

# Gene expression profiling and real-time PCR analyses identify novel potential cancer-testis antigens in multiple myeloma

Maud Condomines<sup>1,2,3</sup>, Dirk Hose<sup>4,5</sup>, Thierry Rème<sup>1,2,3</sup>, Guilhem Requirand<sup>1,2,3</sup>, Michael Hundemer<sup>4,5</sup>, Matthieu Schoenhals<sup>1,2</sup>, Hartmut Goldschmidt<sup>5</sup>, Bernard Klein<sup>1,2,3\*</sup>

<sup>1</sup> *Biothérapie des cellules souches normales et cancéreuses INSERM : U847, Institut de recherche en biothérapie, Université Montpellier I, CHRU Montpellier, IRB - CHRU Saint-Eloi 80 Avenue Augustin Fliche 34295 MONTPELLIER Cedex 5, FR*

<sup>2</sup> *IRB, Institut de recherche en biothérapie CHRU Montpellier, Université Montpellier I, Hôpital Saint-Eloi 34000 Montpellier, FR*

<sup>3</sup> *CHU de Montpellier CHRU Montpellier, FR*

<sup>4</sup> *Nationales Centrum für Tumorerkrankungen INF350, Heidelberg, DE*

<sup>5</sup> *Medizinische Klinik und Poliklinik V Universitätsklinikum Heidelberg, Universitätsklinikum Heidelberg INF410 69115 Heidelberg, DE*

\* Correspondence should be addressed to: Bernard Klein <bernard.klein@inserm.fr >

## Abstract

Cancer-testis (CT) antigens are attractive targets for immunotherapeutic strategies since they are aberrantly expressed in malignant cells and not, or in limited number, in somatic tissues, except germ cells. To identify novel CT genes in multiple myeloma, we used Affymetrix HG-U133 gene expression profiles of 5 testis, 64 primary myeloma cell (MMC) and 24 normal tissue (NT) samples. A 5-filter method was developed to keep known CT genes while deleting non-CT genes. Starting from 44928 probe sets, including probe sets for 18 previously-described CT genes, we have obtained 82 genes expressed in MMC and testis and not detected in more than 6 NT. This list includes 14 of the 18 known CT genes and 68 novel putative CT genes. Real-time RT-PCR was performed for 34 genes in 12 NT, 5 MMC samples and one sample of 5 pooled testes. It has validated the CT status of 23/34 genes (67%). We found one novel “testis-restricted” gene (expression in testis and tumor only) – TEX14 –, 8 “tissue-restricted” (mRNA detected in 1 or 2 non-gametogenic tissues), and 7 “differentially expressed” (mRNA detected in three to six non-gametogenic tissues) CT genes. Further studies are warranted to determine the immunogenicity of these novel CT antigen candidates.

**MESH Keywords** Antigens, Neoplasm ; genetics ; Case-Control Studies ; Gene Expression Profiling ; Humans ; Male ; Middle Aged ; Multiple Myeloma ; genetics ; pathology ; Polymerase Chain Reaction ; RNA ; analysis ; Testis ; immunology ; metabolism ; Tissue Distribution

**Author Keywords** Human ; Tumor antigen ; Immunotherapy

## Introduction

The development of successful immunotherapeutic strategies requires the identification and characterization of tumor-associated antigens (TAA) that will be recognized by the host immune system, leading to tumor rejection. Cancer-testis (CT) antigens expressed by germ cells and malignant cells are shared tumor-specific antigens. To date, 47 CT gene families including more than 90 genes have been identified by either immunological screening methods or expression database analysis (1–4). Scanlan et al. classified CT genes into four categories according to their expression profiles: (i) testis-restricted (mRNA detected in testis and tumor samples only), (ii) “tissue restricted” (mRNA detected in 2 or fewer non-gametogenic tissues), (iii) “differentially expressed” (mRNA detected in three to six non-gametogenic tissues), and (iv) “ubiquitously expressed” (mRNA detected in more than six non-gametogenic tissues) (1). The testis-restricted genes, comprising about one half of the CT genes, in particular genes belonging to the MAGE, GAGE and SSX families, are mainly located on chromosome X. Because of their expression being restricted to germ cells and malignant tissues, there should be no deletion of the high affinity T-cell repertoire and - theoretically - a low risk of pre-existing immune tolerance. Thus, the testis-restricted CT antigens are used as targets in several cancer vaccination trials (5–7).

The expression of several CT genes and gene products by malignant plasma cells isolated from patients with multiple myeloma (MM) has been described (8, 9). Among these, MAGE-C1 is the most frequently expressed one (10, 11). NY-ESO-1 and MAGE-A3 protein expression correlate with higher plasma cell proliferation and poor prognosis (11–13). Using DNA-microarray analysis, we reported that 35 CT genes can be detected at the mRNA level in myeloma cells in at least one out of 64 MM patients. 18 of these 35 CT genes were expressed in more than 10% of the patients (11). This study also confirmed the adverse prognosis value of 6 CT antigens.

CT antigens are immunogenic in MM patients since specific CD8+ T lymphocytes anti-MAGE-A1-4 and LAGE-1 have been detected in peripheral blood of MM patients (14) and at least four testis-restricted CT antigens were the targets of T-cell response in patients receiving allotransplants (15). However, the expression of CT proteins is heterogeneous within the tumor-cell population of an individual patient (10, 12) and an immunotherapy strategy targeting only one antigen could lead to the selection of antigen negative tumor subclones (16). Thus several antigens should be used to vaccinate a given patient. As we reported that the co-expression of 3 testis-restricted CT

genes is found for 70% of the patients (11) and as the immunogenicity of the proteins encoded by most of these CT genes has not been reported in MM yet, it is of great importance to identify a maximum of CT-like genes that are expressed in MMC and for a given patient, to know the antigen cocktail that is co-expressed in MMC.

We and others have shown that microarrays are useful tools to detect TAA expression, in particular CTA, in various tumors, including MM (9, 11), renal cell carcinoma (17), breast cancer (18) and melanoma (19). In this study, our aim was to pick up new CT genes that are aberrantly expressed by malignant plasma cells compared to normal plasma cells and a large panel of normal tissues, using known CTA to validate our bioinformatics-based selection. We detected 68 potential novel CT genes fitting these criteria and the expression of 16 of these was validated by real-time RT-PCR.

## Materials and methods

### Patients and cell samples

Primary MM cells (MMC) were purified from 64 consecutive patients with newly-diagnosed MM (median age, 59 years). According to Durie-Salmon classification, 11 patients were of stage IA, 11 of stage IIA, 39 of stage IIIA and 3 of stage IIIB. Twelve patients had IgAk MM, 7 IgA $\lambda$  MM, 24 IgG $\kappa$  MM, 10 IgG $\lambda$  MM, 6 Bence-Jones  $\kappa$  MM, 3 Bence-Jones  $\lambda$  MM, and 2 non-secreting MM. Bone marrow samples were obtained from healthy donors and patients after informed consent was given in agreement with French or German laws. Normal bone marrow plasma cells (BMPC) and primary MMC were purified using anti-CD138 MACS microbeads. Briefly, bone marrow aspirates were subjected to density centrifugation. Plasma cells were sorted from mononuclear cells to purity > 85% using CD138-microbeads and automated magnetic cell-sorting (Miltenyi Biotec, Bergisch-Gladbach, Germany). Purity was assessed by flow cytometry (FACS Calibur, Becton Dickinson, Heidelberg, Germany) after CD38/CD138 double-staining. Routine smears were assessed by light microscopy. For the isolation of peripheral blood memory B cells (MBC), monocytes, NK and T cells were first removed using anti-CD14, anti-CD16 and anti-CD3 magnetic beads (Dyna), and MBC were then positively selected using anti-CD27 MACS microbeads (Miltenyi Biotec). XG human myeloma cell lines (HMCL) were obtained and characterized in our laboratory (20–23). SKMM, OPM2, LP1 and RPMI8226 HMCLs were purchased from ATTC (LGC Promochem, France).

### Preparation of complementary RNA (cRNA) and microarray hybridization

Testis RNA samples from healthy donors were purchased from Clinisciences (Montrouge, France). RNA extraction was performed using the RNeasy kit (Qiagen, Hilden, Germany), the SV-total RNA extraction kit (Promega, Mannheim, Germany) or Trizol (Invitrogen) in accordance with the manufacturer's instructions. RNA was analyzed using an Agilent 2100 bioanalyzer. Labelled cRNA was generated using the small sample labelling protocol II (Affymetrix, Santa Clara, USA), and hybridized to HG-U133 A+B GeneChip microarrays (Affymetrix) or HG-U133 2.0 plus arrays according to the manufacturer's instructions (24). Fluorescence intensities were quantified and analyzed using the GCOS software (Affymetrix). Arrays were scaled to an average intensity of 100. A threshold of 1 was assigned to values  $\leq 1$ . Gene expression data are deposited in the Array Express public database (<http://www.ebi.ac.uk/microarray-as/ae/>, accession number n° E-MTAB-81).

### Gene expression profiling analysis

All gene expression data were normalized with the MAS5 algorithm using the same global scaling and analyzed with our bioinformatics platforms (RAGE, <http://amazonia.montp.inserm.fr>) and Amazonia!, <http://amazonia.montp.inserm.fr>) (25). The Affymetrix call defined as the statistical assignment of a probe set as "Present" or "Absent" provided a threshold of gene expression. 6% of the probe sets had a marginal call in normal tissue samples using HG-U133A microarray. The "Marginal" call was considered as "Present" because we previously found that some genes interrogated with probe sets with a marginal call often display a positive expression by real-time RT-PCR. As discussed below, the decision to consider as Present or Absent the probe sets with a marginal call did not affect the final results. CT gene expression was assessed in 64 primary MMC, 20 HMCL samples, 8 plasma cell samples from patients with MGUS, 5 normal testis samples, 7 BMPC samples and 7 MBC samples, using HG-U133 Set arrays and in 10 whole bone marrow samples from healthy donors using HG-U133 Plus 2.0 arrays. We also used the gene expression profiling (GEP) from 24 human normal somatic tissues determined with Affymetrix HG-U133A and custom-designed GNFH1 arrays, available from Dr Hogenesch's group on a public database ((26) <http://symatlas.gnf.org/SymAtlas>). For the tissue types containing two to four different samples per tissue, i.e. lung, trachea, heart, prostate, liver, pancreas, kidney, adrenal gland, lymph nodes, thymus, salivary gland, thyroid, pituitary, uterus, tongue, skin, tongue, smooth muscle, psoas muscle, intestine and appendix, a probe set was stated "not expressed" by the tissue if no Present call occurred for 2–4 samples. When data of more than 7 samples of a given tissue were available, i.e. blood, bone marrow, central and peripheral nervous systems, a probe set was stated "expressed" by the tissue if at least 3 Present calls were found. Normal BMPC samples hybridized to the HG-U133 Set arrays were considered as the 25<sup>th</sup> normal tissue sample. Gene expression data of other cancers were obtained from the Oncomine Cancer Microarray database (<http://www.oncomine.org>) (27).

### Real-time RT-PCR

Total RNA derived from pooled normal tissues was obtained from Clontech (San Jose, CA) and Clinisciences. The 5 testis RNA samples used in microarray hybridization were pooled in one testis RNA sample. RNA samples of MMC were those used for the microarray hybridization. We generated cDNA from 500 ng of total RNA using the Superscript™ First-Strand Synthesis System for RT-PCR (Invitrogen, Cergy Pontoise, France), according to manufacturer's instructions. Real-time PCR was performed with Taqman® gene expression assays and the TaqMan® Universal Master Mix from Applied Biosystems (Courtaboeuf, France) using the ABI Prism 7000 Sequence Detection System. Quantitative PCR (qPCR) analysis was completed using ABI PRISM 7000 SDS Software. Ct values were collected for GAPDH and the genes of interest during the log phase of the cycle. Gene of interest levels were normalized to GAPDH for each sample ( $\delta Ct = Ct \text{ gene of interest} - Ct \text{ GAPDH}$ ) and compared with the values obtained for the testis sample, used as a positive control, using the following formula  $100/2^{\delta\delta Ct}$  where  $\delta\delta Ct = \delta Ct \text{ unknown} - \delta Ct \text{ positive control}$ . The testis sample used as a positive control was assigned the arbitrary value of 100. A  $\delta\delta Ct$  value threshold  $<10$  (i.e.;  $< 1/10$  of that in testis) was arbitrarily designed as a negative expression.

### Western blot analysis

Cells were lysed in 10 mM Tris-HCl (pH 7.05), 50 mM NaCl, 50 mM NaF, 30 mM sodium pyrophosphate (NaPPi), 1% triton X-100, 5  $\mu\text{M}$   $\text{ZnCl}_2$ , 100  $\mu\text{M}$   $\text{Na}_3\text{VO}_4$ , 1 mM DTT, 20 mM  $\beta$ -glycerophosphate, 20 mM p-nitrophenolphosphate (PNPP), 20  $\mu\text{g/ml}$  aprotinin, 2.5  $\mu\text{g/ml}$  leupeptin, 0.5 mM PMSF, 0.5 mM benzamidine, 5  $\mu\text{g/ml}$  pepstatin, and 50 nM okadaic acid. Lysates were resolved on 12% sodium dodecyl sulfate-polyacrylamide by gel electrophoresis (SDS-PAGE) and transferred to a nitrocellulose membrane (Schleicher and Schuell, Kassel, Germany). Membranes were blocked for 2 hours at room temperature in 140 mM NaCl, 3 mM KCl, 25 mM Tris-HCl (pH 7.4), 0.1 % Tween 20 (TBS-T), 5% skimmed milk, and then immunoblotted with a mouse anti-MORC (Abnova Corporation, Taipei, Taiwan), a rabbit anti-ARX, anti-ELOVL4 (Abcam, Cambridge, United Kingdom), a rabbit anti-TMEFF2 (Proteintech Group, Chicago, IL), a rabbit anti-CTNNA2 (GeneTex, San Antonio, TX), or with a mouse anti-TMEFF1 (R&D Systems, Minneapolis, MN) antibody. As a control for protein loading, we used a mouse monoclonal anti- $\beta$ -actin antibody (Sigma, St Louis, MO). The primary antibodies were visualized with goat anti-rabbit (Sigma) or goat anti-mouse (Bio-Rad, Hercules, CA) peroxidase-conjugated antibodies by an enhanced chemiluminescence detection system.

## Results

### Selection of candidate CT genes using Affymetrix microarrays

In an attempt to find novel genes showing an expression pattern of known CT genes, i.e. expressed in testis and cancer cells but not or poorly expressed in normal tissues, we have defined 5 consecutive filters making it possible to reduce considerably the probe set number while retaining a maximum of CT genes already known to be expressed in MMC. We used for that the 18 known CT genes expressed in MMC of more than 10% of MM patients identified in our previous study (11). We also used some filters with the Affymetrix-assigned detection call that enables to mix data from different Affymetrix microarrays, avoiding the normalization problem. This call indicates whether a gene is "Present" or "Absent". As depicted in Fig. 1A, we selected probe sets displaying simultaneously a Present call in at least 6/64 MMC samples and in at least 3/5 testis samples leading to a selection of 16982 probe sets out of the 44928 probe sets available on the HG-U133 A+B arrays (filter 1). At this step, 2 known CT genes (MAGE-A1 and XAGE-1) were not retained because their corresponding probe sets were expressed in less than 3 testis samples. Filter 2 then selected 2559 probe sets with mean signal in MMC samples having a Present call at least 2.5-fold higher than the mean signal observed in normal MBC. The 2.5 ratio was chosen because it is the inflexion point of the curves delineating the decrease of the total number of probe sets and that of known CT genes according to increasing ratios (Fig. 1B). This second filter made it possible to focus on genes overexpressed in MMC, while retaining the majority of known CT genes and decreasing the total number of probe sets. Filter 3 consisted of selecting probe sets expressed in less than 7 normal tissues ( $\leq 6$  NT) out of 25, as described in material and methods, leaving 341 probe sets. When several probe sets are available for a given gene, it was retained only if all its corresponding probe sets were expressed in less than 7 NT. For probe sets located on the HG-U133B array, we could find a corresponding probe set by matching gene names on GNFIH array data from Su et al for only 353 out of 1098 probe sets (26) (Fig. 1A). Some poorly working probe sets may give a signal whose height is similar in samples with an Absent or a Present call. In order to eliminate such probe sets, we calculated the ratio between the mean signal in MMC samples with a Present call and the mean signal in MMC samples with an Absent call and filter 4 removed probe sets with a ratio  $< 2.5$ , yielding 143 probe sets including 14 known CT genes. This 2.5 ratio was the inflexion point of the curve drawn in Fig. 1C, making it possible to delete a maximum of probe sets while keeping all known CT gene probe sets. Finally, we apply a bibliography filter (filter 5) removing genes whose expression in human NT reported in the literature differed from the current results with microarrays. The final list contained 82 probe sets/genes, including 14 out of the starting 18 known CT genes. As indicated in Materials and Methods, the 6% of probe sets with a marginal call were considered as probe sets with a Present call. But considering these probe sets as probe sets with an Absent call and running the 5 filters yielded to the same final list of 82 genes. Altogether, these 5 filters result in a 207-fold enrichment in known CT genes, from the initial 16982-probe set list to the final 82-probe set list. Genes were assigned into the three CT gene categories, according to their pattern of expression in NT (Table S1). As shown in Fig. 2, we found 16 (19.5%) testis-restricted CT genes including 10 known CT genes and 6 novel potential CT genes; 31 (37.8%) tissue-restricted CT genes including 2 known CT genes and 29 novel potential CT genes; 35 (42.7%) differentially

expressed CT genes including 2 known CT genes and 33 novel potential CT genes. 16 genes out of 82 were located on chromosome X including 11 known and 5 novel putative CT genes (Fig. 2 ).

### Classification of candidate CT genes according to their expression pattern determined by real-time RT-PCR

We then investigated the pattern of expression of 7 known CT genes and 27 potential CT genes among 12 normal tissues, MMC samples from 5 different patients with MM and one sample of 5 pooled testes by real time RT-PCR. No detectable expression of 6 known CT genes, i.e. DDX43, MAGE-A3, MAGE-C1, MAGE-C2, MORC, and SSSX1 , was observed in each of the 12 normal tissues compared to testis, in agreement with their known “testis-restricted” status (Fig. 3A ). A slight expression (qPCR  $\delta\delta Ct$  value of 12) of CTAG1B/NY-ESO-1 was detected in the brain. Regarding the novel potential CT genes identified by microarray analyses, real-time RT-PCR was performed for 6/6 genes classified in the “testis-restricted” category, for 18/29 genes classified in the “tissue-restricted” category and not expressed in bone marrow and for 3/33 genes in the “differentially expressed” category. Among the 27 genes investigated, 6 were poorly expressed in MMC (mean < 10) and 2 were actually not expressed in testis. Among the remaining 19 genes, 3 were expressed in MMC and testis but were also expressed in more than 6 NT, thus classified in the “ubiquitously expressed” category (Table I ). TEX14 appeared as a novel testis-restricted CT gene (Fig. 3B ). 8 genes were expressed in 1-2 NT, i.e. CTNNA2, LOC130576, RP1-32F7.2/FAM133A, ANKRD45, ELOVL4, IGSF11, TMEFF1 , and TMEFF2 , thus classified in the “tissue-restricted” category (Fig. 3B ) and 7 genes, i.e. ARX, FLJ23577, LOC150763, MGC3040, NOL4, PTPN20A/B and SPAG4 , belong to the “differentially expressed” category since they were expressed in 3 to 6 NT (Fig. 3C ). CTNNA2 and FAM133A were expressed only in brain and testis. Table I summarizes real-time RT-PCR data of all known and novel CT genes investigated and provides the level of expression of MMC compared to testis and NT as well as the frequency of expression of each CT gene within the 64 patients. Altogether, by using a primary selection with microarrays and a real-time RT-PCR validation, we found 16 novel potential CT antigens expressed in MM.

Fig. 4 shows the HG-U133 Set array signal levels and the Affymetrix call of 9 testis and tissue-restricted CT genes in 5 testis samples, purified plasma cells from 8 patients with MGUS and MMC from 64 newly-diagnosed patients. As reported in Table I , we can observe that the expression of all the novel CT genes is heterogeneous within the MM patient population. FAM133A is the more frequently expressed (53% of patients with MMC having a Present call) and TMEFF2 is the less frequently expressed (11%).

### Protein study and link with clinical parameters and patients' survival

As antibodies for 6 out of the 68 new putative CT antigens – MORC, ARX, TMEFF2, TMEFF1, CTNNA2, and ELOVL4 – are commercially available, we investigated the protein expression in testis and 4 myeloma cell lines (2 expressing the gene by RT-PCR and 2 not expressing it). The results were not conclusive because either the protein size in testis had not the expected molecular size or protein expression in myeloma cell lines or testis did not fit with RT-PCR data. We have to point out that these antibodies have to date never been validated by academic teams and their specificity should be questioned. We also found no significant link between the expression of these novel potential CT genes with event-free and overall survivals after treatment with high-dose chemotherapy and autologous stem cell transplantation (results not shown).

### Expression in other cancers

As most known-CT genes are expressed in a broad variety of cancer, we investigated the expression of the 16 microarray and PCR validated novel CT genes in 41 cancer types, using the database ONCOMINE which gathers many published microarray analyses. The novel CT genes were significantly overexpressed in malignant cells compared to normal counterparts of 18 of the 40 cancer types (Table 2 ). 14 out of the 16 genes were overexpressed in at least one cancer type, other than MM. TMEFF1 and SPAG4 were expressed in 6 different cancer types.

## Discussion

We present here a new 5-filter method to find novel potential CT antigens expressed in human myeloma cells based on the GEP of testis, MMC and normal tissues. In order to use publicly available data for normal tissues and to bypass the normalization issue, we developed filters using the Affymetrix call, making it possible to compare data from different hybridization processes or different microarray types. MM is a heterogeneous disease with at least 7 molecular entities defined by GEP (28 ). To avoid loss of information due to this MM heterogeneity, we selected MMC samples with a “Present” call for each probe set to compare Affymetrix signals in MMC to those in normal MBC (R1). The filter criteria were defined in order to keep a maximum of known CT genes while deleting a maximum of non-CT genes, which provided a 207-fold enrichment of known CT genes and yields to 68 novel putative CT genes shared by MMC and testis. We are aware that we likely have missed some new potential CT genes. As our GEP data are publicly available (accession number E-MTAB-81), the current study may encourage the design of new filters to pick up additional CT genes. In addition, as indicated in Table 2 , this 5-filter method can be easily extended to other cancer diseases.

This microarray-based method is a first step to pick up rapidly new CT genes and real-time RT-PCR is mandatory to check the information. First, real-time RT-PCR allowed confirming the pattern of expression of 7 known CT genes among 13 normal tissues

(including testis) and MMC samples and their testis-restricted status. Secondly, real-time RT-PCR confirmed the CT status of 16 novel CT genes out of the 27 tested (60%). This validation highlights the interest of the 5-filter method to rapidly select new putative CT genes. TEX14 is a testis restricted protein localized to germ cell intercellular bridges which is involved in spermatogenesis and fertility (29). CTNNA2 for catenin alpha-2, is a structural component of cytoskeleton associated with cadherin proteins and involved in cell adhesion. In particular, it stabilizes synaptic contacts (30). ELOVL4, for Elongation of very long chain fatty acids-like 4, is a G-protein expressed in retina and particularly in the endoplasmic reticulum of photoreceptor cells (31). A 5-bp deletion in the ELOVL4 gene cause macular degeneration (32). IGSF11 was previously described as expressed in testis and brain only (33). We confirmed here its expression in testis and brain but showed that IGSF11 is also expressed in adrenal gland. This gene is up-regulated in gastrointestinal and hepatocellular carcinomas, encodes a receptor essential for gastric cancer cell growth (34, 35), and has been described as a novel immunogenic target since a natural HLA-A2 restricted T-cell epitope was identified and a subsequent anchor-modified peptide was synthesized (35). TMEFF1 and TMEFF2 for transmembrane protein with EGF-like and two follistatin-like domains 1 and 2, respectively, encode putative growth factors predominantly expressed in the brain (36, 37). Soluble TMEFF2 is an activating ERBB4 ligand (38) and is part of the neuregulin signalling pathway which contributes to the pathogenesis of MM (39). TMEFF2 has been characterized as a survival factor for hippocampal and mesencephalic neurons (36) but is methylated in colorectal cancer (40) and downregulated in androgen-independent prostate cancer (41). TMEFF1 inhibit the proliferation of brain cancer cell lines (37). Thus, TMEFF1 and TMEFF2 may be tumor suppressor genes.

PTPN20 is a nuclear phosphatase whose mRNA was previously detected among normal tissues in testis only, but is overexpressed in cancer cell lines (42). Its biological role is not elucidated yet. ARX, for Aristaless-related homeobox, is a transcription factor playing a crucial role in development with multiple transcript isoforms (43). Mutations in this gene cause X-linked mental retardation syndromes (44, 45). SPAG4, for sperm associated antigen 4, binds ODF proteins in sperm tails involved in spermatozoid motility (46). Its expression has already been reported in malignant cells of different cancers and the authors suggested that SPAG4 could be a novel cancer marker (47).

ANKRD45, FAM133A, NOL4, LOC130576, FLJ23577, LOC150763, and MGC3040 encode putative proteins with unknown functions.

For these 16 microarray- and PCR-validated CT genes, it will be important to check for protein expression. This is also the case for some known CT genes for which protein expression has never been evidenced, such as MORC or DDX43. One of the main difficulties is the lack of validated antibodies. Indeed, we failed to obtain conclusive results in western blot assays with commercially available antibodies directed against ARX, CTNNA2, ELOVL4, TMEFF1 and TMEFF2, because of their poor specificity. No antibodies directed against the remaining proteins could be found. Of note, the commercially-available antibodies have not been validated by academic teams yet.

CT antigens are likely immunogenic because of their restricted expression. Indeed, "non-self" proteins are recognized by high avidity T cells retained during central thymic selection. There is evidence that autoreactive T cells with high avidity for ubiquitous antigens are removed by thymic and peripheral tolerance mechanisms (48, 49). However, some low-avidity T cells specific for tissue-restricted antigens escape thymic negative selection and are able to induce autoimmunity, but have a high activation-threshold (50). As a consequence, immunogenicity of tissue-restricted CT genes should be studied first and particularly genes expressed only in brain and testis since central nervous system can be considered as a relative "immune privilege" site in non-inflammatory conditions (51). Nevertheless, immunogenicity of each of these CT gene products has to be evidenced because peripheral tolerance can efficiently delete the T-cell repertoire of self-reactive T cells (52). Also, the protein expression of each of the candidate CT genes has to be investigated in normal tissues and in tumor samples in order to estimate the degree of tumor overexpression both for immunogenicity and safety concerns. So far, none of the 16 novel putative antigens, except for IGSF11, have been shown to be immunogenic. This is also the case for some previously discovered CT genes which have been categorized in the CT gene list only on the basis on their restricted expression pattern (1). The immunogenicity of a protein can be evidenced in humans either by detecting serum antibodies recognizing the given protein, and/or by detecting specific cellular immunity, i.e. CD4 and CD8 T cells recognizing protein-derived peptides. We are now focussing on this topic for 6 testis-restricted CT genes and also for MORC and MAGE-C1. Putative HLA-A2-restricted 9-mer peptides were designed using the SYFPEITHI and BIMAS softwares (53, 54), their affinity has been validated in a biological HLA-A2 binding assay (unpublished data) and we are now trying to obtain peptide-specific T cells able to specifically kill myeloma cells. The use of HLA-A2 positive HMCL as targets in cytotoxicity assays will ensure that the peptide-directed T cells are able to recognize endogenous processed antigens, naturally presented in HMC molecules of tumor cells.

93 CT genes were previously identified (<http://www.cancerimmunity.org/CTdatabase>) and this study adds 16 microarray- and PCR-validated new putative CT genes that are expressed in multiple myeloma but also in various cancers. This suggests but does not prove that the majority of CT genes, classified in the 4 categories defined by Scanlan et al (1), have been identified. It would be possible

to rapidly answer this question, identifying most CT genes expressed in various cancers. Indeed, the current 5-filter method, avoiding the normalization issue, could be easily applied to other cancers provided that gene expression of malignant cells and their normal counterpart is available.

## Footnotes:

1

This work was supported by grants from the Ligue Nationale Contre le Cancer (équipe labellisée 2006), Paris, France, from INCA (n° R07001FN) and from MSCNET European strep (N°E06005FF), the Hopp-Foundation, Germany, the University of Heidelberg, Germany, the National Centre for Tumor Diseases, Heidelberg, Germany, the Tumorzentrum Heidelberg/Mannheim, Germany.

## Authorship:

MC designed research, performed the experiments and wrote the paper. DH, MH and HG collected bone marrow samples and clinical data. VP performed microarray hybridization. GR performed qPCR experiments. TR and JDV developed bioinformatics tools to analyze the data. DH and MH participated in the writing of the paper. BK is the senior investigator who designed research and wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

## References:

- 1 . Scanlan MJ , Simpson AJ , Old LJ . 2004 ; The cancer/testis genes: review, standardization, and commentary . *Cancer Immun* . 4 : 1 -
- 2 . Wang Z , Zhang Y , Mandal A , Zhang J , Giles FJ , Herr JC , Lim SH . 2004 ; The spermatozoa protein, SLLP1, is a novel cancer-testis antigen in hematologic malignancies . *Clin Cancer Res* . 10 : 6544 - 6550
- 3 . Monji M , Nakatsura T , Senju S , Yoshitake Y , Sawatsubashi M , Shinohara M , Kageshita T , Ono T , Inokuchi A , Nishimura Y . 2004 ; Identification of a novel human cancer/testis antigen, KM-HN-1, recognized by cellular and humoral immune responses . *Clin Cancer Res* . 10 : 6047 - 6057
- 4 . Chen YT , Scanlan MJ , Venditti CA , Chua R , Theiler G , Stevenson BJ , Iseli C , Gure AO , Vasicek T , Strausberg RL , Jongeneel CV , Old LJ , Simpson AJ . 2005 ; Identification of cancer/testis-antigen genes by massively parallel signature sequencing . *Proc Natl Acad Sci U S A* . 102 : 7940 - 7945
- 5 . van Baren N , Bonnet MC , Dreno B , Khammari A , Dorval T , Piperno-Neumann S , Lienard D , Speiser D , Marchand M , Brichard VG , Escudier B , Negrier S , Dietrich PY , Maraninchi D , Osanto S , Meyer RG , Ritter G , Moingeon P , Tartaglia J , van der Bruggen P , Coulie PG , Boon T . 2005 ; Tumoral and immunologic response after vaccination of melanoma patients with an ALVAC virus encoding MAGE antigens recognized by T cells . *J Clin Oncol* . 23 : 9008 - 9021
- 6 . Jager E , Karbach J , Gnjjatic S , Neumann A , Bender A , Valmori D , Ayyoub M , Ritter E , Ritter G , Jager D , Panicali D , Hoffman E , Pan L , Oettgen H , Jold L , Knuth A . 2006 ; Recombinant vaccinia/fowlpox NY-ESO-1 vaccines induce both humoral and cellular NY-ESO-1-specific immune responses in cancer patients . *Proc Natl Acad Sci U S A* . 103 : 14453 - 14458
- 7 . Odunsi K , Qian F , Matsuzaki J , Mhawech-Fauceglia P , Andrews C , Hoffman EW , Pan L , Ritter G , Vilella J , Thomas B , Rodabaugh K , Lele S , Shrikant P , Old LJ , Gnjjatic S . 2007 ; Vaccination with an NY-ESO-1 peptide of HLA class I/II specificities induces integrated humoral and T cell responses in ovarian cancer . *Proc Natl Acad Sci U S A* . 104 : 12837 - 12842
- 8 . Pellat-Deceunynck C , Mellerin MP , Labarriere N , Jago G , Moreau-Aubry A , Harousseau JL , Jotereau F , Bataille R . 2000 ; The cancer germ-line genes MAGE-1, MAGE-3 and PRAME are commonly expressed by human myeloma cells . *European Journal of Immunology* . 30 : 803 - 809
- 9 . Tarte K , Zhan F , De Vos J , Klein B , Shaughnessy J Jr . 2003 ; Gene expression profiling of plasma cells and plasmablasts: toward a better understanding of the late stages of B-cell differentiation . *Blood* . 102 : 592 - 600
- 10 . Dhodapkar MV , Osman K , Teruya-Feldstein J , Filippa D , Hedvat CV , Iversen K , Kolb D , Geller MD , Hassoun H , Kewalramani T , Comenzo RL , Coplan K , Chen YT , Jungbluth AA . 2003 ; Expression of cancer/testis (CT) antigens MAGE-A1, MAGE-A3, MAGE-A4, CT-7, and NY-ESO-1 in malignant gammopathies is heterogeneous and correlates with site, stage and risk status of disease . *Cancer Immun* . 3 : 9 -
- 11 . Condomines M , Hose D , Raynaud P , Hundemer M , De Vos J , Baudard M , Moehler T , Pantesco V , Moos M , Schved JF , Rossi JF , Reme T , Goldschmidt H , Klein B . 2007 ; Cancer/testis genes in multiple myeloma: expression patterns and prognosis value determined by microarray analysis . *Journal of Immunology* . 178 : 3307 - 3315
- 12 . Jungbluth AA , Ely S , DiLiberto M , Niesvizky R , Williamson B , Frosina D , Chen YT , Bhardwaj N , Chen-Kiang S , Old LJ , Cho HJ . 2005 ; The cancer-testis antigens CT7 (MAGE-C1) and MAGE-A3/6 are commonly expressed in multiple myeloma and correlate with plasma-cell proliferation . *Blood* . 106 : 167 - 174
- 13 . van Rhee F , Szmania SM , Zhan F , Gupta SK , Pomtree M , Lin P , Batchu RB , Moreno A , Spagnoli G , Shaughnessy J , Tricot G . 2005 ; NY-ESO-1 is highly expressed in poor-prognosis multiple myeloma and induces spontaneous humoral and cellular immune responses . *Blood* . 105 : 3939 - 3944
- 14 . Goodyear O , Piper K , Khan N , Starczynski J , Mahendra P , Pratt G , Moss P . 2005 ; CD8+ T cells specific for cancer germline gene antigens are found in many patients with multiple myeloma, and their frequency correlates with disease burden . *Blood* . 106 : 4217 - 4224
- 15 . Atanackovic D , Arfsten J , Cao Y , Gnjjatic S , Schnieders F , Bartels K , Schilling G , Faltz C , Wolschke C , Dierlamm J , Ritter G , Eiermann T , Hossfeld DK , Zander AR , Jungbluth AA , Old LJ , Bokemeyer C , Kroger N . 2007 ; Cancer-testis antigens are commonly expressed in multiple myeloma and induce systemic immunity following allogeneic stem cell transplantation . *Blood* . 109 : 1103 - 1112
- 16 . Dudley ME , Wunderlich JR , Yang JC , Sherry RM , Topalian SL , Restifo NP , Royal RE , Kammula U , White DE , Mavroukakis SA , Rogers LJ , Gracia GJ , Jones SA , Mangiameli DP , Pelletier MM , Gea-Banacloche J , Robinson MR , Berman DM , Filie AC , Abati A , Rosenberg SA . 2005 ; Adoptive cell transfer therapy following non-myceloablative but lymphodepleting chemotherapy for the treatment of patients with refractory metastatic melanoma . *J Clin Oncol* . 23 : 2346 - 2357
- 17 . Weinschenk T , Gouttefangeas C , Schirle M , Obermayr F , Walter S , Schoor O , Kurek R , Loeser W , Bichler KH , Wernet D , Stevanovic S , Rammensee HG . 2002 ; Integrated functional genomics approach for the design of patient-individual antitumor vaccines . *Cancer Res* . 62 : 5818 - 5827
- 18 . Cavallo F , Astolfi A , Iezzi M , Cordero F , Lollini PL , Forni G , Calogero R . 2005 ; An integrated approach of immunogenomics and bioinformatics to identify new Tumor Associated Antigens (TAA) for mammary cancer immunological prevention . *BMC Bioinformatics* . 6 : (Suppl 4 ) S7 -
- 19 . Segal NH , Blachere NE , Guevara-Patino JA , Gallardo HF , Shiu HY , Viale A , Antonescu CR , Wolchok JD , Houghton AN . 2005 ; Identification of cancer-testis genes expressed by melanoma and soft tissue sarcoma using bioinformatics . *Cancer Immun* . 5 : 2 -
- 20 . Zhang XG , Gaillard JP , Robillard N , Lu ZY , Gu ZJ , Jourdan M , Boiron JM , Bataille R , Klein B . 1994 ; Reproducible obtaining of human myeloma cell lines as a model for tumor stem cell study in human multiple myeloma . *Blood* . 83 : 3654 - 3663
- 21 . Rebouissou C , Wijdenes J , Autissier P , Tarte K , Costes V , Liautard J , Rossi JF , Brochier J , Klein B . 1998 ; A gp130 interleukin-6 transducer-dependent SCID model of human multiple myeloma . *Blood* . 91 : 4727 - 4737
- 22 . Tarte K , Zhang XG , Legouffe E , Hertog C , Mehtali M , Rossi JF , Klein B . 1999 ; Induced expression of B7-1 on myeloma cells following retroviral gene transfer results in tumor-specific recognition by cytotoxic T cells . *Journal of Immunology* . 163 : 514 - 524
- 23 . Gu ZJ , De Vos J , Rebouissou C , Jourdan M , Zhang XG , Rossi JF , Wijdenes J , Klein B . 2000 ; Agonist anti-gp130 transducer monoclonal antibodies are human myeloma cell survival and growth factors . *Leukemia* . 14 : 188 - 197

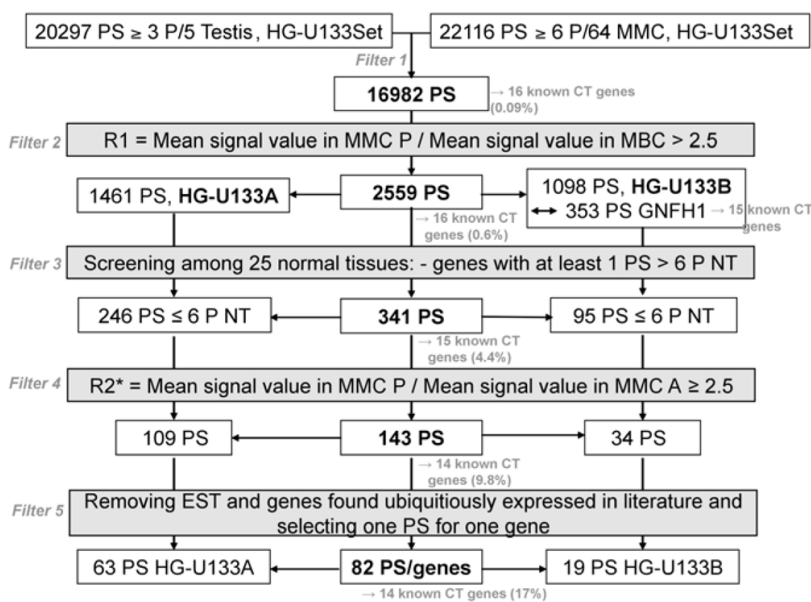
- 24. De Vos J, Thykjaer T, Tarte K, Ensslen M, Raynaud P, Requirand G, Pellet F, Pantesco V, Reme T, Jourdan M, Rossi JF, Orntoft T, Klein B. 2002; Comparison of gene expression profiling between malignant and normal plasma cells with oligonucleotide arrays. *Oncogene*. 21 : 6848 - 6857
- 25. Assou S, Le Carrour T, Tondeur S, Strom S, Gabelle A, Marty S, Nadal L, Pantesco V, Reme T, Hugnot JP, Gasca S, Hovatta O, Hamamah S, Klein B, De Vos J. 2007; A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells*. 25 : 961 - 973
- 26. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. 2004; A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A*. 101 : 6062 - 6067
- 27. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM. 2004; Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci U S A*. 101 : 9309 - 9314
- 28. Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S, Epstein J, Yacoby S, Sawyer J, Burington B, Anaissie E, Hollmig K, Pineda-Roman M, Tricot G, van Rhee F, Walker R, Zangari M, Crowley J, Barlogie B, Shaughnessy JD Jr. 2006; The molecular classification of multiple myeloma. *Blood*. 108 : 2020 - 2028
- 29. Greenbaum MP, Yan W, Wu MH, Lin YN, Agno JE, Sharma M, Braun RE, Rajkovic A, Matzuk MM. 2006; TEX14 is essential for intercellular bridges and fertility in male mice. *Proc Natl Acad Sci U S A*. 103 : 4982 - 4987
- 30. Abe K, Chisaka O, Van Roy F, Takeichi M. 2004; Stability of dendritic spines and synaptic contacts is controlled by alpha N-catenin. *Nat Neurosci*. 7 : 357 - 363
- 31. Ambasudhan R, Wang X, Jablonski MM, Thompson DA, Lagali PS, Wong PW, Sieving PA, Ayyagari R. 2004; Atrophic macular degeneration mutations in ELOVL4 result in the intracellular misrouting of the protein. *Genomics*. 83 : 615 - 625
- 32. Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, Li Y, Metzker ML, Allikmets R, Zack DJ, Kakuk LE, Lagali PS, Wong PW, MacDonald IM, Sieving PA, Figueroa DJ, Austin CP, Gould RJ, Ayyagari R, Petrukhin K. 2001; A 5-bp deletion in ELOVL4 is associated with two related forms of autosomal dominant macular dystrophy. *Nat Genet*. 27 : 89 - 93
- 33. Suzu S, Hayashi Y, Harumi T, Nomaguchi K, Yamada M, Hayasawa H, Motoyoshi K. 2002; Molecular cloning of a novel immunoglobulin superfamily gene preferentially expressed by brain and testis. *Biochem Biophys Res Commun*. 296 : 1215 - 1221
- 34. Katoh M. 2003; IGSF11 gene, frequently up-regulated in intestinal-type gastric cancer, encodes adhesion molecule homologous to CXADR, FLJ22415 and ESAM. *Int J Oncol*. 23 : 525 - 531
- 35. Watanabe T, Suda T, Tsunoda T, Uchida N, Ura K, Kato T, Hasegawa S, Satoh S, Ohgi S, Tahara H, Furukawa Y, Nakamura Y. 2005; Identification of immunoglobulin superfamily 11 (IGSF11) as a novel target for cancer immunotherapy of gastrointestinal and hepatocellular carcinomas. *Cancer Sci*. 96 : 498 - 506
- 36. Horie M, Mitsumoto Y, Kyushiki H, Kanemoto N, Watanabe A, Taniguchi Y, Nishino N, Okamoto T, Kondo M, Mori T, Noguchi K, Nakamura Y, Takahashi E, Tanigami A. 2000; Identification and characterization of TMEFF2, a novel survival factor for hippocampal and mesencephalic neurons. *Genomics*. 67 : 146 - 152
- 37. Gery S, Yin D, Xie D, Black KL, Koefler HP. 2003; TMEFF1 and brain tumors. *Oncogene*. 22 : 2723 - 2727
- 38. Uchida T, Wada K, Akamatsu T, Yonezawa M, Noguchi H, Mizoguchi A, Kasuga M, Sakamoto C. 1999; A novel epidermal growth factor-like molecule containing two follistatin modules stimulates tyrosine phosphorylation of erbB-4 in MKN28 gastric cancer cells. *Biochem Biophys Res Commun*. 266 : 593 - 602
- 39. Mahtouk K, Cremer FW, Reme T, Jourdan M, Baudard M, Moreaux J, Requirand G, Fiol G, De Vos J, Moos M, Quittet P, Goldschmidt H, Rossi JF, Hose D, Klein B. 2006; Heparan sulphate proteoglycans are essential for the myeloma cell growth activity of EGF-family ligands in multiple myeloma. *Oncogene*. 25 : 7180 - 7191
- 40. Young J, Biden KG, Simms LA, Huggard P, Karamatic R, Eyre HJ, Sutherland GR, Herath N, Barker M, Anderson GJ, Fitzpatrick DR, Ramm GA, Jass JR, Leggett BA. 2001; HPP1: a transmembrane protein-encoding gene commonly methylated in colorectal polyps and cancers. *Proc Natl Acad Sci U S A*. 98 : 265 - 270
- 41. Gery S, Sawyers CL, Agus DB, Said JW, Koefler HP. 2002; TMEFF2 is an androgen-regulated gene exhibiting antiproliferative effects in prostate cancer cells. *Oncogene*. 21 : 4739 - 4746
- 42. Fodero-Tavoletti MT, Hardy MP, Cornell B, Katsis F, Sadek CM, Mitchell CA, Kemp BE, Tiganis T. 2005; Protein tyrosine phosphatase hPTPN20a is targeted to sites of actin polymerization. *Biochem J*. 389 : 343 - 354
- 43. Gez J, Cloosterman D, Partington M. 2006; ARX: a gene for all seasons. *Curr Opin Genet Dev*. 16 : 308 - 316
- 44. Stromme P, Mangelsdorf ME, Shaw MA, Lower KM, Lewis SM, Bruyere H, Lutchera V, Gedeon AK, Wallace RH, Scheffer IE, Turner G, Partington M, Frints SG, Fryns JP, Sutherland GR, Mulley JC, Gez J. 2002; Mutations in the human ortholog of Aristaless cause X-linked mental retardation and epilepsy. *Nat Genet*. 30 : 441 - 445
- 45. Kitamura K, Yanazawa M, Sugiyama N, Miura H, Iizuka-Kogo A, Kusaka M, Omichi K, Suzuki R, Kato-Fukui Y, Kamiirisa K, Matsuo M, Kamijo S, Kasahara M, Yoshioka H, Ogata T, Fukuda T, Kondo I, Kato M, Dobyns WB, Yokoyama M, Morohashi K. 2002; Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nat Genet*. 32 : 359 - 369
- 46. Tarnasky H, Gill D, Murthy S, Shao X, Demetrick DJ, van der Hoorn FA. 1998; A novel testis-specific gene, SPAG4, whose product interacts specifically with outer dense fiber protein ODF27, maps to human chromosome 20q11.2. *Cytogenet Cell Genet*. 81 : 65 - 67
- 47. Kennedy C, Sebire K, de Kretser DM, O'Bryan MK. 2004; Human sperm associated antigen 4 (SPAG4) is a potential cancer marker. *Cell Tissue Res*. 315 : 279 - 283
- 48. Starr TK, Jameson SC, Hogquist KA. 2003; Positive and negative selection of T cells. *Annu Rev Immunol*. 21 : 139 - 176
- 49. Palmer E. 2003; Negative selection--clearing out the bad apples from the T-cell repertoire. *Nat Rev Immunol*. 3 : 383 - 391
- 50. Zehn D, Bevan MJ. 2006; T cells with low avidity for a tissue-restricted antigen routinely evade central and peripheral tolerance and cause autoimmunity. *Immunity*. 25 : 261 - 270
- 51. Galea I, Bechmann I, Perry VH. 2007; What is immune privilege (not)? . *Trends Immunol*. 28 : 12 - 18
- 52. Willimsky G, Blankenstein T. 2005; Sporadic immunogenic tumours avoid destruction by inducing T-cell tolerance. *Nature*. 437 : 141 - 146
- 53. Parker KC, Bednarek MA, Coligan JE. 1994; Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol*. 152 : 163 - 175
- 54. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. 1999; SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*. 50 : 213 - 219
- 55. Bredel M, Bredel C, Juric D, Harsh GR, Vogel H, Recht LD, Sikic BI. 2005; Functional network analysis reveals extended gliomagenesis pathway maps and three novel MYC-interacting genes in human gliomas. *Cancer Res*. 65 : 8679 - 8689
- 56. Ginos MA, Page GP, Michalowicz BS, Patel KJ, Volker SE, Pambuccian SE, Ondrey FG, Adams GL, Gaffney PM. 2004; Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. *Cancer Res*. 64 : 55 - 63
- 57. Talantov D, Mazumder A, Yu JX, Briggs T, Jiang Y, Backus J, Atkins D, Wang Y. 2005; Novel genes associated with malignant melanoma but not benign melanocytic lesions. *Clin Cancer Res*. 11 : 7234 - 7242
- 58. Zhan F, Hardin J, Kordsmeier B, Bumm K, Zheng M, Tian E, Sanderson R, Yang Y, Wilson C, Zangari M, Anaissie E, Morris C, Muwalla F, van Rhee F, Fassas A, Crowley J, Tricot G, Barlogie B, Shaughnessy J Jr. 2002; Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood*. 99 : 1745 - 1757
- 59. Lancaster JM, Dressman HK, Whitaker RS, Havrilesky L, Gray J, Marks JR, Nevins JR, Berchuck A. 2004; Gene expression patterns that characterize advanced stage serous ovarian cancers. *J Soc Gynecol Investig*. 11 : 51 - 59
- 60. Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, Lai KM, Ji J, Dudoit S, Ng IO, Van De Rijn M, Botstein D, Brown PO. 2002; Gene expression patterns in human liver cancers. *Mol Biol Cell*. 13 : 1929 - 1939
- 61. Richardson AL, Wang ZC, De Nicolo A, Lu X, Brown M, Miron A, Liao X, Iglehart JD, Livingston DM, Ganesan S. 2006; X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell*. 9 : 121 - 132
- 62. Dyrskjot L, Kruhoffer M, Thykjaer T, Marcussen N, Jensen JL, Moller K, Orntoft TF. 2004; Gene expression in the urinary bladder: a common carcinoma in situ gene expression signature exists disregarding histopathological classification. *Cancer Res*. 64 : 4040 - 4048

- 63 . Andersson A , Ritz C , Lindgren D , Eden P , Lassen C , Heldrup J , Olofsson T , Rade J , Fontes M , Porwit-Macdonald A , Behrendtz M , Hoglund M , Johansson B , Fioretos T . 2007 ; Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status . *Leukemia* . 21 : 1198 - 1203
- 64 . Korkola JE , Houldsworth J , Chadalavada RS , Olshen AB , Dobrzynski D , Reuter VE , Bosl GJ , Chaganti RS . 2006 ; Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors . *Cancer Res* . 66 : 820 - 827
- 65 . Sun L , Hui AM , Su Q , Vortmeyer A , Kotliarov Y , Pastorino S , Passaniti A , Menon J , Walling J , Bailey R , Rosenblum M , Mikkelsen T , Fine HA . 2006 ; Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain . *Cancer Cell* . 9 : 287 - 300
- 66 . Vanaja DK , Chevillet JC , Iturria SJ , Young CY . 2003 ; Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression . *Cancer Res* . 63 : 3877 - 3882
- 67 . Tomlins SA , Mehra R , Rhodes DR , Cao X , Wang L , Dhanasekaran SM , Kalyana-Sundaram S , Wei JT , Rubin MA , Pienta KJ , Shah RB , Chinnaiyan AM . 2007 ; Integrative molecular concept modeling of prostate cancer progression . *Nat Genet* . 39 : 41 - 51
- 68 . Buchholz M , Braun M , Heidenblut A , Kestler HA , Kloppel G , Schmiegel W , Hahn SA , Luttes J , Gress TM . 2005 ; Transcriptome analysis of microdissected pancreatic intraepithelial neoplastic lesions . *Oncogene* . 24 : 6626 - 6636
- 69 . Zou TT , Selaru FM , Xu Y , Shustova V , Yin J , Mori Y , Shibata D , Sato F , Wang S , Oлару A , Deacu E , Liu TC , Abraham JM , Meltzer SJ . 2002 ; Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon . *Oncogene* . 21 : 4855 - 4862
- 70 . Garber ME , Troyanskaya OG , Schluens K , Petersen S , Thaesler Z , Pacyna-Gengelbach M , van de Rijn M , Rosen GD , Perou CM , Whyte RI , Altman RB , Brown PO , Botstein D , Petersen I . 2001 ; Diversity of gene expression in adenocarcinoma of the lung . *Proc Natl Acad Sci U S A* . 98 : 13784 - 13789
- 71 . Graudens E , Boulanger V , Mollard C , Mariage-Samson R , Barlet X , Gremy G , Couillault C , Lajemi M , Piatier-Tonneau D , Zaborski P , Eveno E , Auffray C , Imbeaud S . 2006 ; Deciphering cellular states of innate tumor drug responses . *Genome Biol* . 7 : R19 -
- 72 . Sanchez-Carbayo M , Socci ND , Lozano J , Saint F , Cordon-Cardo C . 2006 ; Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays . *J Clin Oncol* . 24 : 778 - 789
- 73 . Lu KH , Patterson AP , Wang L , Marquez RT , Atkinson EN , Baggerly KA , Ramoth LR , Rosen DG , Liu J , Hellstrom I , Smith D , Hartmann L , Fishman D , Berchuck A , Schmandt R , Whitaker R , Gershenson DM , Mills GB , Bast RC Jr . 2004 ; Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis . *Clin Cancer Res* . 10 : 3291 - 3300
- 74 . Gumz ML , Zou H , Kreinest PA , Childs AC , Belmonte LS , LeGrand SN , Wu KJ , Luxon BA , Sinha M , Parker AS , Sun LZ , Ahlquist DA , Wood CG , Copland JA . 2007 ; Secreted frizzled-related protein 1 loss contributes to tumor phenotype of clear cell renal cell carcinoma . *Clin Cancer Res* . 13 : 4740 - 4749

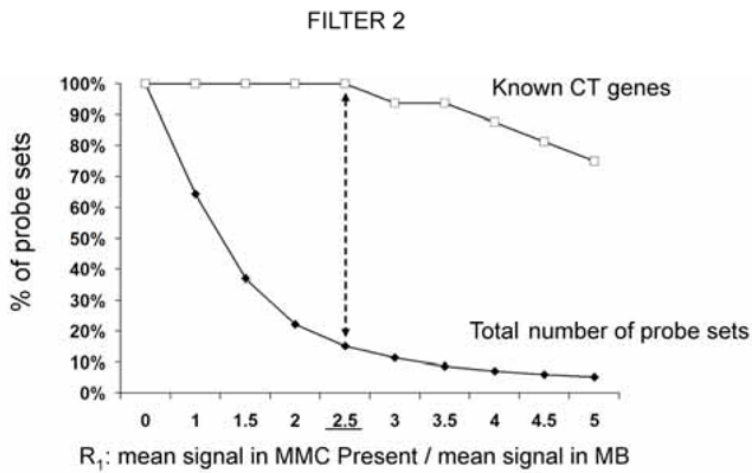
**Figure 1**

A. Filters for selecting based on the Affymetrix signal and detection call PS: probe sets; NT: normal tissues; MMC: primary myeloma cells; MBC: memory B cells; P: Present call; A: Absent call B. Cumulative number of probe sets according to an increasing ratio reflecting the overexpression in MMC compared to MBC cells. For each probe set (PS), the ratio (R1) of the mean Affymetrix signal of MMC samples having a Present call to the mean Affymetrix signal of 7 MBC samples was calculated. The curves represent the percentage of PS for which R1 is superior or equal to the indicated value, starting from 16982 PS selected by filter 1 (full squares) including 16 known CT genes (empty squares). C. Cumulative numbers of probe sets according to an increasing ratio reflecting the overexpression in MMC samples with a Present call compared to MMC samples with an Absent call. For each probe set (PS) with a frequency of expression  $\leq 80\%$  among MMC from 64 patients, the ratio (R2) of the mean Affymetrix signal of MMC samples having a Present call to the mean Affymetrix signal of MMC samples having an Absent call was calculated. Curves represent the percentage of PS for which R2 is superior or equal to the indicated value, starting from 315 PS with a frequency of expression  $\leq 80\%$  out of 341 PS selected by filter 3 (full squares) including 15 known CT genes (empty squares).

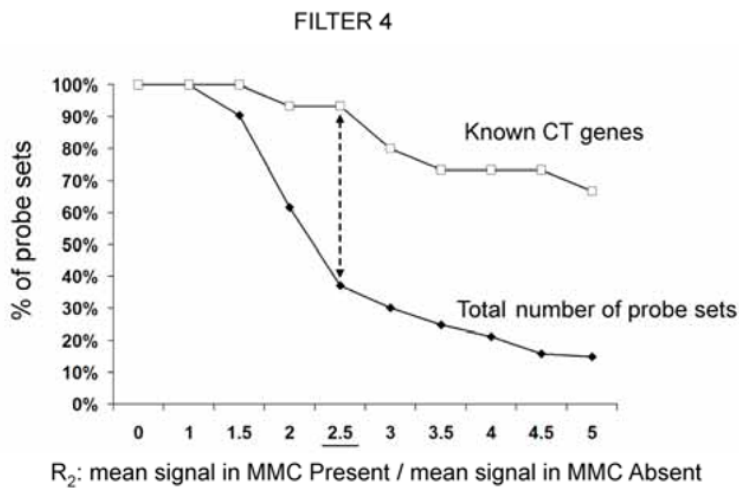
**A**



**B**



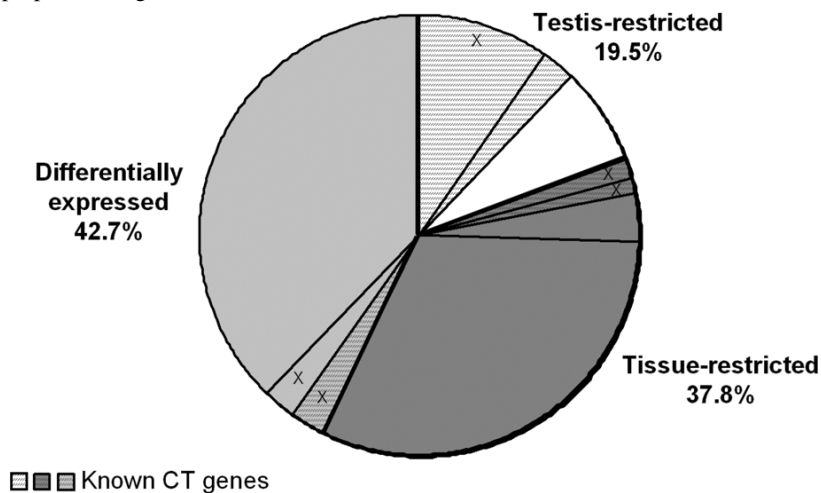
**C**



**Figure 2**

Distribution of the 82 CT genes in the 3 CT gene categories

White background is for the “testis-restricted” category (genes expressed in testis only), dark grey is for the “tissue-restricted” category (genes expressed in testis and fewer than two non-gametogenic tissues) and light grey is for the “differentially expressed” category (genes expressed in testis and in 3–6 non-gametogenic tissues). In each category, hatched parts represent known CT genes. Parts including “X” are the proportion of genes located on chromosome X.



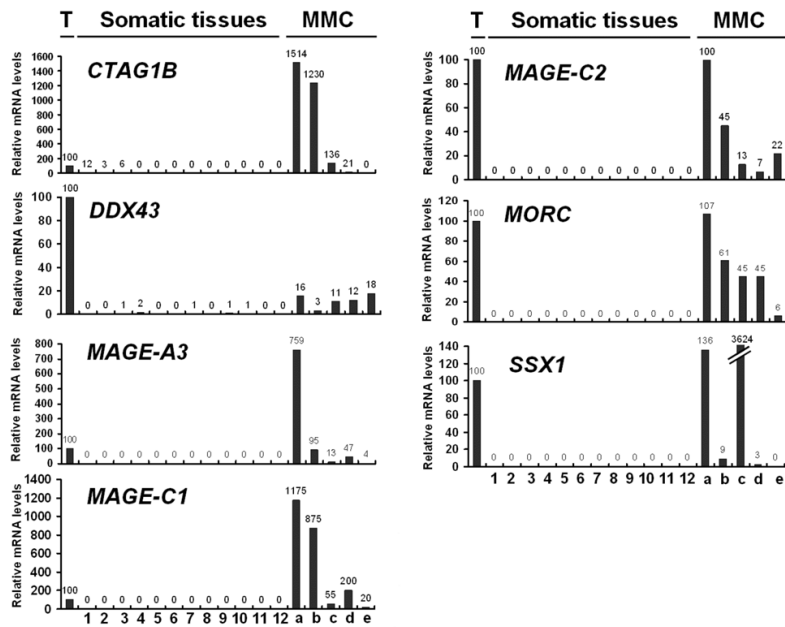
**Figure 3**

Expression patterns of potential novel CT genes determined by real-time RT-PCR

The relative gene mRNA level was determined as described in Materials and Methods. A value of '100' was assigned to the testis sample (T). Expression in testis, 12 somatic tissues and in 5 MMC samples was tested for 7 known CT genes (A), and 16 novel CT genes resulting in 1 testis-restricted (TEX14) and 8 tissue-restricted genes (B) and 7 differentially expressed genes (C). Lane assignments are as follows: 1, brain; 2, liver; 3, stomach; 4, pancreas; 5, kidney; 6, intestine; 7, bone marrow; 8, heart; 9, lung; 10, blood (leucocytes); 11, adrenal gland; 12, skeletal muscle. For each genes, a, b, c are MMC samples showing a Present call with a high Affymetrix signal (above the median); d and e are MMC samples showing a Present call with a low Affymetrix signal (under the median).

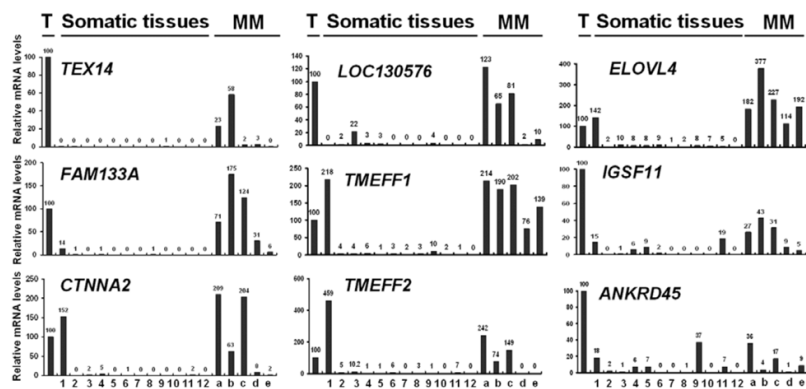
**A**

**Known CT genes**



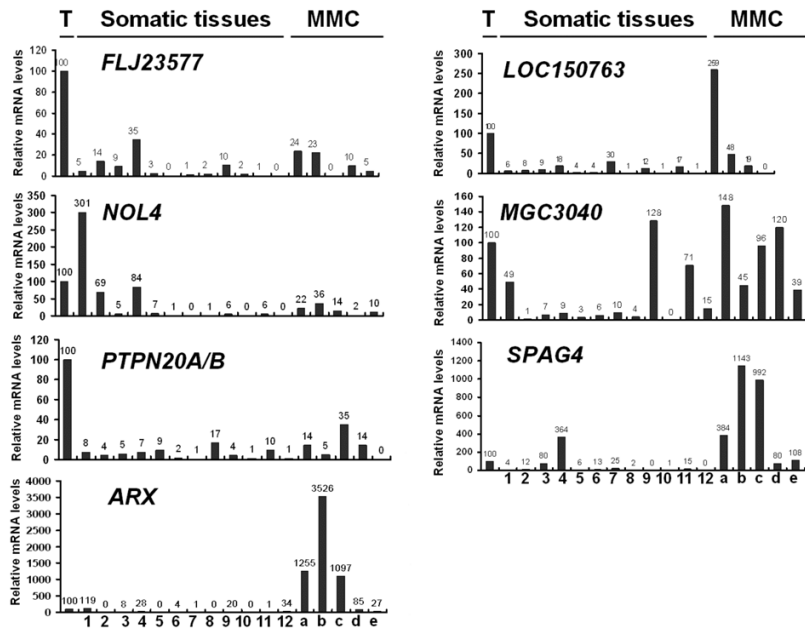
**B**

« Testis- and tissue-restricted » novel CT genes



C

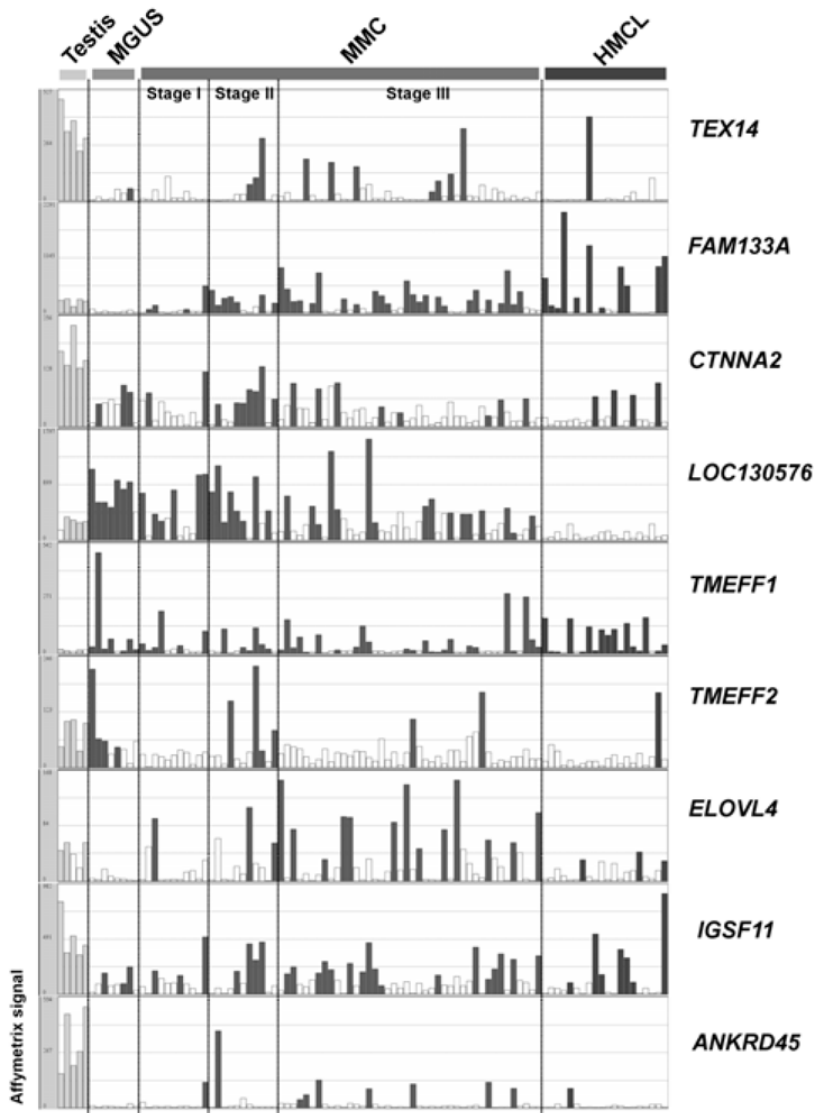
« Differentially expressed » novel CT genes



**Figure 4**

Gene expression profiles of 9 novel CT genes in MGUS and MM samples measured by Affymetrix HG-U133 Set arrays

Histograms show the expression level of 9 tissue-restricted CT genes in five testis samples, eight purified plasma cell samples from patients with monoclonal gammopathy with undetermined significance (MGUS), 64 MMC samples from patients with multiple myeloma (MMC) ordered in stages (I, II, III) and 20 human myeloma cell lines (HMCL) samples. The signal intensity for each gene is shown on the Y axis as arbitrary units determined by the GCOS 1.2 software (Affymetrix). Empty histograms indicate an “Absent” Affymetrix call and filled histograms a “Present” Affymetrix call.



**Table I**

Known and novel CT gene expression patterns determined by real-time PCR.

CT gene category	Taqman assay	gene Gene name	Chromosomal number of positive <sup>1</sup>			R a t i o		
			Location	somatic tissues	Positive somatic tissues	MMC/Testis <sup>2</sup> (qPCR)	Ratio MMC/NT Max <sup>3</sup> (qPCR)	Frequency in MM patients (n=64) <sup>4</sup> (Affymetrix)
Testis- restricted	Hs00218682_m1	<b>DDX43</b>	chr6q12-q13	0		0.1	5.8	34%
	Hs00366532_m1	<b>MAGEA3</b>	chrXq28	0		2.9	28915.2	33%
	Hs00193821_m1	<b>MAGEC1</b>	chrXq26	0		7.0	70208.3	66%
	Hs00212255_m1	<b>MAGEC2</b>	chrXq27	0		0.5	460.7	13%
	Hs00205327_m1	<b>MORC1</b>	chr3q13	0		0.7	204.9	36%
	Hs00846692_s1	<b>SSX1</b>	chrXp11	0		15.4	153600.7	20%
	Hs00258708_m1	TEX14	chr17q22	0		0.3	41.1	16%
Tissue- restricted	Hs00265824_m1	<b>CTAG1B</b>	chrXq28	1	Brain	9.6	78.1	13%
	Hs00189285_m1	CTNNA2	chr2p12-p11.1	1	Brain	1.6	1.0	27%
	Hs00396095_m1	LOC130576	chr2q23.2	1	Stomach	1.0	4.4	47%
	Hs01395126_m1	FAM133A	chrXq21.32	1	Brain	0.7	5.1	53%
	Hs01651017_m1	ANKRD45	chr1q25.1	2	Lung, brain	0.2	0.5	16%
	Hs00224122_m1	ELOVL4	chr6q14	2	Brain, stomach	2.6	1.8	25%
	Hs00541322_m1	IGSF11	chr3q13.32	2	Brain, adrenal gland	0.3	1.8	39%
	Hs00186495_m1	TMEFF1	chr9q31	2	Brain, lung	2.0	0.9	20%
	Hs00249367_m1	TMEFF2	chr2q32.3	2	Brain, stomach	1.6	0.3	11%
Differentially expressed Ubiquitously expressed	Hs00332514_m1	FLJ23577	chr5p13.2	3	Pancreas, liver, lung	0.2	0.4	47%
	Hs01096927_m1	NOL4	chr18q12	3	Brain, pancreas	0.2	0.1	59%
	Hs00417254_m1	PTPN20A	chr10q11.22	3	Heart, adrenal gland, kidney	0.2	1.1	25%
	Hs00292465_m1	ARX	chrXp22.1-3	4	Brain, pancreas, lung, skeletal muscle	19.6	16.5	17%
	Hs00418637_g1	LOC150763	chr2q11.2	4	Pancreas, bone marrow, lung, adrenal gland	1.1	3.6	27%
	Hs00382529_m1	MGC3040	chr3q21	5	Lung, brain, bone marrow, adrenal gland, skeletal muscle	1.0	0.8	31%
	Hs00162127_m1	SPAG4	chr20q11.21	6	Pancreas, stomach, liver, intestine, bone marrow, adrenal gland	8.4	2.3	100%
Ubiquitously expressed	Hs00231877_m1	ETV1	chr7p21.3	7	Lung, adrenal gland, liver, stomach, pancreas, heart, brain	0.2	0.1	45%
	Hs00214398_m1	FLJ20130	chrXq22.3	7	Adrenal gland, liver, pancreas, kidney, intestine, bone marrow, lung	1.5	2.4	41%
	Hs00177193_m1	PTPRG	chr3p21-p14	9	Lung, brain, stomach, liver, pancreas, kidney, intestine, heart, adrenal gland	2.9	0.9	80%

Gene names written in Bold are known CT genes.

<sup>1</sup> A gene was stated expressed by a tissue when the real-time RT-PCR value was  $\geq 10$ ;

<sup>2</sup> The mean value of the three "high" MMC was divided by the testis value;

<sup>3</sup> The ratio represents the mean value of the three "high" MMC to the value of the normal tissue (NT) with the highest expression;

<sup>4</sup> Frequencies are the percentages of MMC samples with a Present call among 64 patients.

**Table 2**

Expression of the novel CT genes in other cancers determined with microarray analysis and extracted with Oncomine database.

CT gene	Cancer vs normal	p-value	reference
TEX14	Oligodendroglioma	0.01	(53)
	Head neck squamous cell carcinoma	0.004	(54)
CTNNA2	Melanoma	2.2E-14	(55)
	Multiple myeloma (other)	1.10E-09	(56)
	Ovarian adenocarcinoma	8.50E-05	(57)
FAM133A	Hepatocellular carcinoma	2.70E-05	(58)
ANKRD45	Breast carcinoma	0.003	(59)
ELOVL4	Bladder	2.40E-05	(60)
	T-cell acute lymphoblastic leukemia	6.5E-4	(61)
TMEFF1	Bladder	2.3E-4	(60)
	Breast carcinoma	8.7E-6	(59)
	Head neck squamous cell carcinoma	1.50E-04	(54)
	B-cell acute lymphoblastic leukemia	2.90E-04	(61)
	Melanoma	7.60E-08	(55)
	Seminoma	3.30E-06	(62)
TMEFF2	Glioblastoma multiforme	8.30E-06	(63)
	Prostate adenocarcinoma	3.00E-04	(64)
	Breast carcinoma	0.005	(59)
	Acute myeloid leukemia	0.007	(61)
	Seminoma	7.60E-04	(62)
FLJ23577	Breast carcinoma	0.002	(59)
	Prostate carcinoma	0.002	(65)

Novel cancer-testis genes in multiple myeloma

	Pancreatic ductal adenocarcinoma	0.004	(66)
NOL4	Breast carcinoma	0.001	(59)
	Colon carcinoma	0.006	(67)
	B-cell acute lymphoblastic leukemia	1.30E-09	(61)
	Small cell lung cancer	0.002	(68)
	Prostate adenocarcinoma	0.004	(64)
PTPN20A1B	Bladder	7.50E-05	(60)
ARX	Breast carcinoma	0.009	(59)
	Seminoma	6.30E-07	(62)
MGC3040	Oligodendroglioma	1.10E-06	(63)
	Colorectal carcinoma	5.7E-5	(69)
SPAG4	Bladder carcinoma	2.80E-11	(70)
	Glioblastoma multiforme	0.003	(63)
	Head neck squamous cell carcinoma	1.90E-04	(54)
	Melanoma	9.90E-06	(55)
	Serous ovarian carcinoma	0.004	(71)
	Clear cell renal cell carcinoma	2.20E-10	(72)