[CANCER RESEARCH 61, 5544–5551, July 15, 2001]

An Overview of the *MAGE* Gene Family with the Identification of All Human Members of the Family[1]

Patrick Chomez,[2,3] Olivier De Backer,[2] Mathieu Bertrand, Etienne De Plaen, Thierry Boon, and Sophie Lucas

Ludwig Institute for Cancer Research, Brussels Branch, and Cellular Genetics Unit, Université Catholique de Louvain, B-1200 Brussels, Belgium

ABSTRACT

The first human members of the *MAGE* gene family that have been described are expressed in tumor cells but silent in normal adult tissues except in the male germ line. Hence, they encode strictly tumor-specific antigens that represent attractive targets for cancer immunotherapy. However, other members of the family were recently found to be expressed in normal cells, indicating that the family is larger and more disparate than initially expected. We therefore performed a database screening to identify all of the recorded members of both classes of human *MAGE* genes. This report provides an overview of the *MAGE* family and proposes a general nomenclature for all of the *MAGE* genes identified thus far. We found that the *MAGE-D* genes were particularly well conserved between man and mouse, suggesting that they exert important functions. In addition, the genomic structure of the *MAGE-D* genes indicates that one of them corresponds to the founder member of the family, and that all of the other *MAGE* genes are retrogenes derived from that common ancestral gene. Intriguingly, the COOH-terminal domain of MAGE-D3 was found to be identical to trophinin, a previously described protein believed to be involved in embryo implantation.

INTRODUCTION

The first member of the human *MAGE* family was identified as a gene encoding a tumor-specific antigen (1). This gene was later found to belong to a cluster of 12 *hMAGE-A* genes located in the q28 region of the X chromosome (2, 3). A sequencing effort directed at the human Xp21 region led to the discovery of a second cluster that was named *hMAGE-B* (4–6) and, more recently, a *hMAGE-C* cluster was identified in Xq26–27 (7, 8). The genes belonging to the *hMAGE-A*, *-B,* and *-C* subfamilies are characterized by a large terminal exon encoding the entire protein. They are completely silent in normal tissues, with the exception of male germ cells, and, for some of them, placenta. Some are expressed in tumor cells of various histological types, where they code for antigens recognized by cytolytic T lymphocytes (9). Because of their specific expression on tumor cells, these antigens are of particular interest for antitumor immunotherapy. Preliminary results of clinical trials suggest that tumor regression can be induced in a significant number of cancer patients by immunization with an antigen encoded by gene *hMAGE-A3* (10–13).

Two groups of *mMage* genes have been identified thus far in the mouse (14–16). Like their human counterparts, they are silent in normal adult tissues with the exception of male germ cells (17, 18), and some of them are expressed in tumor cells (14, 15). These murine genes were named *mMage-a* and *-b* because the sequences and isoelectric points of the corresponding proteins were closest to those of the human MAGE-A and -B proteins, respectively. However, the overall sequence identity between MAGE-A and -B orthologues is weak, and *MAGE-C* genes have not been identified in the mouse, implying that the members of these three subfamilies are poorly conserved during evolution.

More recently, we (19) and others (20) have reported the identification of two sequences that define a fourth subfamily of human *MAGE* genes, *hMAGE-D*. These genes differ from the previously described members of the family by their expression pattern: they are expressed in all normal tissues tested. They also differ by their genomic structure, the open reading frame of *hMAGE-D2* being split over 11 exons. Importantly, MAGE-D1 was recently found to interact with the p75 neurotrophin receptor and to facilitate nerve growth factor-dependent apoptosis (21). MAGE-D1 was also recently reported to interact with members of the Dlx/Msx homeodomain family and to regulate the transcriptional function of Dlx5 (22). These observations suggest that the members of the MAGE-D subfamily exert important functions and prompted us to systematically screen the public databases to identify all of the recorded members of the human *MAGE* family. In this paper, we report the results of this screening as well as the first detailed analysis of the murine *Mage-d* genes.

MATERIALS AND METHODS

Databases Searches. With the aim of identifying new human and murine genes belonging to the *MAGE* family, we performed tblastn searches in the DNA databases (nr, ests, htgs, and the working draft sequence of the human genome) available from the National Center for Biotechnology Information,[4] using the MAGE domain of protein MAGE-D2 as the query. We ran the gapped tblastn program (version 2.1.2), using the BLOSUM-62 substitution matrix and the default values for the gap costs (11 and 1; Ref. 23). All hits with expected values (E) <0.001 were recorded for additional analysis. Nucleotide sequences were aligned to constitute clusters corresponding to individual genes. These alignments were manually evaluated and a consensus sequence was derived for each of them. These sequences were then compared with the sequences of the known members of the *MAGE* family.

Multiple Sequence Alignment. Protein sequences of the MAGE conserved domains were aligned using CLUSTAL W neighbor-joining (24). Calculations were performed with the MacVector 6.5 package (Oxford Molecular Group, Oxford, United Kingdom) using the BLOSUM-30 substitution matrix and the default values for the gap costs (10 and 0.05).

Screening of a PAC Mouse Library. The large-insert RPCI21 PAC library of genomic DNA from 129/SvEvTacfBr female mice was obtained from Peter de Jong (Roswell Park Cancer Institute, Buffalo, NY). The high density filters were screened by hybridization with *mMage-d1* and *mMage-d2* cDNA probes. Six clones containing *mMage-d1*, seven clones containing *mMage-d2,* and four clones containing *mMage-d3* were identified.

Determination of the Genomic Structure of the *mMage-d* Genes. The location and size of the introns of *mMage-d1*, *-d2,* and *-d3* were determined by a combination of sequencing and PCR experiments performed on plasmid clones containing restriction fragments of these genes.

Expression Analyses. The expression of the newly identified *MAGE* genes in normal and tumoral tissues was evaluated by RT-PCR[5] using standard procedures. The primers were: 5′-AAAGAGCAACTGTGCCATTGG and 5′-ACTTTCATCTTACTGGTTTCAAG for *hMAGE-B7*; 5′-CAAGAGCAGA-GATGCAGATGA and 5′-GAGCACACACCCCTATTGCAT for *hMAGE-C4*; and 5′-CCAAGGACACTCCCAGGCTGA and 5′-CATGTTCCTCGGC-

Received 2/2/01; accepted 5/14/01.

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

[1] Supported by the Fonds National de la Recherche Scientifique, Brussels, Belgium (to S. L.) and a fellowship from the Fonds pour la formation à la Recherche dans l'Industrie et l'Agriculture (to M. B.).

[2] These authors contributed equally to this work.

[3] To whom requests for reprints should be addressed, at Ludwig Institute for Cancer Research, Avenue Hippocrate, 74-UCL 74. 59, B-1200 Brussels, Belgium. E-mail: patrick.chomez@bru.licr.org.

[4] Internet address: http://www.ncbi.nlm.nih.gov/.

[5] The abbreviations used are: RT-PCR, reverse transcription-PCR; RACE, rapid amplification of cDNA ends.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010505544.pdf by guest on 26 February 2024

CATATCCA for *hMAGE-D4*. Sequences of the primers used to specifically amplify the other newly identified *MAGE* genes are available on request. A mouse poly(A)$^+$ RNA dot blot (Clontech, Palo Alto, CA) was hybridized with ^{32}P-labeled probes specific for *mMage-d1* (probe 1), *mMage-d2* (probe 2), or *mMage-d3* (probes 3a and 3c). The probes were PCR fragments obtained from cDNA with the following primers: probe 1, sense primer 5′-TGACTGGACT-GCACAGTTC and antisense primer 5′-GCATGCCACTCTCAGTCAA-CAGG; probe 2, sense primer 5′-AGGATCCCAAGGAATGGGCAG and antisense primer 5′-TCACTTGTAGGAGAAACCACAG; probe 3a, sense primer 5′-GACCACAAATACTGACAATG and antisense primer 5′- GGAA-GAAGGGTAACAATA; and probe 3c, sense primer 5′-ACTGCCTAA-CAAGGGAAGAG and antisense primer 5′-CCCAGTTCTATTG TTGGCTT. Radioactive signals were quantified by a phosphorimager analysis. A Northern blot of total mouse brain RNA was hybridized with probe 3b, which is a PCR fragment corresponding to the Mage-conserved domain of *mMage-d3* obtained with sense primer 5′-GTTGGTGAAATACCTGTTGG and antisense primer 5′-CGAGACTAGCAAGATGAAAGTC.

RACE Analyses. The 5′ ends of *mMage-d1*, *-d2,* and *-d3* cDNAs were amplified by PCR from mouse brain mRNA using the Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA). The 5′ ends of trophinin transcripts present in human endometrium were amplified using the 5′ RACE system from Life Technologies, Inc. Total RNA was extracted with Tripure (Roche Molecular Biochemicals) from human endometrium dissected from a surgical sample obtained at the early secretory phase. Poly(A)$^+$ RNA was purified with the mRNA Isolation kit from Roche Molecular Biochemicals. cDNA synthesis was primed with the antisense primer 5′-TACAAGGCAT-GCCACCAAAGC, and two successive rounds of PCR amplification were performed using antisense primers 5′-AAATCTGCTCCAGGCCTGAG and 5′-AACTCTTCCCTTGTTAGGC, respectively. Then the amplified products were cloned into pCR2.1 (Invitrogen, Carlsbad, CA). The clones containing sequences of *mMage-d3* exon 11 were identified by hybridization with oligonucleotide 5′-GACTTTCATCTTGCTAGTCTCG. The clones which did not hybridize were sequenced to determine their 5′ end.

RESULTS

Human *MAGE* Genes We performed tblastn homology searches (23) with the hMAGE-D2 protein sequence (19) to identify all of the human *MAGE* sequences recorded in the databases of the National Center for Biotechnology Information. Most of the new *MAGE* sequences that were retrieved could be grouped in contigs. The new genes were classified on the basis of their sequence homologies and their chromosomal location (Fig. 1). Some were found to belong to previously published *hMAGE* subfamilies (*hMAGE-A13* to *A15*; *hMAGE-B7* to *B17*; *hMAGE-C4* to *C7 and hMAGE-D3* and *-D4*), whereas others defined new subfamilies that were named *hMAGE-E* to *-K* (Table 1).

A RT-PCR analysis of the expression of all these genes was performed on panels of normal and tumoral tissue samples (Fig. 2 and data not shown). Some were found to be silent in all tissues tested, including testis. These genes are probably pseudogenes comparable with the previously described *hMAGE-A7* sequence (2) and are indicated by a "P" in the expression column of Table 1. Genes *hMAGE-B10, hMAGE-B16, hMAGE-B17,* and *hMAGE-C4* displayed the expression pattern characteristic of the *MAGE-A*, *-B,* and *-C* genes, *i.e.,* silent in all normal tissues with the exception of testis. In addition, *hMAGE-C4* was found to be expressed in a minor proportion of tumoral samples, suggesting that it could potentially encode tumor-specific antigens. By contrast, all *hMAGE-D, hMAGE-E, hMAGE-F, hMAGE-G,* and *hMAGE-H* genes were found to be expressed at various levels in many normal tissues, a "ubiquitous" expression pattern resembling that reported for *hMAGE-D1* and *-D2* (19, 20).

By comparing the sequences of the putative proteins encoded by these genes, we identified a stretch of ~200 amino acids which we named the "MAGE conserved domain." This domain corresponds to the only region of homology shared by all of the members of the

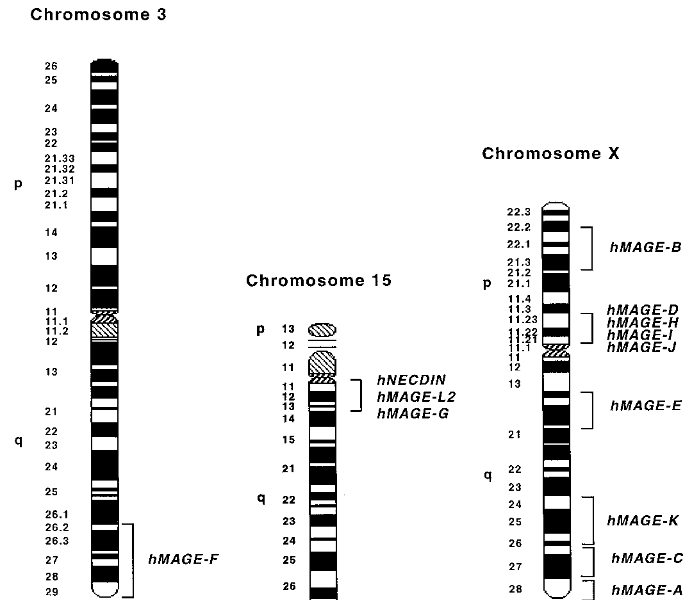


Fig. 1. Chromosomal locations of the human *MAGE* subfamilies. Location of the newly identified *hMAGE* genes were found in the database entries of genomic clones containing these sequences or by searching the *Homo sapiens* genome view of the National Center for Biotechnology Information (see Table 1 for accession numbers of clones and contigs). *hMAGE-H1* corresponds to est NIB1132 (GenBank accession no. T16329) mapped by Miller and Willard (37).

family (Fig. 3*A*). The rest of the protein sequences, and in particular the NH_2-terminal domains, are completely different from one subfamily to the other (Fig. 4). The MAGE conserved domain is usually located close to the COOH termini of the proteins except in the hMAGE-D proteins, where it occupies a more central position. Intriguingly, the MAGE conserved domain is duplicated in hMAGE-E1 and hMAGE-E2 (Fig. 4).

Murine *Mage* Genes. By database screening, we also retrieved a large number of murine *Mage* genes, most of which could be identified as orthologues of human *MAGE* sequences (Table 1). As shown in Fig. 5, the conservation of the MAGE domains between human and mouse differs considerably from one protein to the other. However, these domains are remarkably similar in the two species for the MAGE-D proteins (up to 99% amino acid identity between human and mouse for MAGE-D2; Fig. 3*B*), and the three MAGE-D domains found in mouse are closer to their respective human orthologues than they are to each other (Fig. 5). These observations suggested that the MAGE-D proteins exert important and distinct functions and prompted us to analyze this subfamily in more detail.

The open reading frames of *mMage-d1*, *-d2,* and *-d3* were amplified by RT-PCR from adult mouse brain total RNA and used as probes to isolate the corresponding genomic clones from a mouse PAC library. In addition, we also screened the same library with a human *MAGE-D4* probe, but we only retrieved clones containing *mMage-d1* and *-d3* sequences, strongly suggesting that the mouse genome does not contain a *hMAGE-D4* orthologue. The murine *Mage-d* genes all displayed a genomic structure that closely resembles that observed for the human *MAGE-D2* gene (19). Each gene contains 13 exons, the main open reading frame covering exons 2 to 12, and most of the MAGE conserved domain being encoded by exons 5 to 11 (Fig. 6*A*). Importantly, the existence of multiple exons encoding the MAGE protein seems to be a feature unique to the *MAGE-D* subfamily. Indeed, all of the other human and mouse genes that we identified yielded PCR products of identical sizes when amplified from either cDNA or genomic DNA (Fig. 2 and data not shown), suggesting that the sequence encoding their MAGE conserved domain was entirely

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010005544.pdf by guest on 26 February 2024

Table 1 *Known human and murine members of the MAGE gene family*

Subfamily	Human			Mouse		
	Gene name	Reference/Accession no.[a]	Expression[b]	Gene name	Reference/Accession no.[a]	Expression[b]
MAGE-A[c]	*hMAGE-A1*		T	*mMage-a1*		ND
	hMAGE-A2		T	*mMage-a2*		T
	hMAGE-A3		T	*mMage-a3*		T
	hMAGE-A4		T	*mMage-a4*		T
	hMAGE-A5		T	*mMage-a5*	Ref.14	T
	hMAGE-A6	Ref. 2	T	*mMage-a6*		ND
	hMAGE-A7		P	*mMage-a7*		P
	hMAGE-A8		T	*mMage-a8*		T
	hMAGE-A9		T	*mMage-a9*	AL021127.2$_{(9458-10189)}$	P
	hMAGE-A10		T			
	hMAGE-A11		T			
	hMAGE-A12		T			
	hMAGE-A13	U71148$_{(11970-12608)}$	P			
	hMAGE-A14	NT_011534.1$_{(86877-86203)}$	P			
	hMAGE-A15	NT_025309.1$_{(68501-69500)}$	P[d]			
MAGE-B[c]	*hMAGE-B1*		T	*mMage-b1*		T
	hMAGE-B2	Ref. 5	T	*mMage-b2*	Ref.15	T
	hMAGE-B3		T	*mMage-b3*		T
	hMAGE-B4		T	*mMage-b4*	Ref.16	T
	hMAGE-B5	Ref. 8	T	*mMage-b5*	AC025874.3$_{(174095-173508)}$	T
	hMAGE-B6		T	*mMage-b6*	AC025874.3$_{(99563-94976)}$	ND
	hMAGE-B7	AC005297.1$_{(117306-116890)}$	P	*mMage-b7*	AC025874.3$_{(154024-154611)}$	ND
	hMAGE-B8	AC005297.1$_{(141400-140879)}$	P	*mMage-b8*	AC025874.3$_{(120059-120649)}$	ND
	hMAGE-B9	AC005297.1$_{(171013-170783)}$	P	*mMage-b9*	AC025874.3$_{(164334-164849)}$	ND
	hMAGE-B10	AC011693.5$_{(134036-135079)}$	T	*mMage-b10*	AC025874.3$_{(49446-49961)}$	ND
	hMAGE-B11	AC011693.5$_{(124613-125653)}$	P[e]			
	hMAGE-B12	AC011693.5$_{(116519-117561)}$	P[e]			
	hMAGE-B13	AC011693.5$_{(99610-98559)}$	P			
	hMAGE-B14	NT_025279.3$_{(927601-928600)}$	P[d]			
	hMAGE-B15	NT_011752.1$_{(191672-191259)}$	P[d]			
	hMAGE-B16	NT_025279.3$_{(1018827-1019786)}$	T			
	hMAGE-B17	NT_011766.3$_{(835710-835102)}$	T			
MAGE-C	*hMAGE-C1*	Ref. 7	T			
	hMAGE-C2	Ref. 8	T			
	hMAGE-C3		T			
	hMAGE-C4	AL023279.1$_{(94144-93800)}$	T			
	hMAGE-C5	NT_025337.2$_{(413026-413457)}$	P[d]			
	hMAGE-C6	NT_025337.2$_{(387877-388503)}$	P			
	hMAGE-C7	NT_025337.2$_{(366380-366796)}$	P			
MAGE-D	*hMAGE-D1*	Ref. 20	N	*mMage-d1*	AF319975	N
	hMAGE-D2[f]	Ref. 19	N	*mMage-d2*	AF319976	N
	hMAGE-D3	AB029037	N	*mMage-d3*	AF319977	N
	hMAGE-D4	AF320908	N			
MAGE-E	*hMAGE-E1*	AL096827$_{(103046-100173)}$	N	*mMage-e1*	AF319978	N
	hMAGE-E2	NT_025273.3$_{(1969507-1971078)}$	N	*mMage-e2*	AF319983	N
	hMAGE-E3	NT_011954.3$_{(589500-591000)}$	N			
MAGE-F	*hMAGE-F1*	AC016975.3$_{(167853-166930)}$	N			
MAGE-G	*hMAGE-G1*	AC061965.3$_{(145505-144089)}$	N	*mMage-g1*	AF319979	N
				mMage-g2	AF319980	N
MAGE-H	*hMAGE-H1*	AF320912	N	*mMage-h1*	AF319981	N
MAGE-I	*hMAGE-I1*	NT_011638.3$_{(1237956-1238326)}$	P[d]			
	hMAGE-I2	NT_011638.3$_{(46231-46830)}$	P[d]			
MAGE-J	*hMAGE-J1*	NT_011577$_{(101579-101911)}$	P[d]			
MAGE-K	*hMAGE-K1*	Z81311$_{(15990-16586)}$	P	*mMage-k1*	AF319984	T
MAGE-L2	*hMAGE-L2*	Ref. 30	N	*mMage-12*	Ref.30	N
NECDIN	*hNECDIN*	Ref. 29	N	*mNecdin*	Ref.28	N

[a] Newly identified genes are referred to by their Genbank accession nos. For those sequences that were identified in large genomic sequences, the coordinates corresponding to the *MAGE* sequences are indicated.

[b] The pattern of expression is represented by: P, probable pseudogene; T, specifically expressed in germ cells and/or in tumors; N, expressed in normal adult tissues; ND, not determined.

[c] The members of these subfamilies carrying identical names in human and mouse are not necessarily orthologues.

[d] The expression of these genes was not tested by RT-PCR, but they are probably not functional because multiple stop codons interrupt the sequence that potentially encodes a MAGE-like protein.

[e] These genes are expressed at low levels in testis but are probably not functional because multiple stop codons interrupt the sequence that potentially encodes a MAGE-like protein.

[f] Sequences referred to as "MAGE-D" in Ref. 19 were renamed MAGE-D2.

comprised in a single exon, as previously observed for the *MAGE-A*, *-B,* and *-C* genes.

The homology between the three mMage-d proteins was found to be restricted to a region encompassing the MAGE domain and the 40 amino acids immediately downstream of it (Fig. 6*B*). The NH_2-terminal regions of the proteins appeared to be very different, and the COOH-terminal domain encoded by exon 12 was much longer in mMage-d3 than in the other mMage-d proteins. Data bank searches did not reveal any significant homology between known proteins and the NH_2-terminal regions of mMage-d1, d2, or d3. By contrast, the long COOH-terminal domain of mMage-d3 was found to contain the entire sequence of trophinin, a previously described protein consisting essentially of decapeptide repeats (25, 26). This is in agreement with the recent identification of a human brain cDNA clone (27) that similarly contained *MAGE-D3* and *trophinin* coding sequences in frame. Interestingly, when the trophinin repeats were omitted, the COOH-terminal parts of the mouse and human MAGE-D3 proteins could be aligned to the corresponding regions of MAGE-D1 and -D2 (Fig. 7).

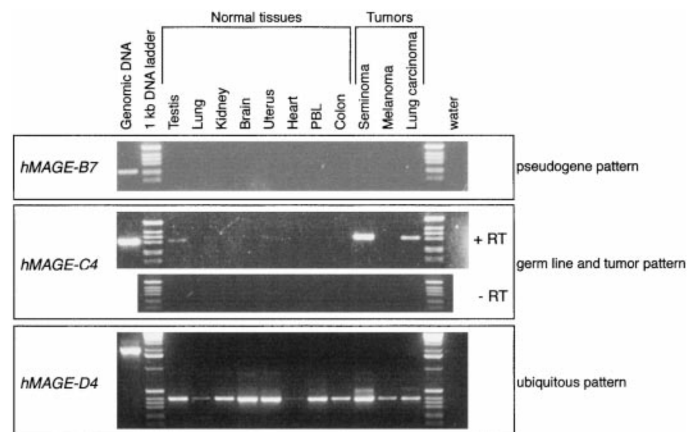


Fig. 2. RT-PCR analysis of the expression of *hMAGE-B7*, *-C4,* and *-D4*, taken as representatives of the newly identified *MAGE* genes. Absence of PCR product in the absence of reverse-transcription is shown for *hMAGE-C4*.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010005544.pdf by guest on 26 February 2024

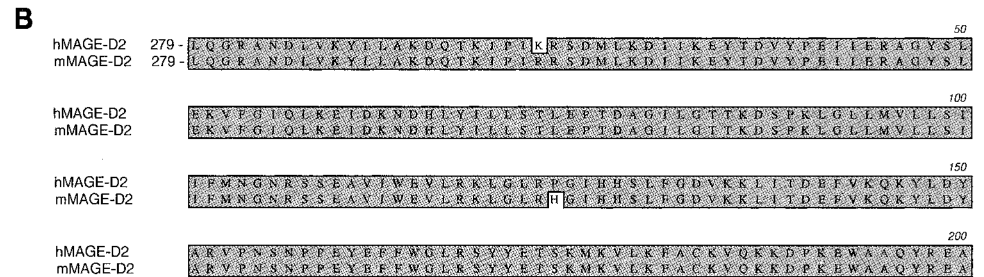


Fig. 3.: Sequence comparison between the MAGE conserved domains of several human MAGE proteins (*A*) and of human and mouse MAGE-D2 proteins (*B*). The location of the first amino acid of the conserved domain in the protein is indicated, except for those proteins whose complete sequence is not known.

Expression Pattern of the *mMage-d* Genes. To determine the relative expression of the three murine *Mage-d* genes in different tissues and at different developmental stages, we performed a semiquantitative mRNA dot blot analysis using cDNA probes specific for each gene. Transcripts corresponding to *mMage-d1* and *-d2* could be detected in all of the tissues analyzed with this method, although at different levels. For instance, *mMage-d1* appeared to be predominantly expressed in the adult brain (Fig. 8*A*),

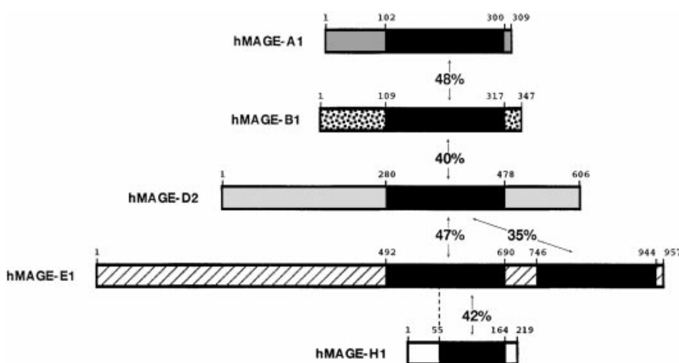


Fig. 4.: Schematic representations of the hMAGE-A1, -B1, -D2, E1, and -H1 proteins. ■, MAGE conserved domains. The percentage of identical amino acids between the MAGE conserved domains are indicated. Different fillings of the NH_2- and COOH-terminal regions of the proteins represent sequences that share no homology.

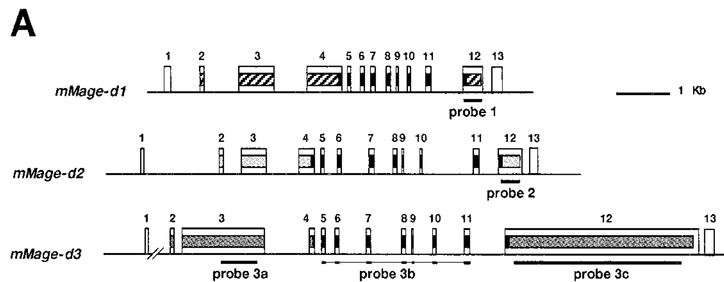
whereas the expression of *mMage-d2* was found to increase steadily during embryonic development, reaching a maximum just before birth (Fig. 8*A*). By contrast, hybridization with a *mMage-d3*-specific probe gave significant signals essentially in the adult brain and in embryos (Fig. 8*A*), although all tissues tested scored positive by RT-PCR (not shown). Importantly, a *trophinin* probe gave similar results (Fig. 8*A*), and a Northern blot analysis of brain tissue detected a single band of about 7 kb that contained both *mMage-d3* and *trophinin* sequences (Fig. 8*B*). We therefore conclude that the *mMage-d3* and *trophinin* exons are expressed as a single transcriptional unit in the mouse brain. In addition, transcripts containing *MAGE-D3* and *trophinin* sequences also could be detected in various tissues by RT-PCR (not shown).

Trophinin was originally described in humans as a homophilic adhesion molecule specifically expressed in trophoblastic and endometrial cells and potentially involved in embryo implantation (25). To determine whether the trophinin transcripts present in human endometrium also contained *hMAGE-D3* exons, we performed a RACE experiment on mRNA prepared from early secretory phase human endometrium. Most of the clones that were obtained (136 of 153) carried both exon 11 and 12 sequences, implying that the *hMAGE-D3* and *trophinin* exons were predominantly expressed as a single transcriptional unit in human endometrium. However, preliminary data also suggested the existence of two discrete transcription start sites, one at the splicing site of

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010105544.pdf by guest on 26 February 2024

mNecdin
hNECDIN
mMage-l2
hMAGE-L2
mMage-d3
hMAGE-D3
mMage-d2
hMAGE-D2
mMage-d1
hMAGE-D1
hMAGE-D4
mMage-g2
mMage-g1
hMAGE-G1
hMAGE-F1
mMage-e1
hMAGE-E1
hMAGE-E2
mMage-k1
hMAGE-K1
mMage-a8
mMage-a3
mMage-a2
mMage-a1
mMage-a5
mMage-a6
mMage-a4
hMAGE-A10
hMAGE-A11
hMAGE-A12
hMAGE-A2
hMAGE-A6
hMAGE-A3
hMAGE-A8
hMAGE-A4
hMAGE-A1
hMAGE-A9
hMAGE-A13
hMAGE-C2
hMAGE-C1
hMAGE-B10
hMAGE-B16
mMage-b8
hMAGE-B3
hMAGE-B5
mMage-b10
mMage-b4
mMage-b7
mMage-b5
mMage-b6
hMAGE-B4
hMAGE-B2
hMAGE-B1
mMage-b3
mMage-b1/2
hMAGE-B6

Fig. 5.: Dendrogram representation of the multiple alignment between MAGE conserved domains of human and mouse MAGE proteins. Human MAGE-C3, -C4, and -H1 and mouse Mage-h1 and -e2 are not represented, because the complete sequence of their MAGE conserved domain are not available.



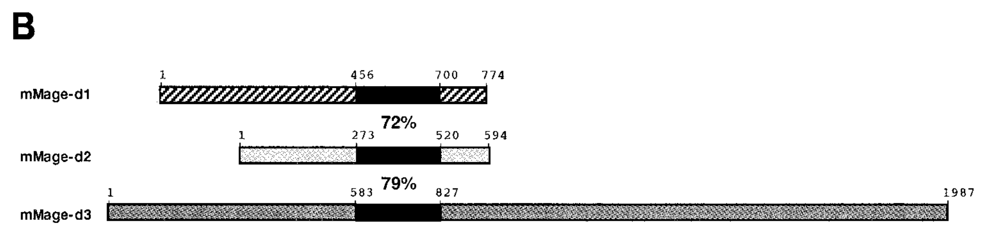


Fig. 6.: Structure of the murine *Mage-d* genes and proteins. *A*, genomic organization of the *mMage-d1*, *-d2,* and *-d3* genes. Coding regions are represented by *filled boxes*, the segments encoding the MAGE conserved domains are in *black*. The different probes used are indicated by *solid bars*. *B*, comparison of the mMage-d1, -d2, and -d3 proteins. The region of homology shared by the three Mage-d proteins is shown in *black*, and the percentage of identical amino acids between these regions are indicated.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010554.pdf by guest on 26 February 2024

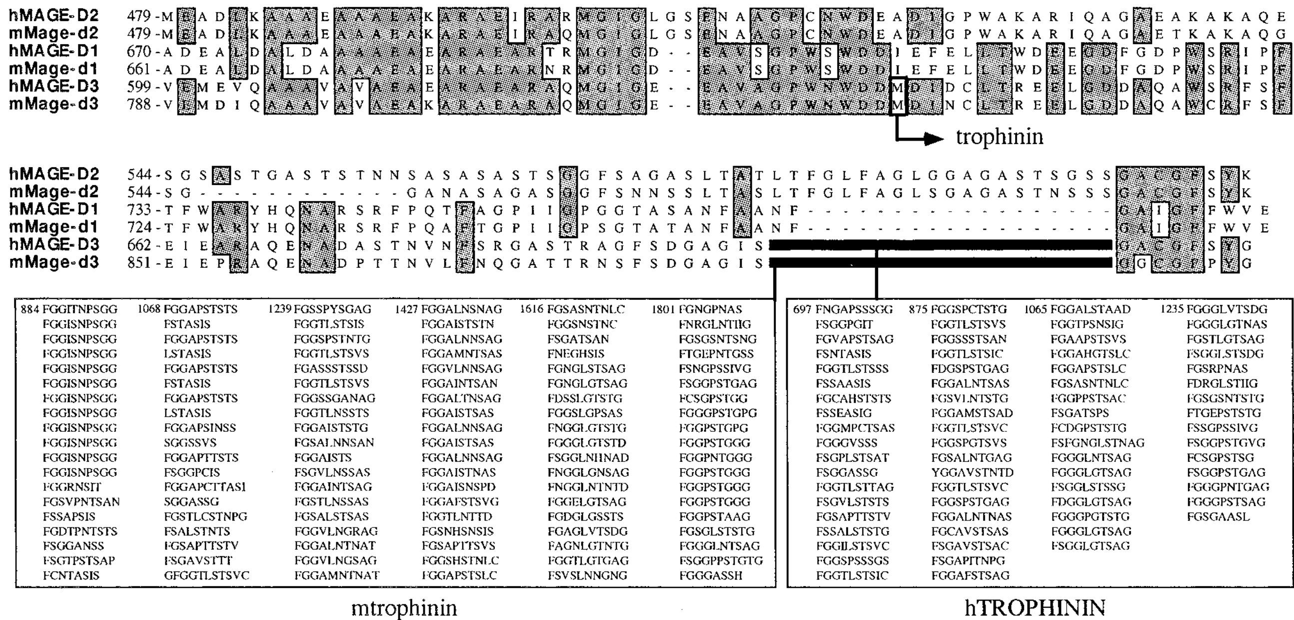


Fig. 7.: Alignment of the COOH-terminal domains of the human and mouse MAGE-D proteins. *Arrow,* the beginning of the trophinin sequence present in the human and mouse MAGE-D3 COOH-terminal domains. The decapeptide repeats that constitute most of trophinin are represented by *horizontal boxes* in the alignment and are detailed in the *bottom* part of the figure.

exon 12 and one in intron 11 (Fig. 9). The possibility therefore remains that the *trophinin* gene can be transcribed independently of *MAGE-D3* in endometrial cells.

DISCUSSION

The human *MAGE*, *GAGE,* and *BAGE* genes were originally described as completely silent in normal adult tissues, with the exception of testis and, for some of them, placenta (9). By contrast, these genes were expressed at a high frequency in various kind of tumors. Therefore, the corresponding proteins represent attractive targets for cancer immunotherapy, because it can be assumed that immunized patients should not be tolerant to such tumor-specific antigens. However, some members of the *MAGE* family are expressed in normal tissues. This is the case for *Necdin*, the first member of the family to be identified (28, 29). More recently, we (19) and others (20, 30), have identified new members of the *MAGE* family that are ubiquitously expressed. In this paper, we report the existence of eight additional members of the human *MAGE* family whose expression is not restricted to tumors and male germ cells. None of the antigenic peptides encoded by the human *MAGE-A* and *-B* genes (9, 31) could be identified in any of the MAGE proteins that show expression in normal somatic tissues. We therefore conclude that immunized patients should not be tolerant to any of these antigens, and that the immune response triggered by the immunization should not have any autoimmune consequences on healthy tissues.

The *MAGE-D* genes contain 13 exons, 11 of which encode the protein. By contrast, all of the other *MAGE* genes share a less complex genomic structure, almost invariably characterized by a large terminal exon carrying the complete coding sequence. This suggests that one or

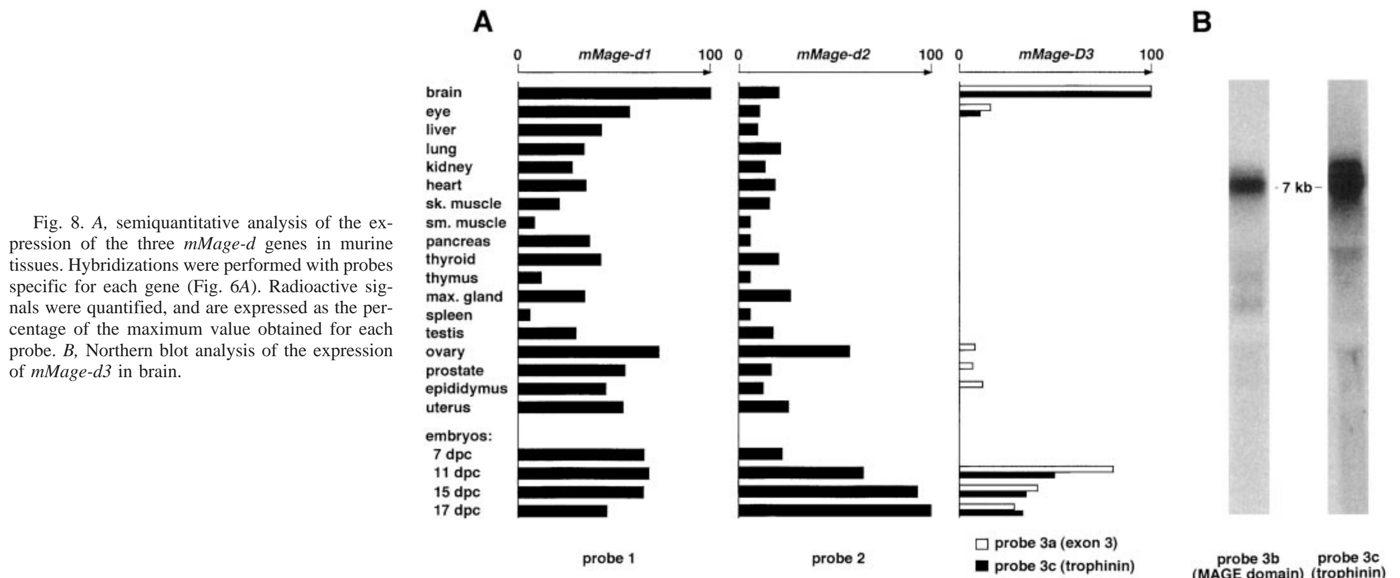


Fig. 8. *A,* semiquantitative analysis of the expression of the three *mMage-d* genes in murine tissues. Hybridizations were performed with probes specific for each gene (Fig. 6*A*). Radioactive signals were quantified, and are expressed as the percentage of the maximum value obtained for each probe. *B,* Northern blot analysis of the expression of *mMage-d3* in brain.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010005544.pdf by guest on 26 February 2024

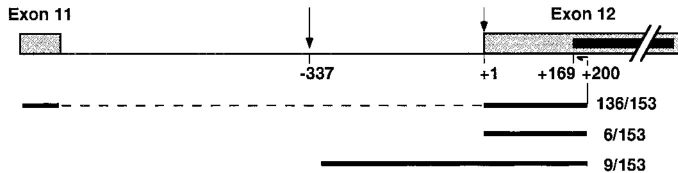


Fig. 9. RACE analysis of the transcripts containing exon 12 of *hMAGE-D3* in human endometrium. Coordinates in intron 11 and exon 12 are relative to the first nucleotide of exon 12. The *black box* in exon 12 corresponds to the *trophinin* coding sequence described in (25). *Horizontal arrow,* the 3′ primer used for the nested RACE PCR. The three major types of transcripts detected are represented by *solid bars*, and the numbers of independent RACE clones obtained for each type are shown. The locations of two putative transcription start sites are indicated by *vertical arrows*.

several *MAGE-D* ancestor genes have generated the first member of other *MAGE* subfamilies by retroposition, a process that frequently occurs in mammalian genomes (32, 33). In addition, gene duplication has obviously contributed to the emergence of the multigenic *MAGE* subfamilies that we observe today. Some of these duplications have occurred recently, indicating an unusually rapid evolution. This is the case for the murine *Mage-a* genes, which are much closer to each other (up to 99% nucleotide identities in their coding sequences) than they are to their human orthologues. By contrast, the duplication events that produced *MAGE-D1*, *-D2,* and *-D3* must be much older. Indeed, the NH_2- and COOH-terminal regions that flank the MAGE conserved domain are completely different for each MAGE-D paralogue but are highly conserved between human and mouse orthologues. This clearly indicates that the *MAGE-D* genes have evolved independently for a long time before the phylogenic separation of the two species. Interestingly, repeat insertion appears to have played a major role during the evolution of the family. For instance, the long COOH-terminal domain of MAGE-D3 was most probably formed by serial duplications of decapeptide repeats, and the NH_2-terminal domains of MAGE-C1 and MAGE-D1, which are also highly repetitive (7, 20), must also have undergone sequential duplication events.

The fact that three very different MAGE-D proteins were conserved during the evolution of mammals strongly suggests that these proteins exert important but distinct functions in this phylum. Importantly, searches in databases also identified MAGE-like genes in nonmammalian species. We identified a zebrafish *MAGE* gene with a structure similar to that of the mammalian *MAGE-D* genes (11 exons, 9 of which encode the MAGE conserved domain; data not shown). In addition, a single *MAGE*-like gene was identified in the genome of the fly *Drosophila melanogaster* (FlyBase accession no. FBgn0037481; Ref. 34). Surprisingly however, we were unable to identify MAGE homologous sequences in the genome of the nematode *Caenorhabditis elegans* or in the yeasts *Saccharomyces cerevisiae* and *Saccharomyces pombe* although MAGE sequences were identified in several vegetal species, including *Arabidopsis thaliana* (GenBank accession no. AF234632; Ref. 34).

Recently, a two-hybrid analysis identified rat MAGE-D1 as a binding partner for the p75 neurotrophin receptor, raising the possibility that it could be a component of its intracellular signaling pathway (21). Although a more refined mapping is clearly required, the available data seems to point to the MAGE conserved domain of MAGE-D1 as the region involved in p75 binding. Presumably, MAGE-D1 signaling to downstream targets could be mediated by a different region of the protein that would be specific to MAGE-D1. Therefore, if other MAGE-D proteins also interact with p75 or related receptors through their MAGE conserved domain, one can assume that these interactions could result in different intracellular responses. Alternatively, receptor binding could be mediated by a MAGE-D1-specific sequence and downstream signaling by the MAGE conserved domain. These issues could be clarified by performing two-hybrid experiments using each MAGE-D protein as a bait.

As suggested above, the different MAGE-D proteins are unlikely to exert redundant functions because their NH_2- and COOH-terminal domains are extremely variable. In this respect, the COOH-terminal part of MAGE-D3 is remarkable because it is identical to trophinin, a previously described protein thought to be involved in embryo implantation (25). In the mouse brain, a single *mMage-d3* transcript is observed that also carries the *trophinin* sequence in frame. Translation of this transcript should generate a large protein composed of a transmembrane trophinin domain and an intracellular MAGE domain that could potentially be involved in an intracellular signaling pathway similar to that proposed for MAGE-D1. However, the size of the protein detected by anti-trophinin antibodies in brain corresponds to the size predicted for trophinin alone (26). This suggests that the physiological form of trophinin is devoid of most, if not all, MAGE-D3 sequences. Whether this discrepancy is attributable to a preferential initiation of translation at the trophinin ATG in exon 12 of *MAGE-D3* or to a rapid processing of a large precursor protein remains to be investigated.

Most of the *MAGE* genes that exist today appear to be retrogenes derived from one or several *MAGE-D* ancestral genes. Retroposition usually results in the acquisition of a defective cDNA copy of the founder gene that degenerates into a pseudogene. However, inactivity is not always a retrogene's fate, and it has been proposed that most intronless genes present in today's eukaryotic genomes are functional retroposons that have lost their characteristic 3′ poly(A) stretches and flanking direct repeats because of their old age (32). The *MAGE* retrogenes obviously belong to this category. *Necdin*, which is highly conserved between man and mouse, is a candidate gene for the Prader-Willi syndrome. Its recent inactivation in the mouse germ line results in perinatal lethality, at least in some genetic backgrounds (35). *MAGE-G1* is also strikingly conserved between man and mouse (91% amino acid identities in the MAGE conserved domain), suggesting that it also exerts important functions. By contrast, many other *MAGE* retrogenes are poorly conserved during evolution. However, despite their old age, many of them still contain an intact open reading frame and are transcribed in male germ cells. The possibility therefore remains that these genes encode proteins whose functions could be related to those of the *MAGE-D* ancestral gene(s). More interestingly, some of these retrogenes could have been recruited during evolution to acquire novel activities, a process referred to as exaptation (36).

ACKNOWLEDGMENTS

We thank Philippe Auquier and Maria Panagiotakopoulos for excellent technical assistance and Dr. Etienne Marbaix for the human endometrium surgical sample.

REFERENCES

1. van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A., and Boon, T. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science (Wash. DC), *254:* 1643–1647, 1991.
2. De Plaen, E., Arden, K., Traversari, C., Gaforio, J. J., Szikora, J-P., De Smet, C., Brasseur, F., van der Bruggen, P., Lethé, B., Lurquin, C., Brasseur, R., Chomez, P., De Backer, O., Cavenee, W., and Boon, T. Structure, chromosomal localization and expression of twelve genes of the *MAGE* family. Immunogenetics, *40:* 360–369, 1994.
3. Rogner, U. C., Wilke, K., Steck, E., Korn, B., and Poustka, A. The *melanoma antigen* gene (*MAGE*) family is clustered in the chromosomal band Xq28. Genomics, *29:* 725–731, 1995.
4. Dabovic, B., Zanaria, E., Bardoni, B., Lisa, A., Bordignon, C., Russo, V., Matessi, C., Traversari, C., and Camerino, G. A family of rapidly evolving genes from the sex reversal critical region in Xp21. Mamm. Genome, *6:* 571–580, 1995.
5. Lurquin, C., De Smet, C., Brasseur, F., Muscatelli, F., Martelange, V., De Plaen, E., Brasseur, R., Monaco, A. P., and Boon, T. Two members of the human *MAGEB* gene

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010105544.pdf by guest on 26 February 2024

family located in Xp.21.3 are expressed in tumors of various histological origins. Genomics, *46:* 397–408, 1997.

6. Muscatelli, F., Walker, A. P., De Plaen, E., Stafford, A. N., and Monaco, A. P. Isolation and characterization of a new *MAGE* gene family in the Xp21.3 region. Proc. Natl. Acad. Sci. USA, *92:* 4987–4991, 1995.
7. Lucas, S., De Smet, C., Arden, K. C., Viars, C. S., Lethé, B., Lurquin, C., and Boon, T. Identification of a new *MAGE* gene with tumor-specific expression by representational difference analysis. Cancer Res., *58:* 743–752, 1998.
8. Lucas, S., De Plaen, E., and Boon, T. *MAGE-B5*, *MAGE-B6*, *MAGE-C2* and *MAGE-C3*: four new members of the *MAGE* family with tumor-specific expression, Int. J. Cancer, *87:* 55–60, 2000.
9. Boon, T., and Van den Eynde, B. J. Shared tumor-specific antigens. *In:* S. A. Rosenberg (ed.), Principles and Practice of the Biologic Treatment of Cancer, Ed. 3, 493–504, 1999.
10. Nestle, F. O., Alijagic, S., Gilliet, M., Sun, Y., Grabbe, S., Dummer, R., Burg, G., and Schadendorf, D. Vaccination of melanoma patients with peptide- or tumor lysate-pulsed dendritic cells. Nat. Med., *4:* 328–332, 1998.
11. Marchand, M., Weynants, P., Rankin, E., Arienti, F., Belli, F., Parmiani, G., Cascinelli, N., Bourlond, A., Vanwijck, R., Humblet, Y., Canon, J-L., Laurent, C., Naeyaert, J-M., Plagne, R., Deraemaeker, R., Knuth, A., Jäger, E., Brasseur, F., Herman, J., Coulie, P. G., and Boon, T. Tumor regression responses in melanoma patients treated with a peptide encoded by gene *MAGE-3*. Int. J. Cancer, *63:* 883–885, 1995.
12. Marchand, M., van Baren, N., Weynants, P., Brichard, V., Dréno, B., Tessier, M-H., Rankin, E., Parmiani, G., Arienti, F., Humblet, Y., Bourlond, A., Vanwijck, R., Liénard, D., Beauduin, M., Dietrich, P-Y., Russo, V., Kerger, J., Masucci, G., Jäger, E., De Greve, J., Atzpodien, J., Brasseur, F., Coulie, P. G., Van der Bruggen, P., and Boon, T. Tumor regressions observed in patients with metastatic melanoma treated with an antigenic peptide encoded by gene *MAGE-3* and presented by HLA-A1. Int. J. Cancer, *80:* 219–230, 1999.
13. Thurner, B., Haendle, I., Roder, C., Dieckmann, D., Keikavoussi, P., Jonuleit, H., Bender, A., Maczek, C., Schreiner, D., von den Driesch, P., Brocker, E. B., Steinman, R. M., Enk, A., Kampgen, E., and Schuler, G. Vaccination with MAGE-3A1 peptide-pulsed mature, monocyte-derived dendritic cells expands specific cytotoxic T cells and induces regression of some metastases in advanced stage IV melanoma. J. Exp. Med., *190:* 1669–1678, 1999.
14. De Plaen, E., De Backer, O., Arnaud, D., Bonjean, B., Chomez, P., Martelange, V., Avner, P., Baldacci, P., Babinet, C., Hwang, S-Y., Knowles, B., and Boon, T. A new family of mouse genes homologous to the human *MAGE* genes. Genomics, *55:* 176–184, 1999.
15. De Backer, O., Verheyden, A-M., Martin, B., Godelaine, D., De Plaen, E., Brasseur, R., Avner, P., and Boon, T. Structure, chromosomal location, and expression pattern of three mouse genes homologous to the human *MAGE* genes. Genomics, *28:* 74–83, 1995.
16. Osterlund, C., Tohonen, V., and Forslund, K. O. *Mage-b4*, a novel *melanoma antigen* (*MAGE*) gene specifically expressed during germ cell differentiation. Cancer Res., *60:* 1054–1061, 2000.
17. Chomez, P., Williams, R., De Backer, O., Boon, T., and Vennström, B. The *SMAGE* gene family is expressed in post-meiotic spermatids during mouse germ cell differentiation. Immunogenetics, *43:* 97–100, 1995.
18. Clotman, F., De Backer, O., De Plaen, E., Boon, T., and Picard, J. Cell- and stage-specific expression of *Mage* genes during mouse spermatogenesis. Mamm. Genome, *11:* 696–699, 2000.
19. Lucas, S., Brasseur, F., and Boon, T. A new *MAGE* gene with ubiquitous expression does not code for known MAGE antigens recognized by T cells. Cancer Res., *59:* 4100–4103, 1999.
20. Pold, M., Zhou, J., Chen, G. L., Hall, J. M., Vescio, R. A., and Berenson, J. R. Identification of a new, unorthodox member of the *MAGE* gene family. Genomics, *59:* 161–167, 1999.
21. Salehi, A. H., Roux, P. P., Kubu, C. J., Zeindlzer, C., Bhakar, A., Tannis, L., Verdi, J. M., and Barker, P. A. NRAGE, a novel MAGE protein, interacts with the p75 neurotrophin receptor and facilitates nerve growth factor-dependent apoptosis. Neuron, *27:* 279–288, 2000.
22. Masuda, Y., Sasaki, A., Shibuya, H., Ueno, N., Ikeda, K., and Watanabe, K. Dlxin-1, a novel protein that binds Dlx5 and regulates its transcriptional function. J. Biol. Chem., *276:* 5331–5338, 2001.
23. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., *25:* 3389–3402, 1997.
24. Thompson, J. D., Higgins, D. G., and Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., *22:* 4673–4680, 1994.
25. Fukuda, M. N., Sato, T., Nakayama, J., Klier, G., Mikami, M., Aoki, D., and Nozawa, S. Trophinin and tastin, a novel cell adhesion molecule complex with potential involvement in embryo implantation. Genes Dev., *9:* 1199–1210, 1995.
26. Suzuki, N., Nadano, D., Paria, B. C., Kupriyanov, S., Sugihara, K., and Fukuda, M. N. Trophinin expression in the mouse uterus coincides with implantation and is hormonally regulated but not induced by implanting blastocysts. Endocrinology, *141:* 4247–4254, 2000.
27. Kikuno, R., Nagase, T., Ishikawa, K., Hirosawa, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and Ohara, O. Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. DNA Res., *6:* 197–205, 1999.
28. Maruyama, K., Usami, M., Aizawa, T., and Yoshikawa, K. A novel brain-specific mRNA encoding nuclear protein (necdin) expressed in neurally differentiated embryonal carcinoma cells. Biochem. Biophys. Res. Commun., *178:* 291–296, 1991.
29. Jay, P., Rougeulle, C., Massacrier, A., Moncla, A., Mattei, M-G., Malzac, P., Roëckel, N., Taviaux, S., Bergé Lefranc, J-L., Cau, P., Berta, P., Lalande, M., and Muscatelli, F. The human *necdin* gene, *NDN*, is maternally imprinted and located in the Prader-Willi syndrome chromosomal region. Nat. Genet., *17:* 357–361, 1997.
30. Boccaccio, I., Glatt-Deeley., H. Watrin, F., Roeckel, N., Laland, M., and Muscatelli, F. The human *MAGEL2* gene and its mouse homologue are paternally expressed and mapped to the Prader-Willi region. Hum. Mol. Genet., *8:* 2497–2505, 1999.
31. Fleischhauer, K., Gattinoni, L., Dalerba, P., Lauvau, G., Zanaria, E., Dabovic, B., van Endert, P. M., Bordignon, C., and Traversari, C. The *DAM* gene family encodes a new group of tumor-specific antigens recognized by human leukocyte antigen A2-restricted cytotoxic T lymphocytes. Cancer Res., *58:* 2969–2972, 1998.
32. Brosius, J. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene, *238:* 115–134, 1999.
33. Weiner, A. M., Deininger, P. L., and Efstratiadis, A. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem., *55:* 631–661, 1986.
34. Pold, M., Pold, A., Ma, H. J., Sjak-Shieb, N. N., Vescio, R. A., and Berensonb, J. R. Cloning of the first invertebrate MAGE paralogue: an epitope that activates T-cells in humans is highly conserved in evolution. Dev. Comp. Immunol., *24:* 719–731, 2000.
35. Gerard, M., Hernandez, L., Wevrick, R., and Stewart, C. L. Disruption of the mouse *necdin* gene results in early post-natal lethality. Nat. Genet., *23:* 199–202, 1999.
36. Brosius, J. Retroposons—seeds of evolution. Science (Wash. DC), *251:* 753, 1991.
37. Miller, A. P., and Willard, H. F. Chromosomal basis of X chromosome inactivation: identification of a multigene domain in Xp11.21-p11.22 that escapes X inactivation. Proc. Natl. Acad. Sci. USA, *95:* 8709–8714, 1998.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/61/14/5544/2862930/ch14010105544.pdf by guest on 26 February 2024