European Journal of Human Genetics (2002) 10, 833–840

www.nature.com/ejhg

ARTICLE

# New *BAGE (B melanoma antigen)* genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile

Myriam Ruault[1], Pierre van der Bruggen[2], Marie-Elisabeth Brun[1], Shelagh Boyle[3], Gérard Roizès[1] and Albertina De Sario*[,1]

[1]*Institut de Génétique Humaine, CNRS UPR 1142, 141, rue de la Cardonille, 34396 Montpellier, France;* [2]*Ludwig Institute for Cancer Research, Brussels Branch, 74 av Hippocrate, B1200 Brussels, Belgium;* [3]*MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, UK*

**A first *BAGE* (*B melanoma antigen*) gene, *BAGE1*, was identified because it encodes a human tumour antigen recognised by a cytolytic T lymphocyte. Here, we characterised five new *BAGE* genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 and nine *BAGE* gene fragments mapping to the juxtacentromeric regions of chromosomes 9, 13, 18, and 21. Genes and gene fragments share extensive regions of 90–99% nucleotide identity. We analysed the expression of *BAGE* genes on 215 tumours of various histological types and on nine normal tissues. Similar to *BAGE1*, the new *BAGE* genes are expressed in melanomas, bladder and lung carcinomas and in a few tumours of other histological types. All the normal tissues were negative, with the exception of testis. Our results show that human juxtacentromeric regions harbour genes, which are transcribed and translated, in addition to gene fragments that are generally not expressed. We suggest that the pattern of expression restricted to cancer/testis is a feature of the few genes mapping to juxtacentromeric regions.**
*European Journal of Human Genetics* (2002) **10,** 833–840. doi:10.1038/sj.ejhg.5200891

**Keywords:** BAGE; chromosome 21; chromosome 13; juxtacentromeric regions; gene family; gene fragment; tumour

By screening a human melanoma cDNA library with an anti-tumour cytolytic T lymphocyte (CTL), we isolated a 1004 bp-long transcript coding for the tumour antigen recognised by the CTL.[1] It was named *BAGE,* for human *B melanoma antigen* (gb U19180) and, hereafter, will be referred to as *BAGE1e*. *BAGE* encodes a putative protein of 43 amino acids and the antigen recognised by the CTL consists of *BAGE1*-encoded peptide AARAVFLAL bound to HLA-Cw16 molecules.[1] *BAGE1a* was found to be expressed in tumours of different histological types. It is silent in normal and foetal tissues except for testis.

Several other unrelated families of genes with a pattern of expression similar to that of *BAGE1* have been identified, notably the *MAGE*, *GAGE* and *LAGE/NY-ESO*-1 families.[2–7] These genes are located on the X chromosome and their function is still unknown. The *MAGE* family contains 18 related genes divided into three clusters (*MAGE-A*, *B*, and *C*) located on the X chromosome.[3,8,9]

A subset of these male germline-specific genes are activated in a wide variety of tumours, in which they code for tumour-specific antigens recognised by autologous T lymphocytes.[10] Male germline cells do not express molecules of the major histocompatibility complex (MHC), which are surface molecules required to present antigenic peptides to T lymphocytes.[11] The antigens recognised by T lymphocytes that are encoded by *BAGE, MAGE, GAGE* and *LAGE/NY-ESO-1* families are, therefore, strictly tumour-specific and may prove useful for cancer immunotherapy. Clinical trials involving defined MAGE and LAGE tumour antigens are proceeding.[12,13] Long-lasting

*Correspondence: A De Sario; Institut de Génétique Humaine, CNRS UPR 1142, 141 rue de la Cardonille, 34396 Montpellier, France.
Tel +334 99 61 99 77; Fax: +334 99 61 99 01; E-mail: desario@infobiogen.fr
Received 2 May 2002; revised 30 July 2002; accepted 31 July 2002

tumour regressions have been observed in a minority of patients.

During annotation of genomic sequences of human chromosome 21,[14] we identified three DNA stretches having 92–99% nucleotide identity with *BAGE1a* mRNA. The *BAGE*-related sequences are located less than 1 Mb away from the centromere: two in 21p and one in 21q. We searched for new *BAGE* transcripts and show here that *BAGE* is a gene family composed of expressed genes that map to the juxtacentromeric regions of chromosomes 13 and 21, and of unexpressed gene fragments that are scattered in the juxtacentromeric regions of several chromosomes. We have analysed the expression of the new *BAGE* genes in tumours and normal tissues.

## Materials and methods

Primer sequences and detailed protocols are available online (http://www.igh.cnrs.fr/equip/centromere/Bage.html).

### Screening of a melanoma library

The MZ2.MEL melanoma cell line cDNA library was screened using probes derived from the 5′ and 3′ regions of *BAGE1a*. Hybridisations were done at stringent conditions according to standard procedures.[15]

### 5′ and 3′ RACE and cDNA cloning

5′ and 3′ RACE experiments were done using the Marathon Ready human testis cDNA (#7414-1, Clontech) and the Expand PCR kit (Boehringer) according to the manufacturer protocols. 5′ and 3′ RACE products were purified using the Wizard kit (Promega) and cloned using the pGem-T PCR cloning kit (Promega). Afterwards, specific primers were designed to conserved regions of the 5′ and 3′ end sequences and were used to reamplify three cDNA libraries: two testis and melanoma (MZ2-Mel43) cDNA libraries[1] and a purchased testis cDNA library (Clontech #7414-1). The obtained 2 kb-products were gel purified and cloned using the pGem-T PCR cloning kit (Promega).

### *BAGE* genomic structure

The genomic structure of *BAGE* genes was determined through alignment of transcripts with genomic sequences. The genomic organisation of *BAGE1* was obtained through alignment of the six transcripts (*BAGE1a*, *BAGE1b*, *BAGE1c*, *BAGE1d*, *BAGE1*e, and *BAGE1f*) with the genomic sequences AF499647 containing (exon 1 and 2) and AC064811 containing (exons 3 to 7). The genomic organisation of *BAGE2* and *BAGE5* was obtained through alignment of the transcripts with the genomic sequences AL163201 and AL161418, respectively.

### Sequencing

Sequencing was done with an Applied Biosystems 373XL sequencer.

### Computer analysis

Nucleotide sequences and predicted amino acid sequences were analysed by BLAST[16] and by CLUSTALW at http://www2.ebi.ac.uk/clustalw

### Chromosome mapping

**In situ *hybridisation*** A 1.8-kb genomic fragment spanning from exon 1 to exon 2 was labelled by nick translation with Bio-16-dUTP (Boehringer) and hybridised to normal human metaphase chromosomes as described in Fantes *et al*.[17] Hybridisation signals were visualised using a Zeiss Axioplan epifluorescence microscope. Images were captured using Digital Scientific Smartcapture software.

***Somatic hybrid cell lines*** Primer pairs specific for *BAGE* were used to amplify DNA from the monochromosome somatic hybrid cell lines (NIGMS mapping panel 2).[18] PCR products were cloned in pGem-T vector (Promega). To map *BAGE1*, 10 colonies per cloning reaction were sequenced with universal primers SP6 and T7.

### Expression analysis

Expression analysis was done on nine normal tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, and testis; Clontech #1420-1; #7414-1) and on 215 tumour samples. Total RNA was extracted from tumour samples by the guanidine-isothiocyanate procedure as described.[19] Reverse transcription was done on 2 $\mu$g of total RNA with 2 mM oligo$(dT)_{15}$ primer and 200 U of MoMLV reverse transcriptase (GIBCO–BRL). The quality of the RNA preparation was tested by PCR amplification of human *β-actin*. Primers used for the expression analysis were located in different exons. A first set of primers, bage14 and bage19, was used to amplify a PCR product of 490 bp corresponding to *BAGE1* and *BAGE5* transcripts. A second set of primers was used to amplify a 368-bp PCR product corresponding to *BAGE2, BAGE3,* and *BAGE5* transcripts. To identify individual genes, 15 $\mu$l of the PCR product was digested with restriction enzymes. Digestion of the 495 bp PCR product with either *Alu*I or *Bst*NI allowed us to identify *BAGE2* and *BAGE3* cDNAs, respectively. Digestion of the 368 bp PCR product with *Bss*KI allowed us to identify *BAGE1* (presence of the restriction site) and *BAGE5* (absence of the restriction site).

To analyse the expression of *BAGE* gene fragments we amplified a testis cDNA library (Clontech #7417-1) with primers bage4 and bage5. The PCR product was cloned in pGEM-T vector (Promega). Twenty colonies were sequenced and compared with genomic sequences in databases.

## Results

### *BAGE* transcripts and genes

We renamed *BAGE1* the gene corresponding to the transcript (gbU19180) previously isolated from a human melanoma cDNA library.[1] The same library was hybridised

with probes corresponding to the 5′ and 3′ regions of *BAGE1* and six alternatively spliced mRNAs, whose sizes range from 1 to 2.3 kb, were isolated: *BAGE1a* (gbU19180, ref. 1), *BAGE1b* (AF527550), *BAGE1c* (AF527551), *BAGE1d* (AF527552), *BAGE1e* (AF527553), *BAGE1f* (AF527554) (Figure 1A).

Afterwards, 5′ and 3′ RACE amplifications were done on a testis Marathon cDNA library (Clontech), using primers specific for *BAGE1* (Figure 1B). We failed to amplify sequences upstream of the 5′ region of *BAGE1* and concluded that the 5′ untranslated region is complete. Interestingly, 3′ RACE amplifications yielded fragments that differed in sequence and length from the *BAGE1* mRNAs. PCR primers were chosen in the conserved 5′ and 3′ regions and full length cDNAs were amplified from two testis and one melanoma cDNA libraries. Full length cDNAs were

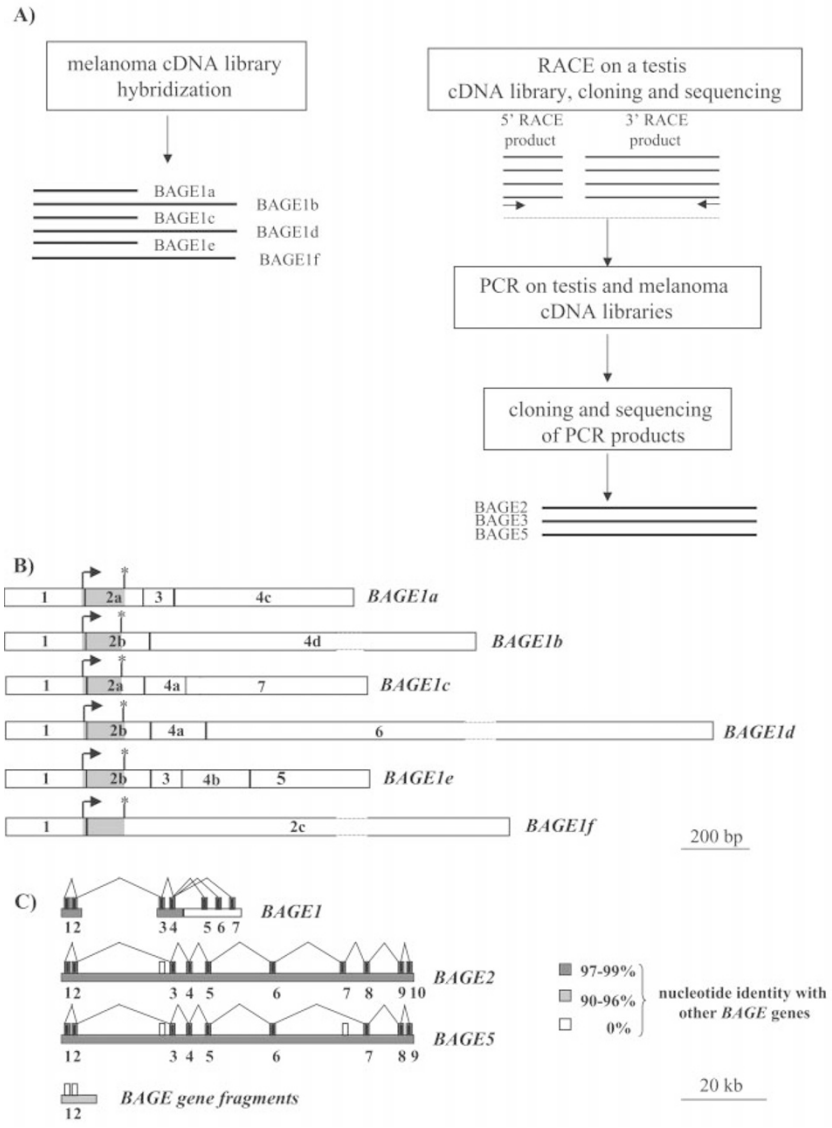


**Figure 1** (**A**) Strategy used to isolate new *BAGE* transcripts: (left side) Hybridisation of a melanoma cDNA library with probes specific for *BAGE1* allowed us to isolate six alternatively spliced *BAGE1* isoforms; (right side) Through RACE experiments on a testis cDNA library we cloned the 5′ and 3′ ends of different *BAGE* transcripts. Then using primers designed in the conserved sequence of different *BAGE* ends, we amplified full length cDNAs on three different (melanoma and testis) cDNA libraries. These transcripts correspond to new *BAGE* genes. (**B**) Exonic structure of the *BAGE1* transcripts. Open rectangles are exons, filled (grey) rectangles correspond to open reading frames (arrows indicate the initiation and stars the stop codons). Exons 2c, 4d, and 6 are not drawn to scale. (**C**) Nucleotide sequence identity among different members of the *BAGE* family. Black rectangles are transcribed exons. White rectangles are predicted exons, i.e. exon 3 of *BAGE1* (an Alu sequence included in the 3′ untranslated region of *BAGE1a* and *BAGE1e* mRNA variants) was also present in *BAGE2* and *BAGE5* genomic sequences, but not in the corresponding transcripts. Likewise, exon 7 of *BAGE2* was also present in the *BAGE5* genomic sequence, but not in the *BAGE5* transcript.

cloned in a plasmid and, for each PCR product, 10 colonies were sequenced. We isolated three 1.9 kb-long transcripts, which were named *BAGE2* (gbAF218570), *BAGE3* (gbAF339514), and *BAGE5* (gbAF339516).

*BAGE2* and *BAGE3* have putative open reading frames (ORFs) of 330 bp encoding predicted proteins of 109 amino acids. *BAGE5*, like *BAGE1*, has an ORF of 132 bp and encodes a predicted protein of 43 amino acids.

We determined the exon/intron structure of the *BAGE* genes (Figure 1B and Table 1). *BAGE1* comprises seven exons. We could not measure the gene length because the corresponding genomic sequence is unordered and uncompleted. *BAGE2* and *BAGE5* span 76 kb and comprise 9 and 10 exons, respectively. All the splicing sites conform the consensus gt/ag sequence. The exon/intron structure of *BAGE3* could not be determined because the corresponding genomic sequence is not available in databases.

### *BAGE* genomic sequences

*BAGE* genomic sequences were searched in databases and we retrieved 14 sequences with a significant nucleotide identity (90–100% $P<10^{-4}$) with the *BAGE* transcripts described above (Table 2). *BAGE* genomic sequences were classified as follows: gene, when a transcript could be assigned to the genomic sequence; predicted gene, when no transcript could be assigned to the genomic sequence, but the predicted ORF was intact; gene fragment, when the gene was truncated and the predicted ORF was disrupted by deleterious mutations (deletions and nucleo-

**Table 1** Exon/intron structure of *BAGE* genes

BAGE1

| *Genbank AC064811* | *Exons*[b] | *Size (bp)* | *5′ Splice site* | *3′ Splice site* | *Introns* | *Size (kb)* |
|---|---|---|---|---|---|---|
| [a] | 1 | ⩾237 | | GGAGCG/**gt**aaga | I | 1.5 |
| [a] | 2a | 170 | ttgt**ag**/GTTTTT | AGGAGG/**gt**aacc | II | nd |
| [a] | 2b | 190 | ttgt**ag**/GTTTTT | CAAGTG/**gt**agga | | |
| [a] | 2c | 1856 | ttgt**ag**/GTTTTT | t | | |
| 133 676–133 586 | 3 | 91 | cagt**ag**AGATGG | TTACAG/**gt**gtga | III | 1.4 |
| 132 178–132 016 | 4a | 163 | tttc**ag**/CGATGT | AAAGAT/**gt**aagt | | |
| 132 215–132 016 | 4b | 200 | ttca**ag**/GTTTTA | AAAGAT/**gt**aagt | | |
| 132 178–131 651 | 4c | 528 | tttc**ag**/CGATGT | t | | |
| 132 178–131 099 | 4d | 1080 | tttc**ag**/CGATGT | t | | |
| 120 733–120 383 | 5[c] | 351 | aaaa**ag**/GATGAG | t | IV | nd |
| 120 345–118 665 | 6[c] | 1680 | ggtc**ag**/TAATAG | t | V | 0.03 |
| 5420–4887 | 7[c] | 533 | tgac**ag**/GGTCTC | t | VI | nd |

BAGE2

| *Genbank HS21C001* | *Exons* | *Size (bp)* | *5′ Splice site* | *3′ Splice site* | *Introns* | *Size (kb)* |
|---|---|---|---|---|---|---|
| –191 680 | 1 | ⩾223 | | GGAGTG/**gt**aaga | I | 1.1 |
| 190 623–190 523 | 2 | 101 | ttgt**ag**/GTTTTT | AGCTCT/**gt**gagt | II | 38.8 |
| 151 755–151 593 | 3 | 163 | tttcag/CGATGT | AAAGAT/**gt**aagt | III | 8.5 |
| 143 054–142 883 | 4 | 172 | ttat**ag**/CGAAGG | CTGCAA/**gt**aagt | IV | 1.8 |
| 141 028–140 914 | 5 | 115 | ttt**ag**/ACAATC | TGCAAA/**gt**aagt | V | 8.1 |
| 132 862–132 161 | 6 | 702 | atac**ag**/TGAATA | AAACAG/**gt**aggg | VI | 9.3 |
| 123 134–123 015 | 7 | 120 | tctt**ag**/GAGGCT | AAAGTG/**gt**acat | VII | 2.8 |
| 120 252–120 136 | 8 | 117 | acat**ag**/GGCCAT | CCTGGG/**gt**gagg | VIII | 5.5 |
| 114 624–114 523 | 9 | 102 | ttat**ag**/GTGTCT | AATCAG/**gt**ttga | IX | 0.2 |
| 114 330– | 10 | ⩾76 | cttc**ag**/GATATG | | | |

BAGE5

| *Genbank AL161418* | *Exons* | *Size (bp)* | *5′ Splice site* | *3′ Splice site* | *Introns* | *Size (kb)* |
|---|---|---|---|---|---|---|
| –76 589 | 1 | ⩾203 | | GGAGCG/**gt**aaga | I | 1.5 |
| 78 261–78 450 | 2 | 190 | ttgt**ag**/GTTTTT | CAAGTG/**gt**agga | II | ⩾38.6 |
| 132 037–132 199 | 3 | 163 | tttc**ag**/CGATGT | AAAGAT/**gt**aagt | III | 8.5 |
| 140 734–140 905 | 4 | 172 | ttat**ag**/CGAAGG | CTGCAA/**gt**aagt | IV | 1.8 |
| 142 760–142 874 | 5 | 115 | ttt**ag**/ACAATC | TGCAAA/**gt**aagt | V | 8 |
| 150 921–151 622 | 6 | 702 | atac**ag**/TGAATA | AAACAG/**gt**aggg | VI | 11.9 |
| 163 514–163 630 | 7 | 117 | acat**ag**/GGCCGT | CCTGGG/**gt**gagg | VII | ⩾5.6 |
| 169 281–169 382 | 8 | 102 | ttat**ag**/GTGTCT | AATCAG/**gt**ttga | VIII | 0.2 |
| 169 575–169 650 | 9 | ⩾76 | cttc**ag**/GATATG | | IX | |

[a]The gene structure was determined using the AF499647 genomic sequence. [b]a, b, c are alternative exons of different length. [c]Order of exons 5–7 is not known because they are alternative terminal exons and the genomic sequence AC064811 is unordered. t=terminal exon and nd=not determined. The ag/gt intron boundaries are in bold type.

tide changes) that introduced stop codons and/or erase the initiation codon.

Overall, we identified three genes (*BAGE2*, *BAGE3*, and *BAGE5*), two predicted genes (*BAGE6* and *BAGE7*), and nine gene fragments (Table 2). *BAGE* gene fragments are 2–8 kb long and correspond to the 5′ region of *BAGE* genes (they span from exon 1 to truncated intron 2).

*BAGE* sequences (genes and gene fragments) share extensive regions of high nucleotide identity: nucleotide identity is higher among genes (97–99%) than between genes and gene fragments (90–96%) (Figure 1C).

**Table 2** List of *BAGE* genomic sequences[a]

| *Genbank* | *Gene type* | *Chromosome* | *Positions* |
|---|---|---|---|
| AC064811 | *BAGE1* gene | 13 | [b] |
| AL163201 | *BAGE2* gene | 21p | 191 902–114 255 |
| AL161418 | *BAGE5* gene | 13 | 76 589–169 650 |
| AL158811 | *BAGE6* predicted gene | 13 | 4 541–147 162 |
| AL049849 | *BAGE7* predicted gene | 21p | 23 819–27 258 |
| AL163203 | gene fragment | 21q | 66 6624–68 522 |
| AC008443 | gene fragment | nd[c] | 18 576–16 465 |
| AC008454 | gene fragment | nd[c] | 32 862–22 256 |
| AL161615 | gene fragment | 13 | 210 981–213 643 |
| AL163539 | gene fragment | 9 | 99 935–101 853 |
| AL356136 | gene fragment | 9 | 6 126–4 209 |
| AL359312 | gene fragment | 9 | 22 315–20 417 |
| AC068255 | gene fragment | 18 | 76 105–78 004 |
| AP001896 | gene fragment | 18q22 | [b] |

[a]Transcripts were tentatively assigned to a specific locus if they had >99.5% nucleotide identity with a given genomic sequence. This threshold was chosen in order to distinguish allelic variations from nucleotide variations corresponding to independent loci, the frequency of which being estimated at 0.1–0.3% in the human genome (ref.14). [b]Unordered genomic sequence. [c]Genomic clone misassigned to chromosome 5. nd=not determined.

### Chromosome mapping

To map *BAGE* genomic sequences, we hybridised a 1.8 kb genomic probe to human metaphase chromosomes. The probe was amplified by PCR from the 5′ region that is common to all the *BAGE* sequences. *BAGE* sequences mapped to the juxtacentromeric regions of chromosomes 9, 13, 14, 15, 18, 21, and 22 (Figure 2). Hybridisation to chromosomes 14 and 15 was observed only in some metaphases suggesting that these *BAGE* sequences were more divergent.

We then amplified exon 1, which is common to all the *BAGE* sequences, on a panel of monochromosome somatic cell lines. Specific amplifications were obtained with chromosomes 9, 13, 15, 18, 21, and 22 (data not shown). No amplification was obtained with chromosome 14: this result can be due to nucleotide divergence between the primers used and the target sequence.

To map individual *BAGE* loci, we analysed the localisation of the genomic sequences retrieved from databases (Table 2). Assignment of genomic sequences to chromo-

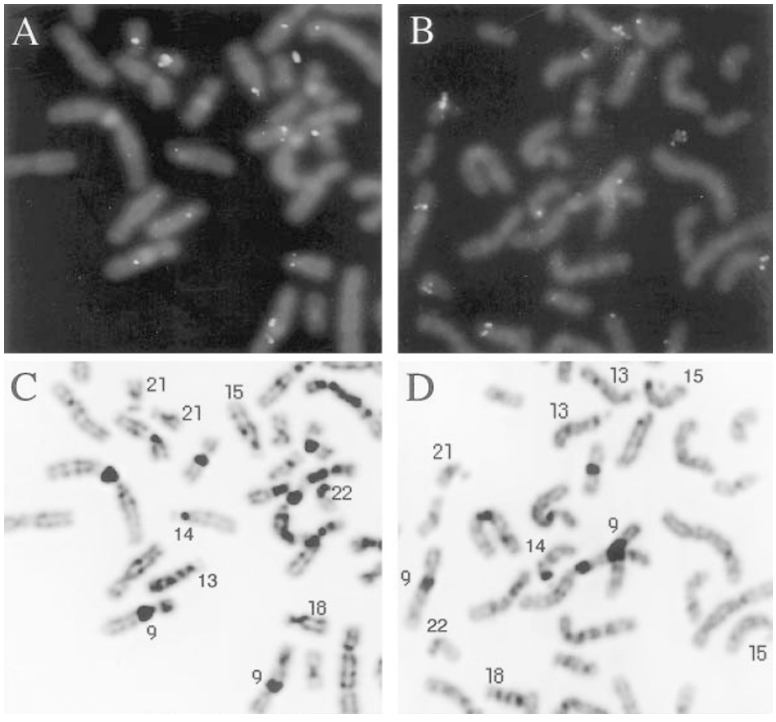


**Figure 2** (**A**) and (**B**) *In situ* hybridization of the 5′ region of *BAGE* to two partial human metaphases. (**C**) and (**D**) DAPI-stained G banding of the partial metaphases.

somes 9, 13, 18, and 21 was consistent with our mapping results; by contrast, assignment to chromosomes 4 and 5 was at variance with both *in situ* hybridisation and somatic hybrid analysis. To ascertain the actual localisation of *BAGE1*, which has 100% nucleotide identity with the genomic clone AC064811 assigned to chromosome 4, we did FISH experiments: the genomic clone AC064811 hybridises to chromosomes 13, 14, 21 and 22 (data not shown). We confirmed this result by PCR on somatic hybrids: a genomic sequence matching (100% nucleotide identity) the *BAGE1* transcripts was amplified from chromosome 13 and from no other chromosome (data not shown).

In conclusion, *BAGE* sequences map to the juxtacentromeric regions of different human chromosomes, each one containing more than one locus: *BAGE* genes map to chromosomes 13 and 21, whereas *BAGE* gene fragments map to chromosomes 9, 13, 18, and 21. *BAGE*-related sequences that are not yet characterised map to chromosomes 14, 15 and 22.

### Predicted BAGE proteins

Multialignment of BAGE predicted proteins shows that amino acid sequence identity ranges from 88 to 98% (Figure 3). Although BAGE2 and BAGE3 differ only in two amino acids, transcripts encoding these proteins are unlikely to be allelic because they have 1.2% nucleotide divergence.

The nonapeptide AARAVFLAL (amino acids 2 to 10 of the BAGE1 predicted protein; boxed in Figure 3) is the sequence of the antigen recognised by a CTL.[1] BAGE predicted proteins have two amino acid changes with respect to the sequence of the BAGE1 antigenic peptide: $(R\rightarrow G)_3$ and $(A\rightarrow V)_4$. Synthetic peptides containing these amino acid variations are not recognised by the CTL that recognises BAGE1 (data not shown).

After database searches, we concluded that BAGE predicted proteins have no significant identity/similarity to any known protein.

### Expression analysis

We analysed the expression of *BAGE* genes in 215 tumour samples of various histological types. To distinguish individual genes, we took advantage of the few nucleotide variations that characterise different transcripts and we designed a strategy based upon PCR followed by restriction enzyme digestion.

A first set of primers (bage14/bage19) amplified a PCR product of 490 bp corresponding to *BAGE1* and *BAGE5* transcripts. Fifteen per cent (15 out of 103) of melanomas scored positive, whereas all the normal tissues were negative, with the exception of testis (Table 3). In a previous study with a different set of primers that amplified the same transcripts, the percentage of positive melanomas was comparable (22%).[1]

A second set of primers (bage20/bage21) was also used to amplify a PCR product of 368 bp corresponding to *BAGE2*, *BAGE3*, and *BAGE5* transcripts. Thirty-four per cent (35 out of 103) of melanomas scored positive, indicating that a significant number of tumours expressed *BAGE2* or *BAGE3* without expressing the other *BAGE* genes, in particular in primary melanomas (Table 3). Here again, all the normal tissues were negative, with the exception of testis.

PCR products were then digested with restriction enzymes *Bss*KI, *Alu*I, *Bst*NI, to identify individual transcripts for *BAGE1*/*BAGE5*, *BAGE2* or *BAGE3*, respectively. In melanomas, *BAGE2* (23%) was more frequently expressed than *BAGE1* (14%), *BAGE3* (14%) and *BAGE5* (9%) (Table 3). Individual tumours generally expressed more than one *BAGE* gene simultaneously (data not shown).

In accordance to a previous analysis,[1] our results confirmed that *BAGE* genes are expressed in melanoma, bladder and lung carcinomas and in a few tumours of other histological types. Leukemias, colorectal carcinomas, and renal carcinomas scored negative. Here, no head and neck tumours expressed *BAGE*, whereas 8% were found positive in the previous study. This could be explained by the small number of samples analysed.

To analyse transcription of *BAGE* truncated genes we amplified a testis and a melanoma cDNA libraries with primers derived from the 5′ region that is common to genes and gene fragments. After having cloned the obtained PCR product, we sequenced 20 colonies and no transcript corresponding to a truncated gene was obtained. We tentatively concluded, therefore, that *BAGE* gene fragments are not expressed.

## Discussion

In this paper, we show that the *BAGE* gene family comprises genes that are transcribed and translated (as indicated by the antigenic properties of BAGE1) and gene fragments that are not expressed. Genes and predicted genes map to the juxtacentromeric regions of human chromosomes 13 and 21, whereas gene fragments map to the juxtacentromeric regions of chromosomes 9, 13, 18, and 21. The list of *BAGE* genes and gene fragments is incom-

```
                  10        20        30        40        50        60        70        80        90       100
BAGE1     M AARAVFLAL SAQLLQARLMKEESPVVSWRLEPEDGTALCFIF----------------------------------------------------------
BAGE2     M AAGVVFLAL SAQLLQARLMKEESPVVSWRLEPEDGTALDVHFVSTLEPLSNAVKRNVPRCIIILVLQEPTAFRISVTSSCFVQNTLTKLLKDRRKMQTVQCATARETS
BAGE3     M AAGVVFLAL SAQLLQARLMKEESPVVSWRLEPEDGTALDVHFVSTLEPLSNAVKRNVPRCIIILVLQEPTPFRISVTSSCFVQNTLTKLLKDRRKMQTVQCATAQETS
BAGE5     M AAGAVFLAL SAQLLQARLMKEESPVVSWRLEPEDGTALCFIF----------------------------------------------------------
BAGE6     M AAGAVFLAL SAQLLQARLMKEESPVVSWRLEPEDGTAL
BAGE7     M AAGAVFLAL SAQLLQARLMKEESPVVSWRLEPEDGTAL
```

**Figure 3** Multialignment of BAGE predicted proteins. BAGE6 and BAGE7 are translations of predicted genes. Non-conserved amino acids are shaded. The frame encloses the sequence of the BAGE1 antigenic nonapeptide.

**Table 3** Expression of *BAGE* genes in normal tissues and tumour samples

| *Histological type* | *Number tested* | *Number of* BAGE-*positive samples* | | | | | |
|---|---|---|---|---|---|---|---|
| | | PCR | | PCR-restriction enzyme | | | |
| | | [a] *BAGE1* *BAGE5* | [b] *BAGE2* *BAGE3* *BAGE5* | BssKI+ *BAGE1* | AluI+ *BAGE2* | BstNI+ *BAGE3* | BssKI− *BAGE5* |
| *Tumour samples* | | | | | | | |
| Melanoma | *103* | 15 | 35 | 14 | 24 | 14 | 9 |
| primary lesion | *26* | 3 | 10 | 2 | 4 | 4 | 2 |
| metastatic lesion | *77* | 12 | 25 | 12 | 20 | 10 | 7 |
| Bladder carcinoma | *20* | 4 | 6 | | | | |
| Lung carcinoma NSCLC | *13* | 0 | 3 | | | | |
| Mammary carcinoma | *16* | 0 | 1 | | | | |
| Sarcoma | *10* | 0 | 1 | | | | |
| Prostatic carcinoma | *6* | 0 | 1 | | | | |
| Neuroblastoma | *1* | 1 | 1 | | | | |
| Head & neck tumour | *11* | 0 | 0 | | | | |
| Leukemia | *10* | 0 | 0 | | | | |
| Renal carcinoma | *10* | 0 | 0 | | | | |
| Colorectal carcinoma | *9* | 0 | 0 | | | | |
| Uterus | *2* | 0 | 0 | | | | |
| Brain tumour | *2* | 0 | 0 | | | | |
| *Normal tissues* | | | | | | | |
| Testis | *1* | 1 | 1 | | | | |
| Placenta | *1* | 0 | 0 | | | | |
| Heart | *1* | 0 | 0 | | | | |
| Brain | *1* | 0 | 0 | | | | |
| Lung | *1* | 0 | 0 | | | | |
| Liver | *1* | 0 | 0 | | | | |
| Skeletal muscle | *1* | 0 | 0 | | | | |
| Kidney | *1* | 0 | 0 | | | | |
| Pancreas | *1* | 0 | 0 | | | | |

[a]RT–PCR with primers bage14/bage19. [b]RT–PCR with primers bage20/bage21.

plete: (i) *BAGE*-related sequences that are not yet characterised map to chromosomes 14, 15, and 22; and (ii) most of human juxtacentromeric regions are still unsequenced.[20]

The sizes of the new *BAGE* transcripts isolated in this work are consistent with those of the two mRNA species (1 and 2.4 kb) previously identified in Northern blot experiments.[1] *BAGE1* transcripts have a great variety of alternative terminal exons. Eleven per cent of genes undergoing alternative splicing have alternative terminal exons.[21]

*BAGE* genes and gene fragments share extensive regions of 97–99% nucleotide identity which may be accounted for by concerted evolution. Concerted evolution is supposed to be the molecular mechanism responsible for sequence conservation among human ribosomal genes. Ribosomal genes map to the short arms of the five acrocentric chromosomes and participate in the formation of a common nucleolus. In the human germline cells, acrocentrics undergo frequent interchromosome DNA exchanges.[22,23] *BAGE* sequences may therefore undergo similar interchromosome exchanges that promote sequence homogeneity.

In addition, we observed that *BAGE* genes have a higher nucleotide identity than *BAGE* gene fragments. Interestingly, we have found that *BAGE* genes, but not the *BAGE* gene fragments, are under selective pressure and that the *BAGE* gene family was generated by chromosome rearrangements during the evolution of hominoids (De Sario, personal communication).

Similar to *BAGE1*, the new *BAGE* genes are expressed in different cancer cells but silent in normal tissues other than testis. Chromatin compaction and/or DNA methylation may account for gene silencing. DNA methylation was already shown to be the primary mechanism responsible for the silencing of the genes encoding the MAGE antigens.[24,25]

Given their expression profile restricted to cancer cells, the *BAGE* genes isolated during this work may encode new tumour antigens useful for cancer immunotherapy. *BAGE2,* which is expressed in 22% of melanomas and has a putative coding region of 309 residues, is the most promising.

Other genes located in the regions flanking human centromeres have a cancer/testis expression profile: *TPTE*

(*Transmembrane Phosphatase with Tensin homology*)[26] and *CT2* (*Creatin Transporter2*)[27] mapping to the juxtacentromeric regions of chromosomes 21 and 16, respectively, are exclusively expressed in testis; an *NF1*-related gene mapping next to the centromere of chromosome 15 is only expressed in neuroblastoma.[28] These results lead us to suggest that the restricted pattern of expression is a feature of the few genes mapping to juxtacentromeric regions and that these are candidate genes encoding tumour antigens.

**Acknowledgements**

*A De Sario acknowledges the Ligue contre le Cancer Comité du Gard and the French Embassy in Brussels. M Ruault was supported by French MRT (Ministère de la Recherche et Technologie) and by a FEBS short term fellowship. We thank Patrick Atger for the artwork.*

**References**

1 Boël P, Wildmann C, Sensi ML *et al*: BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity* 1995; **2**: 167–175.
2 Van der Bruggen P, Traversari C, Chomez P *et al*: A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science* 1991; **254**: 1643–1647.
3 De Plaen E, Arden K, Traversari C *et al*: Structure, chromosomal localization, and expression of 12 genes of the MAGE family. *Immunogenetics* 1994; **40**: 360–369.
4 Chen YT, Scanlan MJ, Sahin U *et al*: A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc Natl Acad Sci USA* 1997; **94**: 1914–1918.
5 Lethé B, Lucas S, Michaux L *et al*: LAGE-1, a new gene with tumor specificity. *Int J Cancer* 1998; **76**: 903–908.
6 Van den Eynde B, Peeters O, De Backer O, Gaugler B, Lucas S, Boon T: A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma. *J Exp Med* 1995; **182**: 689–698.
7 De Backer O, Arden KC, Boretti M *et al*: Characterization of the GAGE genes that are expressed in various human cancers and in normal testis. *Cancer Res* 1999; **59**: 3157–3165.
8 Lurquin C, De Smet C, Brasseur F *et al*: Two members of the human MAGEB gene family located in Xp21.3 are expressed in tumors of various histological origins. *Genomics* 1997; **46**: 397–408.
9 Lucas S, De Smet C, Arden KC *et al*: Identification of a new MAGE gene with tumor-specific expression by representational difference analysis. *Cancer Res* 1998; **58**: 743–752.
10 Van den Eynde BJ, Morel S: Differential processing of class-I-restricted epitopes by the standard proteasome and the immunoproteasome. *Curr Opin Immunol* 2001; **13**: 147–153.
11 Haas GG, O'Cruz OJ, De Bault LE: Distribution of human leukocyte antigen-ABC and –D/DR antigens in the unfixed human testis. *Am J Reprod Immunol Microbiol* 1988; **18**: 47–51.
12 Marchand M, Brichard V, van Baren N, Coulie PG: Biological and clinical developments in melanoma vaccines. *Expert Opin Biol Ther* 2001; **1**: 497–510.
13 Jäger E, Jäger D, Knuth A: Clinical cancer vaccine trials. *Curr Opin Immunol* 2002; **14**: 178–182.
14 Hattori M, Fujiyama A, Taylor TD *et al*: The DNA sequence of human chromosome 21. *Nature* 2000; **405**: 311–319.
15 Sambrook J, Fritsch EF, Maniatis T: Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989.
16 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic Local Alignment Search Tool. *J Mol Biol* 1990; **215**: 403–410.
17 Fantes JA, Bickmore WA, Fletcher JM, Ballesta F, Hanson IM, van Heyningen V: Submicroscopic deletions at the WAGR locus, revealed by nonradioactive in situ hybridization. *Am. J. Hum. Genetics* 1992; **51**: 1286–1294.
18 Dubois BL, Naylor SL: Characterization of NIGMS human/rodent somatic cell hybrid mapping panel 2 by PCR. *Genomics* 1993; **16**: 315–319.
19 Davis LG, Battey JF: *Basic Methods in Molecular Biology*. New York: Elsevier, 1986, pp 130–135.
20 Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE: Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res* 2001; **11**: 1005–1017.
21 Mironov AA, Fickett JW, Gelfand MS: Frequent alternative splicing of human genes. *Genome Res* 1999; **9**: 1288–1293.
22 Ferguson-Smith MA: The sites of nucleolus formation in human pachytene chromosomes. *Cytogenetics* 1964; **3**: 124–134.
23 Mirre C, Hartung M, Stahl A: Association of ribosomal genes in the fibrillar center of the nucleolus: a factor influencing translocation and nondisjunction in the human meiotic oocyte. *Proc Natl Acad Sci USA* 1980; **77**: 6017–6021.
24 De Smet C, De Backer O, Faraoni I, Lurquin C, Brasseur F, Boon T: The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proc Natl Acad Sci USA* 1996; **93**: 7149–7153.
25 De Smet C, Lurquin C, Lethé B, Martelange V, Boon T: DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with CpG-rich promoter. *Mol Cell Biol* 1999; **19**: 7327–7335.
26 Guipponi M, Yaspo ML, Riesselman L *et al*: Genomic structure of a copy of the human TPTE gene which encompasses 87 kb on the short arm of chromosome 21. *Hum Genet* 2000; **107**: 127–131.
27 Iyer GS, Krahe R, Goodwin LA *et al*: Identification of a testis-expressed creatine transporter gene at 16p11.2 and confirmation of the X-linked locus to Xq28. *Genomics* 1996; **34**: 143–146.
28 Legius E, Marchuk DA, Hall BK *et al*: NF1-related locus on chromosome 15. *Genomics* 1992; **13**: 1316–1318.