GCG
          M   P   R   G   Q   K   S   K   L   R   A   R   E   K   R   R   K   A>
CGA GAG GAG ACC CAG GGT CTC AAG GTT CGT CAC GCC ACT GCA GCA GAG AAA GAG GAG TGC
 R   E   E   T   Q   G   L   K   V   R   H   A   T   A   A   E   K   E   E   C>
CCC TCC TCC TCT CCT GTT TTA GGG GAT ACT CCC ACA AGC TCC CCT GCT GCT GGC ATT CCC
 P   S   S   S   P   V   L   G   D   T   P   T   S   S   P   A   A   G   I   P>
CAG AAG CCT CAG GGA GCT CCA CCC ACC ACC ACT GCT GCT GCA GCT GTG TCA TGT ACC GAA
 Q   K   P   Q   G   A   P   P   T   T   T   A   A   A   A   V   S   C   T   E>
TCT GAC GAA GGT GCC AAA TGC CAA GGT GAG GAA AAT GCA AGT TTC TCC CAG GCC ACA ACA
 S   D   E   G   A   K   C   Q   G   E   E   N   A   S   F   S   Q   A   T   T>
TCC ACT GAG AGC TCA GTC AAA GAT CCT GTA GCC TGG GAG GCA GGA ATG CTG ATG CAC TTC
 S   T   E   S   S   V   K   D   P   V   A   W   E   A   G   M   L   M   H   F>
ATT CTA CGT AAG TAT AAA ATG AGA GAG CCC ATT ATG AAG GCA GAT ATG CTG AAG GTT GTT
 I   L   R   K   Y   K   M   R   E   P   I   M   K   A   D   M   L   K   V   V>
GAT GAA AAG TAC AAG GAT CAC TTC ACT GAG ATC CTC AAT GGA GCC TCT CGC CGC TTG GAG
 D   E   K   Y   K   D   H   F   T   E   I   L   N   G   A   S   R   R   L   E
CTC GTC TTT GGC CTT GAT TTG AAG GAA GAC AAC CCT AGT AGC CAC ACC TAC ACC CTC GTC
 L   V   F   G   L   D   L   K   E   D   N   P   S   S   H   T   Y   T   L   V>
AGT AAG CTA AAC CTC ACC AAT GAT GGA AAC CTG AGC AAT GAT TGG GAC TTT CCC AGG AAT
 S   K   L   N   L   T   N   D   G   N   L   S   N   D   W   D   F   P   R   N>
GGG CTT CTG ATG CCT CTC CTG GGT GTG ATC TTC TTA AAG GGC AAC TCT GCC ACC GAG GAA
 G   L   L   M   P   L   L   G   V   I   F   L   K   G   N   S   A   T   E   E>
GAG ATC TGG AAA TTC ATG AAT GTG TTG GGA GCC TAT GAT GGA GAG GAG CAC TTA ATC TAT
 E   I   W   K   F   M   N   V   L   G   A   Y   D   G   E   E   H   L   I   Y>
GGG GAA CCC CGT AAG TTC ATC ACC CAA GAT CTG GTG CAG GAA AAA TAT CTG AAG TAC GAG
 G   E   P   R   K   F   I   T   Q   D   L   V   Q   E   K   Y   L   K   Y   E>
CAG GTG CCC AAC AGT GAT CCC CCA CGC TAT CAA TTC CTA TGG GGT CCG AGA GCC TAT GCT
 Q   V   P   N   S   D   P   P   R   Y   Q   F   L   W   G   P   R   A   Y   A>
GAA ACC ACC AAG ATG AAA GTC CTC GAG TTT TTG GCC AAG ATG AAT GGT GCC ACT CCC CGT
 E   T   T   K   M   K   V   L   E   F   L   A   K   M   N   G   A   T   P   R>
GAC TTC CCA TCC CAT TAT GAA GAG GCT TTG AGA GAT GAG GAA GAG AGA GCC CAA GTC CGA
 D   F   P   S   H   Y   E   E   A   L   R   D   E   E   E   R   A   Q   V   R>
TCC AGT GTT AGA GCC AGG CGT CGC ACT ACT GCC ACG ACT TTT AGA GCG CGT TCT AGA GCC
 S   S   V   R   A   R   R   R   T   T   A   T   T   F   R   A   R   S   R   A>
CCA TTC AGC AGG TCC TCC CAC CCC ATG TGA GAA CTC AGG CAG ATT GTT CAC TTT GTT TTT
 P   F   S   R   S   S   H   P   M   _
GTG GCA AGA TGC CAA CCT TTT GAA GTA GTG AGC AGC CAA GAT ATG GCT AGA GAG ATC ATC
ATA TAT ATC TCC TTT GTG TTC CTG TTA AAC ATT AGT ATC TTT CAA GTG TTT TTC TTT TAA
TAG AAT GTT TAT TTA GAG TTG GGA TCT ATG TCT ATG AGC GAC ATG GAT CAC ACA TTT ATT
GGT GCT GCC AGC TTT AAG CAT AAG AGT TTT GAT ATT CTA TAT TTT TCA AAT CCT TGA ATC
TTT TTT GGG TTG AAG AAG AAG AAA GCA TAG CTT TAG AAT AGA GAT TTT CTC AGA AAT GTG
TGA AGA ACC TCA CAC AAC ATA ATT GGA GTC TTA AAA TAG AGG AAG AGT AAG CAA AGC ATG
TCA AGT TTT TGT TTT CTG CAT TCA GTT TTG TTT TTG TAA AAT CCA AAG ATA CAT ACC TGG
TTG TTT TTA GCC TTT TCA AGA ATG CAG ATA AAA TAA ATA GTA ATA AAT TAA AAA AAA AAA
AAA AA
```

FIG. 3. Nucleotide and deduced amino acid sequence of *MAGE-Xp*. The predicted start of translation is at nucleotide 324, in exon 3, with the first in-frame stop codon (TGA) at nucleotide 1366, yielding an ORF of 1041 bp that encodes a predicted protein of 347 amino acids. Splice junctions between different exons are indicated by arrowheads. The start and stop codons are underlined. Two consensus polyadenylylation signals are boxed in the 3′ untranslated region.

exon. The three first exons are not predicted to be translated. The 3′ untranslated region is 499 bp long and two consensus polyadenylylation signals, AAUAAA, were identified. Intron–exon boundaries were determined through sequence comparisons between cDNA clones and PCR products from genomic DNA spanning the introns.

**Homology with the *MAGE* Gene Family.** Nucleic acid or amino acid homology searches of the nonredundant data bases were done by using BLAST through the National Center for Biotechnology Information server (20). A significant similarity was found between the *MAGE-Xp* protein and the 11 MAGE proteins located in Xq27–qter (6). The strongest similarity is with *MAGE-10* (50% identity and 69% conservation; Fig. 4) and the lowest is the *MAGE-5* (26% identity, 43% conservation). However, we did not detect any cross-hybridization under normal stringency conditions with the *MAGE* genes from Xq when cDNA 7H-3, 5H-4, or 6H-1 was used (data not shown).

The typical structure of the *MAGE* genes (6), characterized by the presence of three or four exons with the coding region in the last exon, is also conserved in *MAGE-Xp*. The presence of multiple copies of the last exon detected in Xp21.3 reflects a similar organization to the *MAGE* family of genes in Xq27–qter.

**Expression Pattern of *MAGE-Xp*.** The expression of *MAGE-Xp* was first tested by Northern blot analysis using the 5H-4 cDNA clone. An mRNA transcript of 1.8–1.9 kb was detected only in human adult testis and not in adult peripheral blood, colon, small intestine, ovary, prostate, thymus, or spleen RNA (Fig. 5*A*); no expression was detected in adult human heart, brain, placenta, lung, liver, skeletal muscle, kidney, or pancreas and human fetal heart, brain, lung, liver, or kidney tissues (data not shown). Because of the strong similarity with the *MAGE* gene family and the presence of multiple copies of exon 4 in the Xp21.3 region, we tested the expression of *MAGE-Xp* by reverse transcriptase PCR using a primer specific for the 5′ end of *MAGE-Xp*, which is unique in the genome (Fig. 5*B*). Again, we observed expression only in human testis and no expression in stomach, breast, muscle, uterus, thymocytes, and lymphocytes. In adult testis, we were able to detect both transcripts with a lower level of the transcript containing the extra exon (estimated 10–20%). A lower expression of the smaller transcript was detected in fetal testis. This expression pattern is similar to that of the *MAGE* family of genes (6) in normal tissues. We also tested the expression of *MAGE-Xp* in a panel of tumors including four colon carcinomas, six leukemias, seven melanomas, four lung cancers, three sarcomas, two breast carcinomas, four kidney tumors, six seminomas, two neuroblastomas, two brain tumors, one stomach tumor, and two pharyngeal tumors. We did not find expression in any of the tumors tested. This reveals a major difference from the *MAGE* gene family in Xq27–qter, whose expression was detected in one or several kinds of tumors.

**The Putative *MAGE-Xp* Protein.** The predicted protein of *MAGE-Xp* is 347 amino acids, slightly longer than the other putative MAGE proteins, which are 309–319 amino acids, except *MAGE-10* protein, which is 369 amino acids. It is devoid of a signal sequence. *MAGE-Xp* protein has moderate homology (35% identity, 55% conservation) with the mouse protein necdin which is higher than the homology found between necdin and the other MAGE proteins. Hydrophobic cluster analysis (21) was used to compare the amino acid sequences of *MAGE-Xp* and *MAGE-1* proteins (data not shown). This analysis demonstrated a remarkable conservation of the main hydrophobic regions, including the putative transmembrane domain, suggesting a conservation of function with the *MAGE-Xq* and necdin proteins.

## DISCUSSION

**A Cluster of *MAGE*-Related Genes in Xp21.3.** The gene that we have isolated by exon-trapping in Xp21.3 is clearly related to the *MAGE* family of genes recently characterized and localized to Xq27–qter. We describe the presence of at least five copies of the last exon and translated portion of the *MAGE-Xp* transcript in a 30-kb interval, while the untranslated exons are unique. Because the testis cDNA library was screened by using the second untranslated exon, it remains to be seen if other copies are expressed with different 5′ untranslated ends. However, it seems that *MAGE-Xp* defines a second *MAGE* gene family, in Xp21.3, in which the genes are clustered, probably resulting from sequential duplications. There may be a parallel evolution of the two *MAGE* gene families similar to that of the immunoglobulin gene families.

**Sequence and Structure Homologies.** Although the similarity between *MAGE-Xp* and *MAGE-10* is deduced from amino acid sequence similarity, the gene structure, with two to three untranslated exons, is also well conserved. The role of these untranslated exons is not known, and only a few genes with a similar structure have been described (1). They might repre-

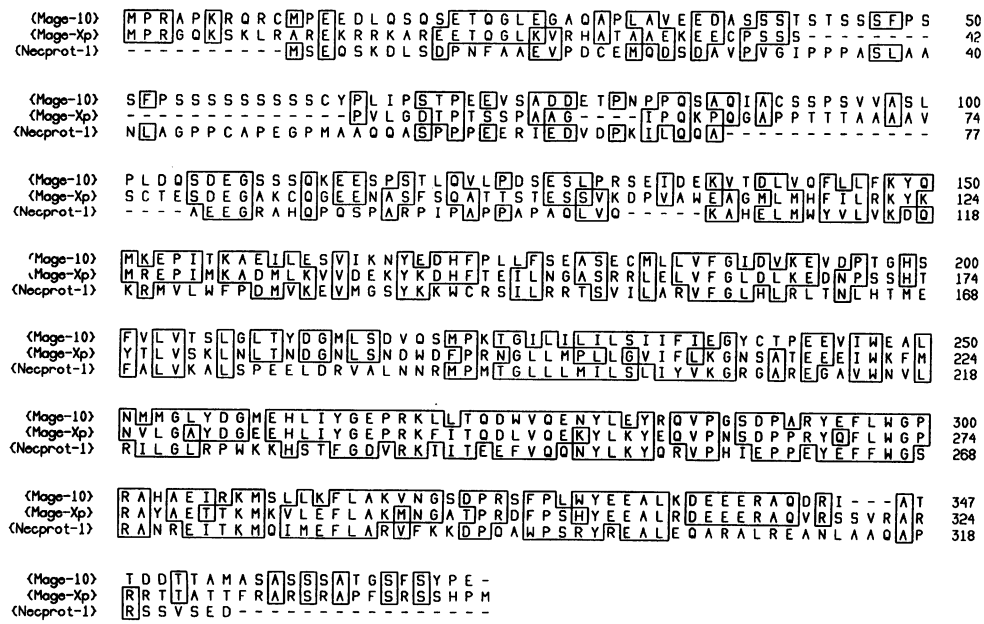


FIG. 4. Alignment of the predicted protein of human *MAGE-Xp* with human *MAGE-10* protein and the mouse Necdin protein (Necprot-1). The GCG Pretty Plot parameters were plurality 2.0; threshold 1.0; average weight 1.0; average match 0.54; average mismatch −0.40.

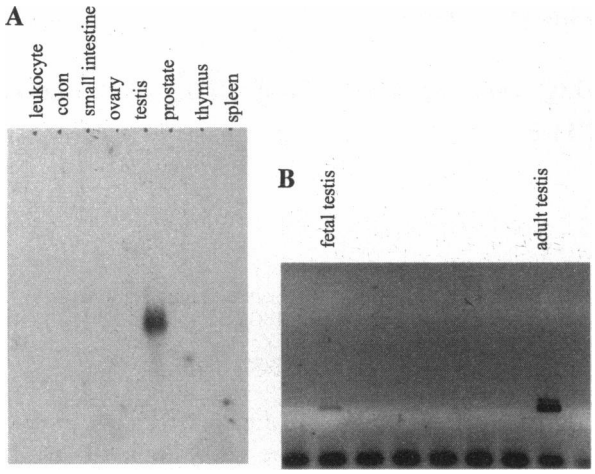


FIG. 5. (*A*) Hybridization of cDNA 5H-4 to a Northern blot of poly(A)$^+$ RNA from eight adult human tissues (Clontech, 2 μg per lane). The size of the transcript detected in testis is 1.8 kb. (*B*) Expression of *MAGE-Xp* in normal tissues detected by using PCR amplification (*Materials and Methods*). Expression is detected in adult and fetal testis.

sent translation control regions. The putative protein structure is also well conserved and suggests a conservation of function of the MAGE and mouse necdin proteins. *MAGE-Xp* has a stronger homology with necdin than the MAGE-Xq family does with necdin, an observation that could reflect the evolutionary process with *MAGE-Xp* as an ancestral gene to the *MAGE-Xq* genes. Recently, two mouse genes homologous to the human *MAGE* genes and located apparently in the same region have been isolated (25). It appears that *MAGE-Xp* has a stronger homology with these mouse genes than does *MAGE-Xq*. No function has yet been assigned to these proteins, but necdin is a nuclear protein expressed at different developmental stages of the mouse central nervous system.

**The Absence of *MAGE-Xp* Expression in Tumors.** All the *MAGE* genes described in Xq27–qter are switched on and up-regulated in tumors, and they are not expressed in normal tissues, with the exception of testis and placenta. They all have the potential to encode antigenic peptides which are combined with HLA class I molecules through a consensus motif (22, 23). The peptide–HLA complexes are presented at the cell surface and are recognized by cytotoxic T lymphocytes. They are potential tumor rejection antigens and have important implications for cancer immunotherapy. The mechanism whereby the tumor rejection antigens arise is unknown, and one hypothesis is the activation of a gene which is silent in normal tissues and for which no strong natural tolerance has been established (24).

*MAGE-Xp* is specifically expressed in testis, as are the other *MAGE* genes, but it is not expressed in any tumor samples tested. This difference may be important in understanding the molecular mechanism involved in the switching on and up-regulating of normally silent genes in tumors. This could be due to a different chromatin environment or the existence of different regulatory factors or promoters.

***MAGE-Xp* and the *DSS* Locus.** The *DSS* X-linked locus has been recently mapped to a 160-kb duplication interval including QST-59 (*DXS319*) and the *DAX-1* gene (13, 14). *MAGE-Xp* is located between these two probes and therefore is a potential candidate for the sex reversal phenotype when it is expressed in double dosage. This could be tested in mice by transgenesis of an extra copy of this gene or cluster of genes.

We have previously characterized some male patients with deletions of this region including *DAX-1* and the *MAGE-Xp* gene (15). These patients have been diagnosed as having AHC only, and their phenotype is the same as that observed in patients with point mutations in *DAX-1*. Therefore, it appears that the "knock-out" of *MAGE-Xp* is not associated with a particular phenotype. One cause of the absence of phenotype could be the existence of a redundant function between *MAGE-Xp* and the *MAGE* genes in Xq27–qter. If this is the case, only the "knock out" of both regions should result in a functional disturbance.

We thank Dr. O. De Backer for helpful discussion and Y. Ishikawa Brush for help with Fig. 4. This work was funded by the Imperial Cancer Research Fund and the International Human Frontiers Science program.

1. van der Bruggen, C., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A. & Boon, T. (1991) *Science* **254,** 1643–1647.
2. Traversari, C., van der Bruggen, P., Luescher, E., Lurquin, C., Chomez, P., Van Pel, A., De Plaen, E., Amar-Costesec, A. & Boon, T. (1992) *J. Exp. Med.* **176,** 1453–1457.
3. Brasseur, F., Marchand, M., Vanwijck, R., Herin, M., Lethe, B., Chomez, P. & Boon, T. (1992) *Int. J. Cancer* **52,** 839–841.
4. De Smet, C., Lurquin, C., van der Bruggen, P., De Plaen, E., Brasseur, F. & Boon, T. (1994) *Immunogenetics* **39,** 121–129.
5. Weynants, P., Lethé, B., Brasseur, F., Marchaud, M. & Boon, T. (1994) *Int. J. Cancer.* **56,** 826–829.
6. De Plaen, E., Arden, K., Traversari, K., Gaforio, J. J., Szikora, P., De Smet, C., Brasseur, F., van der Bruggen, P., Lethé, B., Lurquin, C., Brasseur, R., Chomez, P., De Backer, O., Cavenee, W. & Boon, T. (1994) *Immunogenetics* **40,** 360–369.
7. Oaks, M. K., Hanson, J. P. & O'Malley, D. P. (1994) *Cancer Res.* **1,** 1627–1629.
8. Gaugler, B., Van den Eynde, B., van der Bruggen, P., Romero, P., Gaforio, J. J., DePlaen, E., Lethé, B., Brasseur, F. & Boon, T. (1994) *J. Exp. Med.* **179,** 921–930.
9. Maruyama, K., Usami, M., Aizawa, T. & Yoshikawa, K. (1991) *Biochem. Biophys. Res. Commun.* **178,** 291–296.
10. Aizawa, T., Maruyama, K., Kondo, H. & Yoshikawa, K. (1992) *Dev. Brain Res.* **68,** 265–274.
11. Walker, A. P., Chelly, J., Ishikawa Brush, Y., Récan, D., Chaussain, J. L., Oley, C., Connor, J., Yates, J., Price, D., Super, M., Bottani, A., Steinman, B., Kaplan, J. C., Davies, K. & Monaco, A. P. (1992) *Hum. Mol. Genet.* **1,** 579–585.
12. Worley, K. C., Ellison, K. A., Zhang, Y. H., Wang, D. F., Mason, J., Roth, E. J., Adams, V., Fogt, D. D., Zhu, X. M., Towbin, J. A., Chinault, A. C., Zoghbi, H. & McCabe, E. R. B. (1993) *Genomics* **16,** 407–416.
13. Bardoni, B., Zanaria, E., Guioli, S., Floridia, G., Worley, K. C., Tonini, G., Ferrante, E., Chiumello, G., McCabe, E. R. B., Fraccaro, M. L., Zuffardi, O. & Camerino, G. (1994) *Nat. Genet.* **7,** 497–501.
14. Zanaria, E., Muscatelli, F., Bardoni, B., Strom, T. M., Guioli, S., Guio, W., Lalli, E., Moser, C., Walker, A. P., McCabe, E., Meitinger, T., Monaco, A. P., Sassone-Corsi, P. & Camerino, G. (1994) *Nature (London)* **372,** 635–641.
15. Muscatelli, F., Strom, T. M., Walker, A. P., Zanaria, E., Récan, D., Meindl, A., Bardoni, B., Guioli, S., Zehetner, G., Rabl, W., Schwarz, H. P., Kaplan, J. C., Camerino, G., Meitinger, T. & Monaco, A. P. (1994) *Nature (London)* **372,** 672–676.
16. Buckler, A. J., Chang, D. D., Graw, S. L., Brook, D. J., Haber, D. A., Sharp, P. A. & Housman, D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 4005–4009.
17. Walker, A. P., Muscatelli, F. & Monaco, A. P. (1993) *Hum. Mol. Genet.* **2,** 107–114.
18. Larin, Z., Monaco, A. P. & Lehrach, H. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 4123–4127.
19. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148.
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
21. Gaboriaud, C., Bissery, V., Benchetrit, T. & Mornon, J. P. (1987) *FEBS Lett.* **224,** 149–155.
22. Celis, E., Tsai, V., Crimi, C., DeMars, R., Wentworth, P. A., Chesnut, R. W., Grey, H. M., Sette, A. & Serra, H. M. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2105–2109.
23. Ding, M., Beck, R. J., Keller, J. & Fenton, R. G. (1994) *Biochem. Biophys. Res. Commun.* **202,** 549–555.
24. Boon, T., De Plaen, E., Lurquin, C., Van den Eynde, B., van der Bruggen, P., Traversari, C., Amar-Costesec, A. & Van Pel, A. (1992) *Cancer Surv.* **13,** 23–37.
25. De Backer, O., Verheyden, A.-M., Martin, B., Godelaine, D., De Plaen, E., Brasseur, R., Avner, P. & Boon, T. (1995) *Genomics*, in press.