Oncogene (2005) 24, 2461–2473
© 2005 Nature Publishing Group All rights reserved 0950-9232/05 $30.00
www.nature.com/onc

Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma

Michela Mattioli[1,5], Luca Agnelli[1,5], Sonia Fabris[1], Luca Baldini[1], Fortunato Morabito[2], Silvio Bicciato[3], Donata Verdelli[1], Daniela Intini[1], Lucia Nobili[1], Lilla Cro[1], Giancarlo Pruneri[4], Vincenzo Callea[2], Caterina Stelitano[2], Anna Teresa Maiolo[1], Luigia Lombardi[1] and Antonino Neri*[,1]

[1]*Laboratorio di Ematologia Sperimentale e Genetica Molecolare and U.O. Ematologia 1, Dipartimento di Scienze Mediche, Università degli Studi di Milano, Ospedale Maggiore IRCCS, Milano, Italy;* [2]*Divisione di Ematologia and Centro Trapianto di Midollo, Azienda Ospedaliera 'Bianchi-Melacrinò-Morelli', Reggio Calabria, Italy;* [3]*Dipartimento dei Processi Chimici dell'Ingegneria, Università di Padova, Padova, Italy;* [4]*Divisione di Patologia, Istituto Europeo di Oncologia, Milano, Italy*

Multiple myeloma (MM) is the most common form of plasma cell dyscrasia, characterized by a marked heterogeneity of genetic lesions and clinical course. It may develop from a premalignant condition (monoclonal gammopathy of undetermined significance, MGUS) or progress from intramedullary to extramedullary forms (plasma cell leukemia, PCL). To provide insights into the molecular characterization of plasma cell dyscrasias and to investigate the contribution of specific genetic lesions to the biological and clinical heterogeneity of MM, we analysed the gene expression profiles of plasma cells isolated from seven MGUS, 39 MM and six PCL patients by means of DNA microarrays. MMs resulted highly heterogeneous at transcriptional level, whereas the differential expression of genes mainly involved in DNA metabolism and proliferation distinguished MGUS from PCLs and the majority of MM cases. The clustering of MM patients was mainly driven by the presence of the most recurrent translocations involving the immunoglobulin heavy-chain locus. Distinct gene expression patterns have been found to be associated with different lesions: the overexpression of *CCND2* and genes involved in cell adhesion pathways was observed in cases with deregulated *MAF* and *MAFB*, whereas genes upregulated in cases with the t(4;14) showed apoptosis-related functions. The peculiar finding in patients with the t(11;14) was the downregulation of the α-subunit of the IL-6 receptor. In addition, we identified a set of cancer germline antigens specifically expressed in a subgroup of MM patients characterized by an aggressive clinical evolution, a finding that could have implications for patient classification and immunotherapy.
Oncogene (2005) **24,** 2461–2473. doi:10.1038/sj.onc.1208447
Published online 7 February 2005

*Correspondence: A Neri, U.O. Ematologia 2, Centro 'G.Marcora', Dipartimento Scienze Mediche, Università degli Studi di Milano, Ospedale Maggiore di Milano, IRCCS, Via Francesco Sforza 35, 20122 Milan, Italy; E-mail: neri.a@policlinico.mi.it
[5]These two authors contributed equally to this work
Received 28 September 2004; revised 26 November 2004; accepted 14 December 2004; published online 7 February 2005

Keywords: gene expression profiling; IGH translocations; MGUS; multiple myeloma; plasma cell leukemia

Introduction

Multiple myeloma (MM) is a neoplastic disorder of plasma cells (PCs) characterized by a highly heterogeneous clinical course; it can be preceded by the presumably premalignant condition called monoclonal gammopathy of undetermined significance (MGUS), and may progress from a truly overt intramedullary form to extramedullary myeloma (plasma cell leukemia, PCL) (Anderson *et al*., 2002; Barille-Nion *et al*., 2003).

MM shows a marked genomic instability involving both numerical and structural chromosomal aberrations making it difficult to identify the molecular mechanisms leading to distinct disease entities, prognostic risk groups and new targeted therapies. Chromosomal translocations involving the immunoglobulin heavy-chain (IGH) locus at 14q32 are a genetic hallmark of MM: their prevalence varies with the clinical stage of the disease, being detected in about 50% of MGUS patients, 60–70% of MM patients and in more than 80% of PCL patients (Ho *et al*., 2002; Seidl *et al*., 2003). Unlike in the case of other hematological malignancies, multiple loci have been described as being involved in IGH translocations (Bergsagel and Kuehl, 2001). The most recurrent are 11q13 (Chesi *et al*., 1996; Ronchetti *et al*., 1999), 4p16.3 (Chesi *et al*., 1997, 1998b; Richelda *et al*., 1997), 16q23 (Chesi *et al*., 1998a), 6p21 (Shaughnessy *et al*., 2001) and 20q12 (Hanamura *et al*., 2001), which, respectively, lead to the deregulation of *CCND1*, *FGFR3* and *WHSC1(MMSET)*, *MAF*, *CCND3* and *MAFB* genes. Moreover, several reports indicate that the partner loci are still unknown in about 40% of the cases with IGH translocations (Seidl *et al*., 2003), thus suggesting that new deregulated target genes may be identified.

DNA microarray technology has been used in recent studies to define the gene expression profiles of normal

and malignant PCs (De Vos *et al*., 2002; Zhan *et al*., 2002, 2003; Davies *et al*., 2003). It has been shown that MM PCs can be differentiated from normal PCs on the basis of the expression of a number of genes, mainly including oncogenes, growth-related and translation machinery genes, and it has been reported that at least two different classes of MM can be identified: one more similar to MGUS and the other more similar to human myeloma cell lines (Zhan *et al*., 2002). A more recent analysis (Davies *et al*., 2003) of a limited number of cases has confirmed previous findings and shown that the differences between MGUS and MM are smaller than those between normal and clonal MGUS/MM PCs. Despite these recent data, a number of biological and clinical issues remain to be clarified in relation to the pathogenesis of MM, particularly the molecular events leading to the transition from MGUS to MM, the contribution of the activated IGH-translocated genes to the biological and clinical variability of MM and the molecular pathways deregulated in patients without IGH translocations.

To contribute to a better understanding of the changes in gene expression associated with the neoplastic transformation of PCs in MM, we analysed highly purified PCs of patients affected by different forms of plasma cell dyscrasias (MGUS, MM and PCL) using high-density oligonucleotide microarrays representative of $\sim$18 000 transcripts. The gene expression profiles were investigated for correlations with clinical parameters and distinct types of chromosomal lesions.

Results

Gene expression profiling differentiates MGUS from MM and PCL PCs

To determine whether gene expression profiling of PCs from MGUS, MM and PCL patients could identify the distinct clinical entities, we performed an unsupervised analysis using the hierarchical clustering algorithm (Eisen *et al*., 1998). The 52 samples, described by 2249 highly variable genes, generated a dendrogram (Figure 1) with two major branches: one branch (*green*) contained all the MGUS and five of the 39 MM cases, whereas the second branch grouped the PCL samples and most of the MM. Neither the PCL nor MM samples could be identified as a distinct cluster of the dendrogram, although 4/6 PCL cases were grouped together. The MM samples did not cluster according to their clinical–pathological characteristics, such as clinical stage, the occurrence of bone lesions or β2-microglobulin levels, but were mainly grouped on the basis of the presence of IGH translocations (see below).

In order to identify the genes specifically distinguishing the three different clinical entities, we used a supervised approach to compare the expression profiles of MGUS, MM and PCL cases. A total of 120 probesets were differentially expressed in MGUS *versus* PCL samples (80 downregulated and 40 upregulated), in addition to 56 transcripts encoding immunoglobulin molecules upregulated in MGUS samples (data not shown). Figure 2a shows the matrix representing the expression of the identified genes, grouped according to their predicted biological functions. The main difference involved the expression of genes related to DNA replication and metabolism. A number of proliferation-associated genes were upregulated in the PCL samples, including genes controlling DNA synthesis (e.g. *NUSAP1*, *CAD*, *PCNA*, *PAICS*) and genes involved in DNA mismatch repair pathway (*MCM2, MCM4*, *MCM5*, *MCM6*, *MSH6*). Furthermore, the PCL PCs showed a selective upregulation of several genes encoding molecules involved in RNA metabolism (*EXOSC4*, *RBM10*), protein metabolism (*PMSC4*, *CCT4*, *CCT3*) and nuclear–cytoplasmic transport (*NUP62*, *NUP210*, *RAN*). Consistently with these results, the genes whose overexpression has been linked to growth arrest, such as the negative cell-cycle regulator *GADD45B* (Jin *et al*., 2002), were upregulated in the MGUS samples. Interestingly, among the genes differentially expressed in MGUS *versus* PCL PCs, several encode for MYC molecular targets such as *CAD*, *PRDX3* and *GADD45B* (Fernandez *et al*., 2003).

The comparison of MGUS *versus* PCL PCs also showed changes in the expression of genes codifying for transcriptional repressors or cofactors that play critical roles in chromatin remodeling, such as ENO1, SAP30 and SMARCA4. This functional category also included the *WHSC1* gene, targeted by t(4;14) (Chesi *et al*., 1997; Richelda *et al*., 1997), which was the most statistically significant selected gene and showed a 15-fold induction in PCL *versus* MGUS. The upregulation of the *WHSC1* gene was confirmed by Real-time quantitative RT–PCR (Q-RT–PCR) analysis (Figure 2b), which showed that the PCL samples, although negative for t(4;14) and structural alterations of the 4p16.3 region (Fabris *et al*., 2005) had higher levels of *WHSC1* transcripts than MGUS ($P=0.0318$) and MM patients negative for t(4;14) ($P<0.0001$).

When the supervised approach was applied to the comparison of MM with either PCL or MGUS PCs, we were able to identify only a few genes specifically differentiating MM from MGUS (Figure 2c). Nine of these 15 genes overlapped those identified in the comparison between MGUS and PCL (Figure 2a, genes in red), and showed progressive modulation during the transition from MGUS through MM to the PCL phenotype. In particular, the chemokine *CXCL8* (*IL8*) was the most upregulated gene in MGUS together with the *GADD45B* and the cytoskeletal organizer *VIL2* genes.

MM patients clustering is mainly driven by the presence of the most recurrent IGH chromosomal translocations

To investigate further the correlation between gene expression profiles and clinical-pathological characteristics of MM patients, we performed an unsupervised analysis on the 39 MM dataset described by the 1899

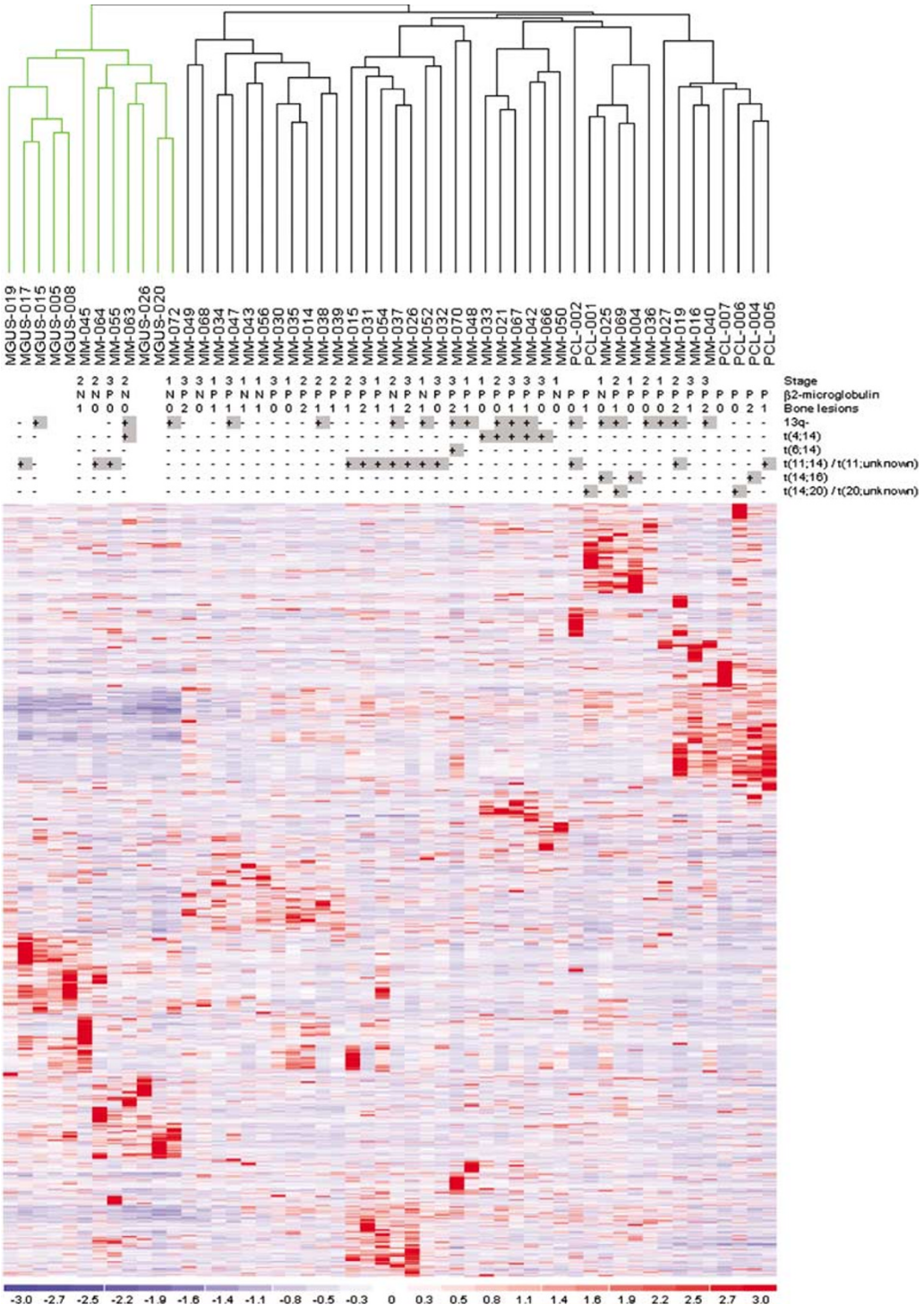

Figure 1 Unsupervised analysis of gene expression profiles from purified CD138$^+$ PC samples of seven MGUS, 39 MM and six PCL cases. The dendrogram was generated using a hierarchical clustering algorithm based on the average-linkage method. In the matrix, each column represents a sample, and each row a gene. The 52 samples are grouped according to the expression levels of the 2249 most variable genes. The color scale bar represents the relative gene expression changes normalized by the standard deviation. The clinical and molecular characteristics of the patients are shown. *Stage*: 1, stage I; 2, stage II; 3, stage III. *Serum β2-microglobulin:* N, normal values; P, pathological values ($\geqslant$2.4 mg/l). *Bone lesions*: 0, none; 1, $\leqslant$3; 2, $>$3. Patients positive for 13q deletion and for different IGH chromosomal translocations are identified with + symbol. *t(11;unknown)* and *t(20;unknown)* indicate different rearrangement, respectively targeting the 11q13 and 20q12 loci without any apparent involvement of Ig loci

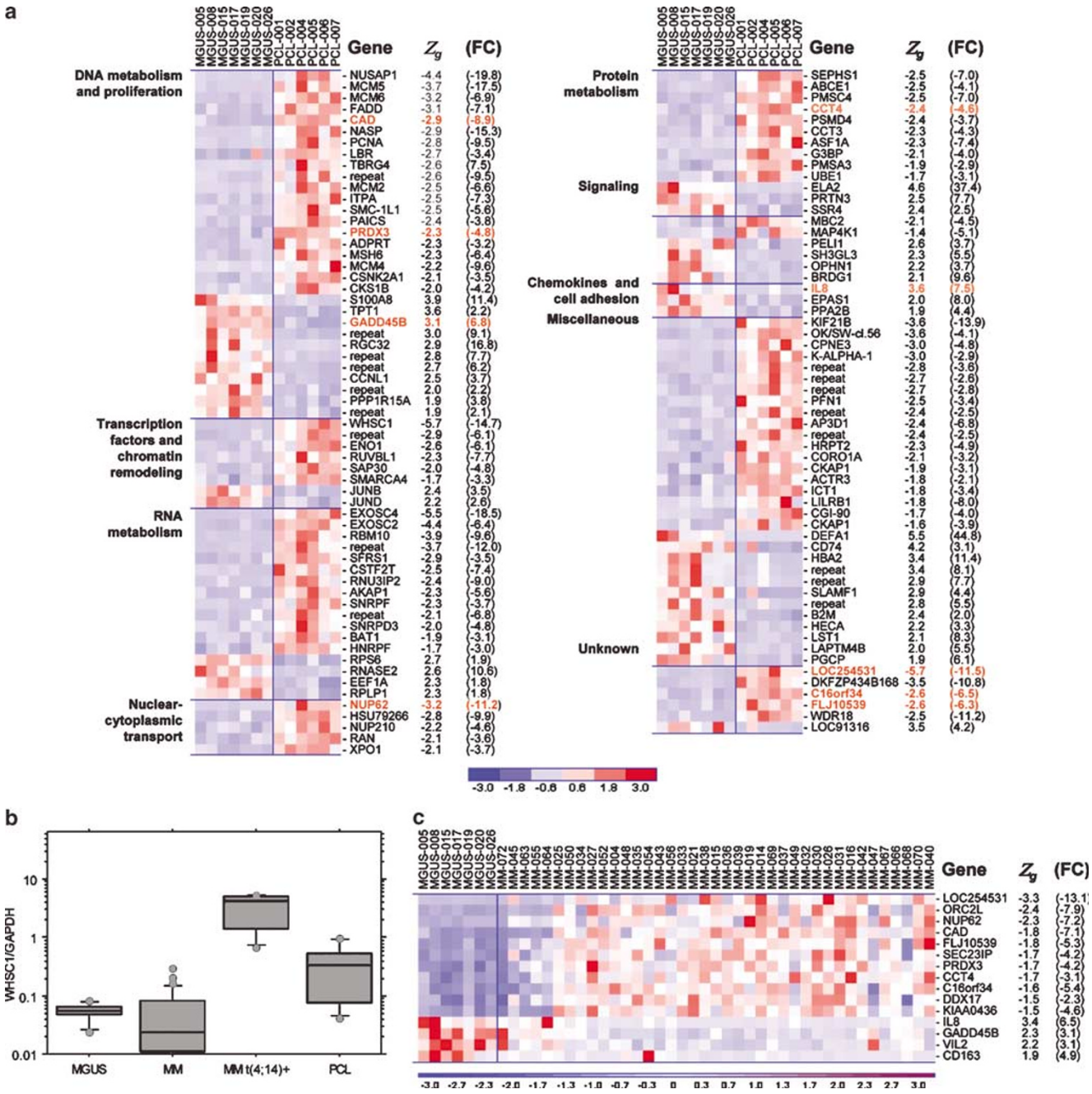

Figure 2 Identification of genes specifically expressed in MGUS, MM and PCL by supervised analysis. (**a**) The genes differentially expressed in MGUS *versus* PCL PCs are grouped according to their functional categories and ranked within each category according to their z_g score. The color changes in each row represent the gene expression relative to the mean across the samples. The genes, z_g scores and average fold changes (FC) are indicated. The z_g scores and average fold changes are expressed using the MGUS group as baseline. Some of the genes appear more than once (*repeat*). Genes overlapping with those reported in Figure 2c are in red. (**b**) Box plot of *WHSC1/GAPDH* mRNA expression levels in purified MGUS ($n=7$), MM ($n=33$), t(4;14)-positive MM ($n=6$) and PCL ($n=6$) PCs as detected by Q-RT–PCR analysis (values on the *y*-axis are $\log_{10}$ scaled). (**c**) Identification of genes distinguishing the MGUS and MM samples by supervised analysis performed on seven MGUS and 39 MM patients; the support value used for the analysis was $n=n_0-2$. For GenBank Accession numbers, see Supplementary Table 1

most highly variable genes. Figure 3 shows that the algorithm grouped the MM samples into two major clusters: one (*orange*) mainly including cases negative for the most recurrent IGH translocations (17/18) and the other (*green*) mainly grouping positive cases (18/21). The association between the clinical–prognostic variables (clinical stage, serum β2-microglobulin levels, focal bone lesions, 13q deletion) and these two defined clusters was evaluated. The only significant correlation was found between the cluster grouping the patients positive for IGH translocations and a lower incidence of bone lesions ($P=0.0271$). Interestingly, the majority of

the patients with focal bone lesions included in this cluster were characterized by t(11;14) ($P = 0.0166$). Our database does not indicate any overall correlation between the occurrence of bone lesions and *DKK1* gene expression levels, recently reported to be associated with bone disease (Tian *et al*., 2003). However, *DKK1* transcriptional levels tended to become increasingly discriminant when the IGH translocation-negative cluster was compared with either the whole IGH translocation-positive cluster (*green*, Figure 3; $P = 0.0699$) or the IGH translocation-positive cluster without the t(11;14) cases ($P = 0.0357$).

The analysis of the strength of meaningful clusters by BRB tool quantified the internal stability of the two major branches (R-index = 0.872, D-index = 2.9). Cutting the dendrogram at lower levels (i.e. selecting more than two clusters) led to the most stable subgrouping (R-index = 0.891, D-index = 2) when seven clusters were considered (Figure 3, blue line), thus further supporting the significance of the clustering associated with the presence of specific IGH translocations. Within the translocation-positive group, the samples were clustered on the basis of the translocation type, being further subdivided in two main branches: one grouping mainly t(11;14) samples, and the other containing the t(4;14) and MAF/MAFB positive cases in two distinct subgroups. The patients without any known 14q rearrangements were clustered in three subgroups (Figure 3), one of which included cases grouping with PCL (MM-027, 019, 016, 040) as determined by the unsupervised analysis of the whole samples data set (Figure 1).

In an attempt to identify gene expression patterns associated with the main IGH translocations, we compared the expression profiles of the cases characterized by distinct translocations with those of all the other MMs by means of supervised analyses (Figure 4). Seven genes were specifically upregulated in the patients carrying the t(4;14), in addition to *FGFR3* and *WHSC1* genes, targeted by the translocation (Figure 4, upper panel). The highest consistency was reached by Kruppel-like factor 4 (*KLF4*), a zinc-finger transcriptional repressor, whose constitutive expression suppresses cell proliferation by blocking G1/S progression (Chen *et al*., 2001). In addition, we found a significant upregulation of the genes codifying for the DNA-binding protein CSDA, also known to act as a negative regulator of transcription through polymerase II promoter (Coles *et al*., 1996), and for the NGFR-associated protein 1, which is involved in apoptosis induction upon nerve growth factor binding (Mukai *et al*., 2000).

In all, 19 genes (18 upregulated and one downregulated) were identified as differentially expressed in the three patients carrying translocations deregulating either *MAF* or *MAFB* genes (Figure 4, middle panel). Among the overexpressed genes, we identified *CCND2* and *ITGβ7*, which have recently been reported to be upregulated in MM as consequence of MAF deregulation (Hurt *et al*., 2004), as well as several other genes codifying proteins involved in cell adhesion processes: the chemokine receptor CX3CR1 (Imaizumi *et al*., 2004) and the chemokine CXCL12, specific ligand of the CXCR4 receptor (Burger *et al*., 1999), the PKP2 constituent of desmosomal plaque (Chen *et al*., 2002) and the toll-like receptor 4 (TLR4), a trans-membrane receptor expressed in endothelial and B-cells (Aderem and Ulevitch, 2000). In addition, we also found overexpression of the *ARK5* gene, which codifies for a protein kinase whose activation triggered by the AKT pathway induces tumor cells survival during nutrient starvation (Suzuki *et al*., 2003). Moreover, these data were confirmed by analyses of the three PCL samples carrying deregulated *MAF*/*MAFB* and five human myeloma cell lines positive for t(14;16) translocation (data not shown).

Immunohistochemical analysis of bone marrow biopsies from 10 MM cases selected on the basis of the microarray data (three positive for translocations involving *MAF* or *MAFB* genes, and seven negative for *CCND2* mRNA expression) showed that only the PCs of the patients carrying deregulated *MAF*/*MAFB* loci expressed high CCND2 protein levels (Figure 5).

Supervised analysis of cases positive for the t(11;14) was performed on all but one patient (MM-019 was excluded because characterized by 4p16.3 locus rearrangement and *WHSC1* overexpression). Only five genes with heterogeneous functions resulted specifically upregulated in patients with t(11;14) in addition to *CCND1* (Figure 4, bottom panel). The only gene consistently downregulated in *CCND1*-translocated patients codifies for the ligand-binding α-subunit of interleukin-6 receptor (IL6-R, CD126). The analysis of *IL6-R* mRNA levels in the entire MMs panel indicated heterogeneous expression, with the highest levels being reached in the patients positive for either *WHSC1* or *MAF*/*MAFB* deregulation (Figure 6a and b). The analysis of IL6-R protein in MM subgroups defined on the basis of the translocations (Figure 6b, right panel) showed that PCs from patients carrying t(11;14) had significantly lower amount of CD126 protein in comparison with the majority of other MM patients, and particularly the patients characterized by t(4;14) or *MAF* genes deregulation.

Finally, a supervised analysis aimed at identifying the genes differentially expressed by the 13q+ and 13q− cases was performed. No gene could be identified, even by relaxing analysis criteria, thus confirming the evidences coming from the unsupervised clustering.

The expression of multiple tumor-associated antigens characterizes MM cases clustering with PCL samples

To investigate the changes in gene expression that may correlate with the clinical progression of MM, we applied supervised analysis on two groups of patients identified by unsupervised analysis (Figure 1): one including five MM cases shown to cluster with MGUS (MM-045, 064, 055, 063, 072) and the other four cases clustering with PCL (MM-027, 019, 016, 040).

Cases clustering with MGUS showed downregulation of eight probe-sets (Figure 7, upper panel), including

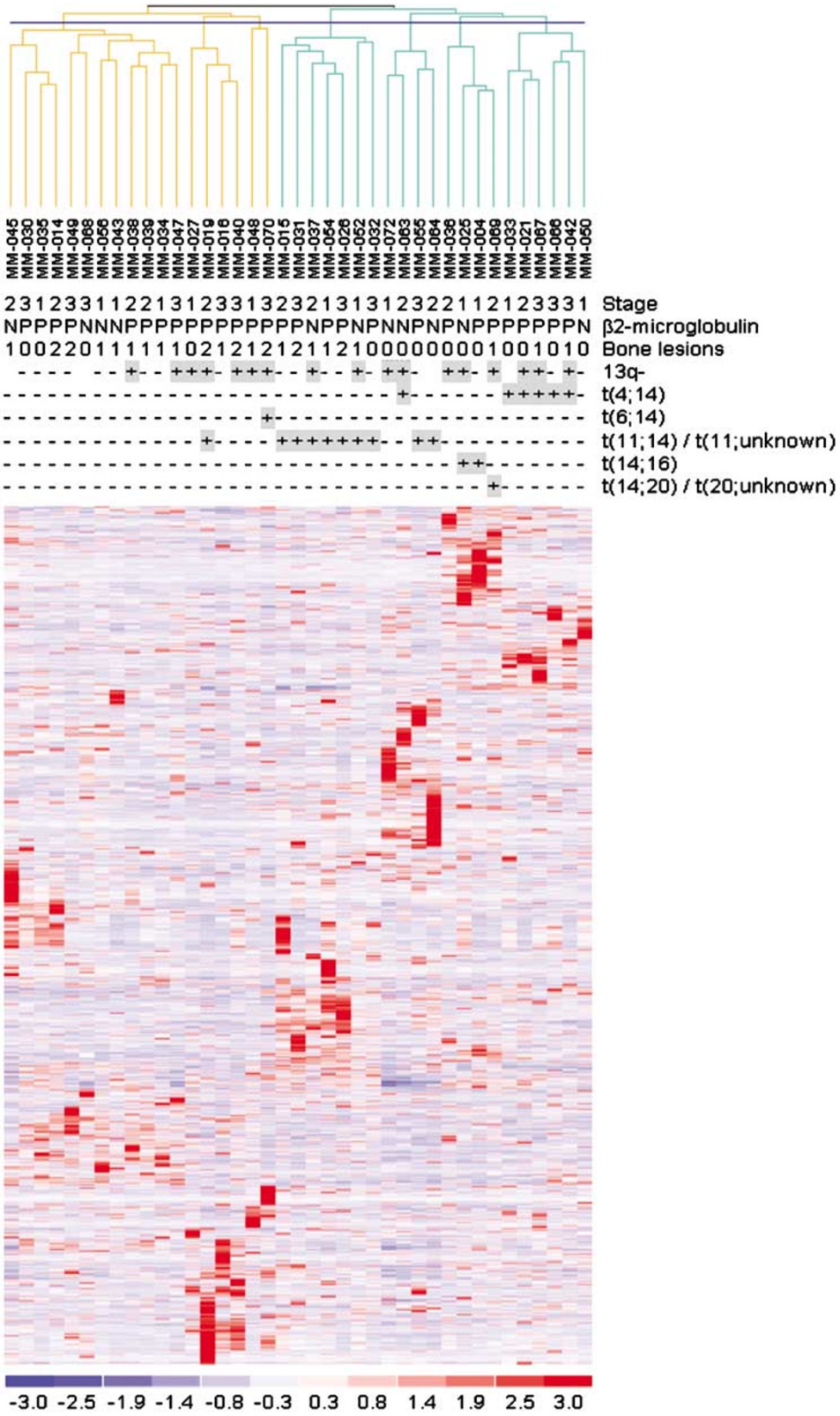

Figure 3 Unsupervised analysis of gene expression profiles from purified CD138$^+$ PC samples of 39 MM cases. The dendrogram was generated using a hierarchical clustering algorithm based on the average-linkage method. Samples are grouped according to the expression levels of the 1899 most variable genes. The blue line represents the cutting of the dendrogram at seven clusters level (see text). See Figure 1 for a description of the matrix

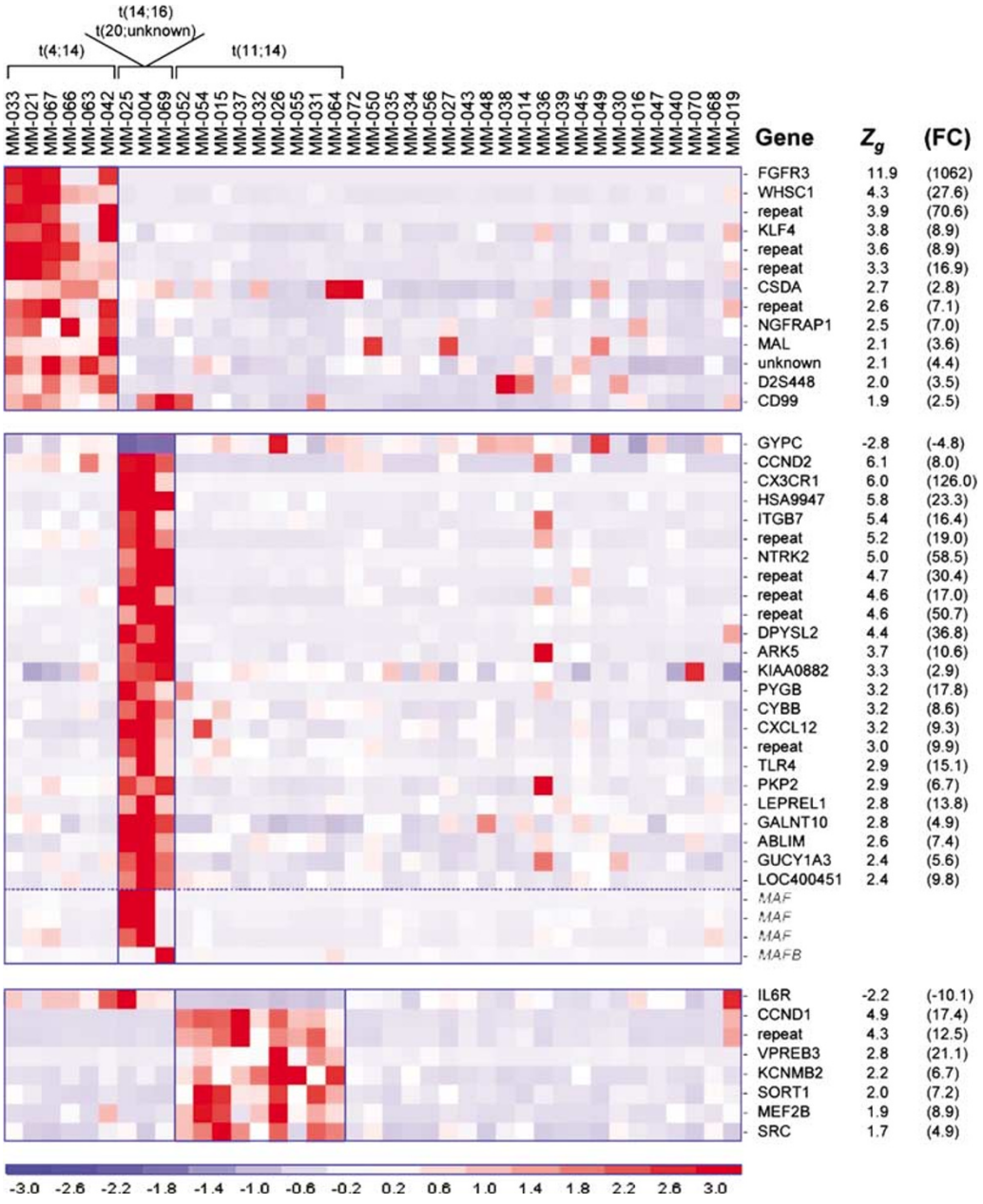

Figure 4 Identification of genes expressed in MM patients carrying specific 14q32 translocations. Supervised analyses were performed using the Gene@Work software platform to compare the MM patients positive for t(4;14) (upper panel), t(14;16)/t(20;unknown) (middle panel) or t(11;14) (lower panel) with all of the other MM patients. See Figure 2 for a description of the matrix. The support value $n = n_0 - 2$ was chosen for the t(4;14) and t(11;14) analyses; full support was used in the *MAF*/*MAFB* analysis. A δ-value of 0.05 was used in the t(11;14) analysis. *MAF* and *MAFB* expression levels are shown at the bottom of the middle panel, although the analysis criteria did not allow their selection. The z_g scores and average fold (FC) changes are expressed using the specific translocation group as baseline. For GenBank Accession numbers, see Supplementary Table 2

those for the *RUVBL1*, *CAD* and *SMARCA4* genes found expressed at significant lower levels in the MGUS than in the PCL samples (Figure 2a). In particular, the *CAD* gene, previously identified as being significantly downregulated in MGUS *versus* both MM and PCL PCs, was also downregulated in these patients, thus supporting the relationship of these MM cases to the MGUS phenotype.

The supervised analysis performed on the four cases clustering with PCL identified a set of 12 overexpressed genes (Figure 7, lower panel), mostly codifying for two known classes of antigenic peptides (GAGE and MAGE) whose normal expression is restricted to male germ cells and which are known to be deregulated in a large variety of tumors including MM (De Backer *et al*., 1999; van Baren *et al*., 1999).

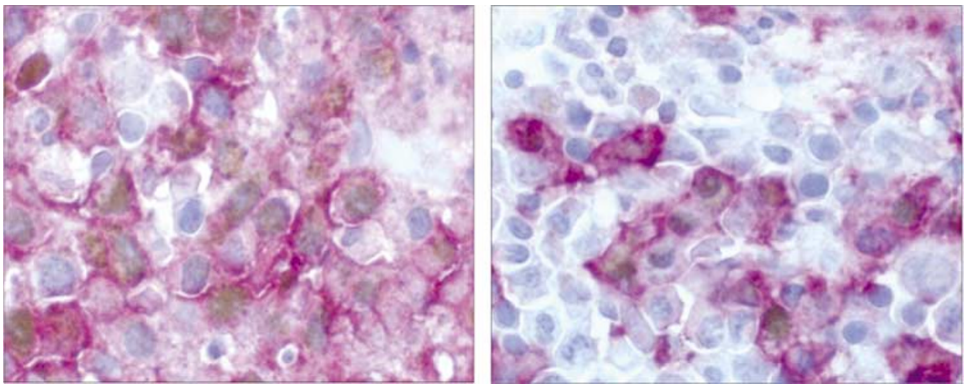

Figure 5 Immunohistochemical staining of bone marrow biopsies for CCND2. Representative results of two different MM cases carrying the *MAF*/*MAFB* translocations are reported showing nuclear CCND2 immunoreactivity (brown) in most of the CD138$^+$ PCs (red). Original magnification: $\times 40$

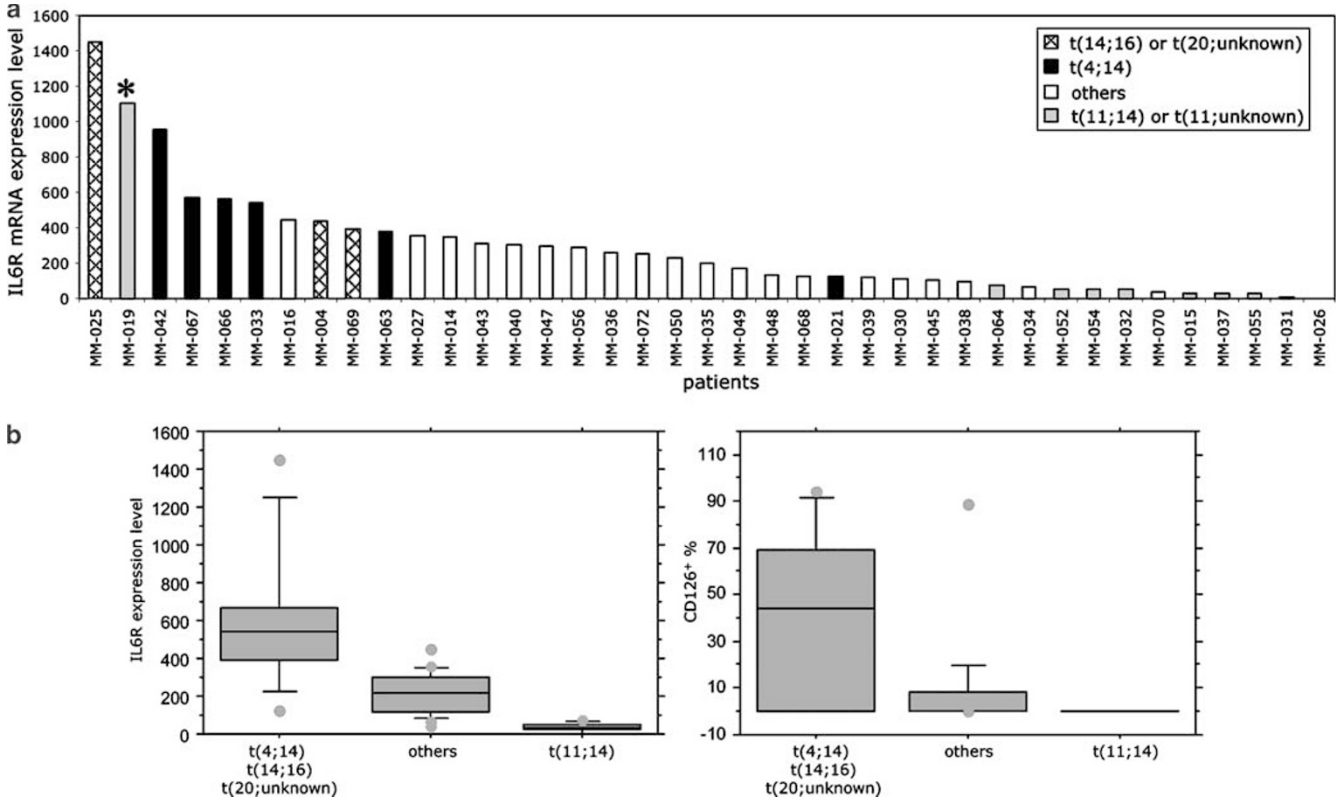

Figure 6 IL6-R expression in MM PCs. (**a**) *IL6-R* mRNA expression levels in the 39 MM cases as assessed by microarray analysis. (**b**) Box plot of mRNA (left) and CD126 protein (right) expression levels obtained by means of DNA microarray and flow cytometry analyses of CD138$^+$ PCs, respectively. MM cases are grouped according to the presence of different translocations. MM-019 patient (*) was excluded because positive for 4p16.3 locus rearrangement and *WHSC1* overexpression

GAGE genes were mapped on chromosome Xq11.23, reported to be a breakpoint site in synovial sarcoma (de Leeuw *et al*., 1994). Given the high degree of genomic instability that characterizes MM, we investigated by FISH analysis the region for the presence of any alterations. The results (data not shown) indicated the absence of any structural rearrangement targeting this locus, thus supporting the notion of DNA demethylation as the primary inducing mechanism of these genes expression (De Backer *et al*., 1999).

Discussion

Recent reports concerning the gene expression profiling of normal and malignant PCs (including MGUS and

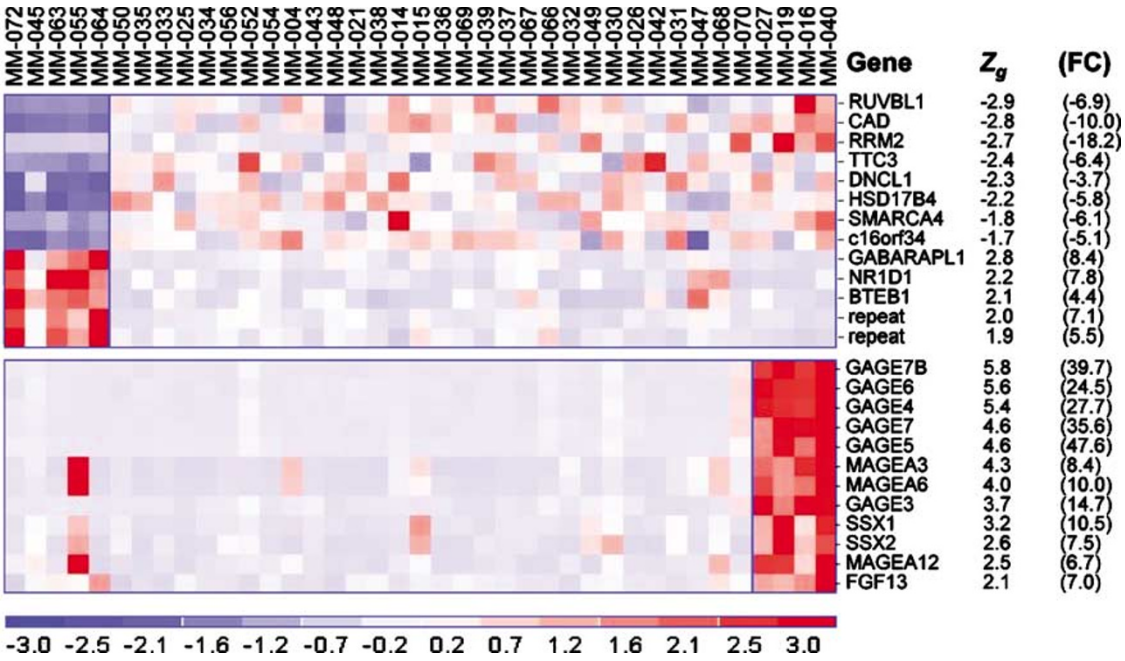

Figure 7 Identification of genes differentially expressed in '*MGUS-resembling*' and '*PCL-resembling*' MM patients. Supervised analyses were performed comparing the five MM samples clustering with MGUS in the unsupervised analysis (top panel) or the four patients clustering with PCL (bottom panel) with all the other MM samples. The genes, z_g scores and fold changes (FC) are indicated. For GenBank Accession numbers, see Supplementary Table 3

MM but not PCL cases) have suggested that normal PCs can be differentiated from MGUS and MM, whereas MGUS and MM cannot be easily distinguished (Zhan *et al*., 2002; Davies *et al*., 2003). One of the major findings of our study (the first to include PCL cases) is that the MM phenotype is highly heterogeneous at transcriptional level. MM and PCL samples are not distinguishable, and some MM cases show features closely related to MGUS. Conversely, MGUSs are defined as a clearly recognizable group, differentiated from PCLs on the basis of the expression of a large number of genes mainly involved in controlling cell proliferation and DNA metabolism, a finding that is consistent with the more proliferative phenotype of PCL PCs.

The gene that was most significantly overexpressed in PCL *versus* MGUS PCs was *WHSC1*, which is targeted by the t(4;14) in MM. The six PCL cases were indeed negative for structural alterations of the 4p16.3 region, as evaluated by FISH (Fabris *et al*., 2005). Although the function of this gene is still unknown, it has been suggested that it may be involved in chromatin remodeling (Huang *et al*., 1998; Garlisi *et al*., 2001). Even though the expression of *WHSC1* in PCLs is not as high as that observed in MM with t(4;14), we can speculate that its induction may be related to the activation of the chromatin remodeling pathway as a consequence of the proliferative activity of PCL PCs. Overall, this finding suggests that *WHSC1* may contribute to the progression of the disease and further supports its putative role in the neoplastic transformation of tumors with the t(4;14).

Notably, the majority of genes distinguishing MGUS from MM were also differentially expressed in MGUS *versus* PCL, suggesting that their expression may be progressively modulated during the transition of PCs from a preneoplastic to a fully malignant phenotype. Among these genes, three MYC-targets (*CAD*, *PRDX3* and *GADD45B*) (Fernandez *et al*., 2003) were included. This finding, together with the observation of a progressive induction of *MYC* expression associated with the worsening in the clinical presentation of MM cases, as assessed by clinical staging (see Supplementary Figure 1), suggests that the induction of the MYC pathway may play a role in MM. Moreover, the finding of *IL-8* downregulation in both MM and PCL samples possibly reflects a progressive independence of malignant PCs from microenvironment homing factors.

An interesting result of our study is that the hierarchical clustering of MM PCs defined two major groups mainly driven by the presence/absence of the most recurrent IGH translocations. Previous gene expression profiling data defined four major MM clusters (Zhan *et al*., 2002); however, no information was provided concerning the correlation with the IGH translocations. We found no significant association with the clinical staging or the prognostic parameters analysed, except for a lower incidence of bone lesions within the group positive for IGH translocations, although an overall higher frequency of bone lesions was correlated with the presence of t(11;14), in line with recently reported data (Hideshima *et al*., 2004; Robbiani *et al*., 2004).

The overexpression of *FGFR3/WHSC1* in t(4;14) PCs leads to the upregulation of *KLF4*, *CSDA* and *NGFRAP1* genes reported to be involved in the induction of apoptosis (Coles *et al*., 1996; Mukai *et al*., 2000; Chen *et al*., 2001). In particular, the overexpression of *KLF4* is intriguing as it has been reported to mediate the transcriptional repression of the *CCND1* promoter (Shie *et al*., 2000); the low or background levels of *CCND1* found in the patients with the t(4;14) (data not shown) may therefore be related to its increased expression. This finding may suggest that cases with the t(4;14) are characterized by the activation of proapoptotic pathways, and thus may be more likely to respond to chemotherapy regimens, consistently with recently reported data (Stewart *et al*., 2004). During the revision process of this manuscript, a study was published describing the gene expression profiling of t(4;14) MM cases (Dring *et al*., 2004) whose results only partially overlap with our own, probably because of the different analysis criteria. Instead of the generally used fold-ratio to describe gene expression changes among cell subsets, the algorithm applied in our supervised analysis: (i) discovers gene expression signatures common to an entire set of experiments; and (ii) considers the consistency of the expression levels throughout the phenotype dataset (Califano, 2000).

The downregulation of *IL6-R* was the most consistent finding in patients with t(11;14). It has been reported that *IL6-R* (CD126) expression is restricted to neoplastic PCs, with little or no detectable expression by normal PCs (Rawstron *et al*., 2000). Our data indicate a heterogeneous expression of CD126 in MM PCs and show that patients characterized by t(4;14) or translocated *MAF* genes exhibit the highest levels. Although the significance of CD126 expression in neoplastic PCs remains to be fully clarified, we can argue that lower levels of IL6-R may reduce the proliferative capacity of myeloma cells in response to IL-6. It is worth noting that patients with t(11;14) have improved survival especially after high-dose therapy (Fonseca *et al*., 2002; Moreau *et al*., 2002), whereas those with t(4;14) and t(14;16) show a shorter survival after both standard and high-dose therapy (Keats *et al*., 2003). Further studies are warranted to investigate the biological significance of this finding and its potential clinical relevance.

Despite the limited number of cases bearing chromosomal translocations involving *MAF/MAFB* loci (consistent with their low incidence in MM), our analysis identified, as part of the MAF signature, genes that have been previously reported as transcriptional target of the *MAF* gene (Hurt *et al*., 2004), as well as a number of genes involved in cell adhesion processes. These data support the hypothesis that the deregulated expression of MAF transcription factors targets the interaction between PCs and the microenvironment through the activation of cell adhesion and chemotaxis pathways. Furthermore, the ectopic expression of chemokines also suggests the presence of the autocrine production of factors promoting cell adhesion to the bone marrow. The corroboration of these finding on a larger number of cases carrying translocations targeting the *MAF* genes is needed to assess their potential significance.

Finally, MM cases clustering with PCL are characterized by the expression of GAGE and MAGE tumor antigens previously described in MM (van Baren *et al*., 1999). In our panel, their expression is specifically restricted to the MM cases characterized by an aggressive clinical evolution (course <1 years) and whose gene expression profiles resemble those of PCLs, which on the contrary do not show deregulated expression of these genes. A more aggressive phenotype has been described as being associated with high levels of MYC and a lack of CD56 (Sahara *et al*., 2002; Rasmussen *et al*., 2003). Concerning our restricted panel of cases with a 'PCL-like' phenotype, we detected high levels of *MYC* expression that only partially correlated with the lack of CD56 expression (as detected by both microarray and FACS analysis) (data not shown). The analysis of a larger number of MM cases is needed to assess the clinical and therapeutic implications of these findings, because the expression of these tumor antigens seem to identify a restricted number of high-risk patients who may represent good candidates for antibody-mediated immunotherapy.

In conclusion, our data extend previous evidence that the marked heterogeneity of MM patients is associated with a wide spectrum of molecular diversity based on global gene expression profiling. We showed that this diversity is mainly driven by different chromosomal translocations, and identified potential molecular pathways specifically related to such genetic lesions. These data provide a contribution to the investigation of the molecular pathogenesis of the neoplasia and toward a molecular-based classification of MM for the identification of prognostically relevant entities.

Materials and methods

Patients

Bone marrow samples were collected after informed consent, during standard diagnostic procedures, from seven MGUS, 39 newly diagnosed MM and six PCL patients. The median age of the patients (22 female and 30 male subjects) was 65 years (range 39–82 years). The diagnoses were made by means of conventional morphology and immunophenotype analyses; the clinical staging of the 39 MM patients (13 stage I, 12 stage II and 14 stage III) was defined according to the criteria described by Durie and Salmon (1975). The monoclonal components were as follows: IgG (35 patients); IgA (seven); IgD (one patient); κ chain (five); λ chain (one patient); double monoclonal component (three patients); κ/λ ratio 2.5 for MGUS, 1.6 for MM and 5 for PCL patients. β2-microglobulin levels were $\geqslant 2.4$ mg/l in 34/52 patients. Focal bone lesions (detected by radiography and/or magnetic resonance imaging) were present in 26 patients: score 1 was assigned to patients having $\leqslant 3$ lesions ($n=18$), and 2 to patients with more than three lesions ($n=8$). All of the MM and PCL and 2/7 MGUS patients were characterized by means of fluorescence *in situ* hybridization (FISH) for the presence of the most recurrent IGH translocations: t(4;14)(p16.3;q32) was detected in six MM

cases; t(11;14)(q13;q32) in 11 (one MGUS, nine MM, one PCL); t(14;16)(q32;q23) in two MM and one PCL; t(14;20)(q32;q12) in one PCL and t(6;14)(p21;q32) in one MM patient. In addition, two cases (MM-054; PCL-002) showed a *CCND1* translocation at the 11q13 and two cases (MM-069; PCL-006) a *MAFB* translocation at the 20q not involving the Ig loci; MM-019 showed a rearrangement in the region 4p16.3 in addition to a t(11;14) translocation. Finally, chromosome 13 deletions were found in 17/35 MM, 1/5 PCL and 1/2 MGUS patients (Fabris *et al*., 2005).

PCs purification

PCs were purified from bone marrow samples, after red blood cell lysis with 0.86% ammonium chloride, using CD138 immunomagnetic microbeads according to the manufacturer's instructions (MidiMACS® system, Miltenyi Biotec, Auburn, CA, USA). The purity of the positively selected PCs was assessed by morphology and flow cytometry and was $\geqslant 90\%$ in all cases.

Total RNA extraction and purification

Total RNA was extracted from frozen PCs samples using the TRIZOL® reagent (Invitrogen, Carlsbad, CA, USA), and purified using the RNeasy® total RNA Isolation Kit (Qiagen, Valencia, CA, USA). RNA integrity was verified by means of an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany).

Generation of gene expression profiles

Biotin-labeled cRNA was prepared using the GeneChip® Expression Analysis Technical Manual protocol (Affymetrix Inc., Santa Clara, CA, USA) as previously described (Klein *et al*., 2001). Briefly, 5 μg of total RNA were converted into double-strand cDNA using the Custom SuperScript™ Double-Stranded cDNA Synthesis kit (Invitrogen, Carlsbad, CA, USA). The cDNA was phenol–chloroform extracted with Phase Lock Gel Light™ (Eppendorf, Hamburg, Germany). Biotinylated cRNA was generated by means of an *in vitro* transcription reaction using the Enzo BioArray™ HighYield™ RNA transcript Labeling Kit (Enzo Life Science, Inc., Farmingdale, NY, USA) and purified using the RNeasy® total RNA Isolation Kit (Qiagen, Valencia, CA, USA). In all, 15 μg of fragmented cRNA were hybridized on GeneChip® Human Genome U133A Arrays (Affymetrix Inc., Santa Clara, CA, USA) after quality checking on GeneChip® Test3 Arrays (Affymetrix Inc., Santa Clara, CA, USA). After scanning, the images were processed using Affymetrix® MicroArray Suite (MAS) 5.0 software to generate gene expression intensity values. The data discussed in this publication have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE2113.

Microarray data analysis

Arrays normalization was performed using MAS 5.0 'global scaling' procedure, which normalizes the signals of different experiments to the same target intensity. The genes whose expression did not vary across the dataset were not considered for further analyses. Unsupervised analyses were applied to a subset of genes whose average change in expression levels varied at least twofold from the mean across the whole panel. Hierarchical agglomerative clustering and dendrogram generation were used to search for natural groupings in the profiles using the DNAChip Analyzer (Schadt *et al*., 2001). Before clustering, the expression values for a gene across all samples were standardized to have a mean of 0 and a standard deviation of 1, and these standardized values were used to calculate correlations between genes and samples, and served as the basis for merging nodes. The distance between two genes or individual samples was calculated as one *minus* Pearson correlation coefficient between the standardized expression values, and the average-linkage method (Eisen *et al*., 1998) was used for cluster merging.

The presence of significant clusters was assessed by applying the global test of clustering stability implemented in the BRB ArrayTools microarray analysis package (McShane *et al*., 2002). This statistical method allows computing the strength of meaningful clusters as a function of the level at which the dendrogram are 'cut' (e.g. depending on the distance at which the agglomerative process was stopped). Two cluster-specific reproducibility measures that can aid in assessing the meaningfulness of individual clusters of interest were defined: the robustness R-index, which measures the proportion of pairs of samples within a cluster that remain grouped together despite any data perturbation, and the discrepancy D-index, which measures the number of added and deleted specimens when comparing a cluster before and after data perturbation. The overall cluster stability was computed perturbing the original data 100 times using artificial experimental noise in the form of a Gaussian error estimate from the data.

The supervised gene expression analyses were performed using the Gene@Work software platform, which is a gene expression analysis tool based on the pattern discovery algorithm Structural Pattern Localization Analysis by Sequential Histograms (SPLASH) (Califano, 2000; Klein *et al*., 2001). Genes@Work discovers global gene expression 'signatures' that are common to an entire set of at least n experiments (the support set), where n is a user-selectable parameter called the minimum support. Briefly, differentially expressed genes are identified by comparing an expected gene expression probability density $p(e)$, empirically computed from the experimental set with a predefined threshold (the parameter δ). Patterns of differentially expressed genes are then ranked according to their statistical significance (the pattern score or *z-score*). Unless otherwise stated, the value of δ for the analysis was set to 0.02, and the support value was chosen as $n = n_0 - 1$, being n_0 the number of samples in the phenotype set. For each gene, the statistical significance of the differential expression across the phenotype and control sets (gene *z-score*, z_g) was computed using the formula $z_g = (\mu_p - \mu_c)/(\sigma_p + \sigma_c)$, where μ_p and σ_p are, respectively, the mean and s.d. computed from the gene expression values for that gene in the phenotype group, and μ_c and σ_c are their corresponding values computed from the control group.

Q-RT–PCR

In total, 50 ng of total RNA were transcribed into cDNA with random hexamer primers and analysed according to published protocol (Heid *et al*., 1996) using an ABI PRISM™ 7700 Sequence Detection System (Applied Biosystems, Foster City, CA, USA). TaqMan probe and primers were designed using Primer Express (Applied Biosystems, Foster, CA, USA). Human *WHSC1* (type II, Accession Number AF071593) and endogenous control *GAPDH* (TaqMan® Pre-Developed Assay Reagents, PDARs; Applied Biosystems, Foster, CA, USA) were analysed and the results were expressed as *WHSC1* RNA (ng)/*GAPDH* RNA (ng). *WHSC1* levels were detected with the forward primer 5′-CTCTTGGG TTTAGCAAAAGTTC ATC-3′, the reverse primer 5′-TTTCGTCTGCACTTTCG TATGG-3′ and the probe 6-*FAM*-CCTTAACTGAGAAT

GAGGTCTCGGACAGC-*TAMRA*. The reactions were performed in triplicate.

Immunohistochemistry

Paraffin-embedded sections from bone marrow biopsies fixed and decalcified in Mielodec® (Bio-Optica, Milan, Italy) were pretreated with an antigen retrieval solution at 99°C for 30 min, and then incubated with a 1/100 dilution of anti-CCND2 polyclonal antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA) in Tris-buffered saline (TBS) overnight at 4°C. Detection was performed using the LSAB™ kit (Dako, Glostrup, Denmark) according to the manufacturer's instructions. Peroxidase activity was developed with 3-3-diaminobenzidine-copper sulfate (Sigma Chemical Co., St Louis, MO, USA). The cases that proved to be immunoreactive for CCND2 were then incubated with a 1/2000 dilution of anti-CD138 monoclonal antibody (clone B-B4, Dako, Glostrup, Denmark) in TBS for 1 h at room temperature. Alkaline phosphatase activity was developed with the fuchsine substrate system (Dako, Glostrup, Denmark), according to the manufacturer's instructions.

Flow cytometry analysis

Dual color immunophenotyping analysis was performed on heparinized bone marrow samples by CD138-FITC (Immuno Quality Products, Groningen, The Netherlands) and CD126-PE (Immunotech, Marseilles, France) antibodies. After staining, the cells were analysed using a FACSCalibur™ Flow Cytometer (Becton Dickinson, Franklin Lakes, NJ USA).

Acknowledgements
This work was supported by grants from the Associazione Italiana Ricerca sul Cancro (to AN), Ministero della Salute; Ministero dell'Universita' e Ricerca Scientifica (FIRB projects RBAU012E9A to AN, RBNE01TZZ8 and RBAU01935A to SB) and Associazione Studio Malati Ematologici, Milano.

References

Aderem A and Ulevitch RJ. (2000). *Nature*, **406,** 782–787.
Anderson KC, Shaughnessy Jr JD, Barlogie B, Harousseau JL and Roodman GD. (2002). *Hematology (Am. Soc. Hematol. Educ. Program)*, **2002,** 214–240.
Barille-Nion S, Barlogie B, Bataille R, Bergsagel PL, Epstein J, Fenton RG, Jacobson J, Kuehl WM, Shaughnessy J and Tricot G. (2003). *Hematology (Am. Soc. Hematol. Educ. Program)*, **2003,** 248–278.
Bergsagel PL and Kuehl WM. (2001). *Oncogene*, **20,** 5611–5622.
Burger JA, Burger M and Kipps TJ. (1999). *Blood*, **94,** 3658–3667.
Califano A. (2000). *Bioinformatics*, **16,** 341–357.
Chen X, Bonne S, Hatzfeld M, van Roy F and Green KJ. (2002). *J. Biol. Chem.*, **277,** 10512–10522.
Chen X, Johns DC, Geiman DE, Marban E, Dang DT, Hamlin G, Sun R and Yang VW. (2001). *J. Biol. Chem.*, **276,** 30423–30428.
Chesi M, Bergsagel PL, Brents LA, Smith CM, Gerhard DS and Kuehl WM. (1996). *Blood*, **88,** 674–681.
Chesi M, Bergsagel PL, Shonukan OO, Martelli ML, Brents LA, Chen T, Schrock E, Ried T and Kuehl WM. (1998a). *Blood*, **91,** 4457–4463.
Chesi M, Nardini E, Brents LA, Schrock E, Ried T, Kuehl WM and Bergsagel PL. (1997). *Nat. Genet.*, **16,** 260–264.
Chesi M, Nardini E, Lim RS, Smith KD, Kuehl WM and Bergsagel PL. (1998b). *Blood*, **92,** 3025–3034.
Coles LS, Diamond P, Occhiodoro F, Vadas MA and Shannon MF. (1996). *Nucleic Acids Res.*, **24,** 2311–2317.
Davies FE, Dring AM, Li C, Rawstron AC, Shammas MA, O'Connor SM, Fenton JA, Hideshima T, Chauhan D, Tai IT, Robinson E, Auclair D, Rees K, Gonzalez D, Ashcroft AJ, Dasgupta R, Mitsiades C, Mitsiades N, Chen LB, Wong WH, Munshi NC, Morgan GJ and Anderson KC. (2003). *Blood*, **102,** 4504–4511.
De Backer O, Arden KC, Boretti M, Vantomme V, De Smet C, Czekay S, Viars CS, De Plaen E, Brasseur F, Chomez P, Van den Eynde B, Boon T and van der Bruggen P. (1999). *Cancer Res.*, **59,** 3157–3165.
de Leeuw B, Suijkerbuijk RF, Olde WD, Meloni AM, Stenman G, Kindblom LG, Balemans M, van den Berg E, Molenaar WM and Sandberg AA. (1994). *Cancer Genet. Cytogenet.*, **73,** 89–94.
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 and Klein B. (2002). *Oncogene*, **21,** 6848–6857.
Dring AM, Davies FE, Fenton JA, Roddam PL, Scott K, Gonzalez D, Rollinson S, Rawstron AC, Rees-Unwin KS, Li C, Munshi NC, Anderson KC and Morgan GJ. (2004). *Clin. Cancer Res.*, **10,** 5692–5701.
Durie BG and Salmon SE. (1975). *Cancer*, **36,** 842–854.
Eisen MB, Spellman PT, Brown PO and Botstein D. (1998). *Proc. Natl. Acad. Sci. USA*, **95,** 14863–14868.
Fabris S, Agnelli L, Mattioli M, Baldini L, Ronchetti D, Morabito F, Verdelli D, Nobili L, Intini D, Callea V, Stelitano C, Lombardi L and Neri A. (2005). *Genes Chromosomes Cancer*, **42,** 117–127.
Fernandez PC, Frank SR, Wang L, Schroeder M, Liu S, Greene J, Cocito A and Amati B. (2003). *Genes Dev.*, **17,** 1115–1129.
Fonseca R, Blood EA, Oken MM, Kyle RA, Dewald GW, Bailey RJ, Van Wier SA, Henderson KJ, Hoyer JD, Harrington D, Kay NE, Van Ness B and Greipp PR. (2002). *Blood*, **99,** 3735–3741.
Garlisi CG, Uss AS, Xiao H, Tian F, Sheridan KE, Wang L, Motasim BM, Egan RW, Stranick KS and Umland SP. (2001). *Am. J. Resp. Cell Mol. Biol.*, **24,** 90–98.
Hanamura I, Iida S, Akano Y, Hayami Y, Kato M, Miura K, Harada S, Banno S, Wakita A, Kiyoi H, Naoe T, Shimizu S, Sonta SI, Nitta M, Taniwaki M and Ueda R. (2001). *Jpn. J. Cancer Res.*, **92,** 638–644.
Heid CA, Stevens J, Livak KJ and Williams PM. (1996). *Genome Res.*, **6,** 986–994.
Hideshima T, Bergsagel PL, Kuehl WM and Anderson KC. (2004). *Blood*, **104,** 607–618.
Ho PJ, Campbell LJ, Gibson J, Brown R and Joshua D. (2002). *Rev. Clin. Exp. Hematol.*, **6,** 276–300.
Huang N, vom Baur E, Garnier JM, Lerouge T, Vonesch JL, Lutz Y, Chambon P and Losson R. (1998). *EMBO J.*, **17,** 3398–3412.
Hurt EM, Wiestner A, Rosenwald A, Shaffer AL, Campo E, Grogan T, Bergsagel PL, Kuehl WM and Staudt LM. (2004). *Cancer Cell*, **5,** 191–199.

Imaizumi T, Yoshida H and Satoh K. (2004). *J. Atheroscler. Thromb.*, **11,** 15–21.
Jin S, Tong T, Fan W, Fan F, Antinore MJ, Zhu X, Mazzacurati L, Li X, Petrik KL, Rajasekaran B, Wu M and Zhan Q. (2002). *Oncogene*, **21,** 8696–8704.
Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ, Belch AR and Pilarski LM. (2003). *Blood*, **101,** 1520–1529.
Klein U, Tu Y, Stolovitzky GA, Mattioli M, Cattoretti G, Husson H, Freedman A, Inghirami G, Cro L, Baldini L, Neri A, Califano A and Dalla-Favera R. (2001). *J. Exp. Med.*, **194,** 1625–1638.
McShane LM, Radmacher MD, Freidlin B, Yu R, Li MC and Simon R. (2002). *Bioinformatics*, **18,** 1462–1469.
Moreau P, Facon T, Leleu X, Morineau N, Huyghe P, Harousseau JL, Bataille R and Avet-Loiseau H. (2002). *Blood*, **100,** 1579–1583.
Mukai J, Hachiya T, Shoji-Hoshino S, Kimura MT, Nadano D, Suvanto P, Hanaoka T, Li Y, Irie S, Greene LA and Sato TA. (2000). *J. Biol. Chem.*, **275,** 17566–17570.
Rasmussen T, Theilgaard-Monch K, Hudlebusch HR, Lodahl M, Johnsen HE and Dahl IM. (2003). *Br. J. Haematol.*, **123,** 253–262.
Rawstron AC, Fenton JA, Ashcroft J, English A, Jones RA, Richards SJ, Pratt G, Owen R, Davies FE, Child JA, Jack AS and Morgan G. (2000). *Blood*, **96,** 3880–3886.
Richelda R, Ronchetti D, Baldini L, Cro L, Viggiano L, Marzella R, Rocchi M, Otsuki T, Lombardi L, Maiolo AT and Neri A. (1997). *Blood*, **90,** 4062–4070.
Robbiani DF, Chesi M and Bergsagel PL. (2004). *N. Engl. J. Med.*, **351,** 197–198.
Ronchetti D, Finelli P, Richelda R, Baldini L, Rocchi M, Viggiano L, Cuneo A, Bogni S, Fabris S, Lombardi L, Maiolo AT and Neri A. (1999). *Blood*, **93,** 1330–1337.
Sahara N, Takeshita A, Shigeno K, Fujisawa S, Takeshita K, Naito K, Ihara M, Ono T, Tamashima S, Nara K, Ohnishi K and Ohno R. (2002). *Br. J. Haematol.*, **117,** 882–885.
Schadt EE, Li C, Ellis B and Wong WH. (2001). *J. Cell Biochem. Suppl.*, **37** (Suppl.)**,** 120–125.
Seidl S, Kaufmann H and Drach J. (2003). *Lancet Oncol.*, **4,** 557–564.
Shaughnessy Jr J, Gabrea A, Qi Y, Brents L, Zhan F, Tian E, Sawyer J, Barlogie B, Bergsagel PL and Kuehl M. (2001). *Blood*, **98,** 217–223.
Shie JL, Chen ZY, Fu M, Pestell RG and Tseng CC. (2000). *Nucleic Acids Res.*, **28,** 2969–2976.
Stewart JP, Thompson A, Santra M, Barlogie B, Lappin TR and Shaughnessy Jr J. (2004). *Br. J. Haematol.*, **126,** 72–76.
Suzuki A, Kusakai G, Kishimoto A, Lu J, Ogura T and Esumi H. (2003). *Oncogene*, **22,** 6177–6182.
Tian E, Zhan F, Walker R, Rasmussen E, Ma Y, Barlogie B and Shaughnessy Jr JD. (2003). *N. Engl. J. Med.*, **349,** 2483–2494.
van Baren N, Brasseur F, Godelaine D, Hames G, Ferrant A, Lehmann F, Andre M, Ravoet C, Doyen C, Spagnoli GC, Bakkus M, Thielemans K and Boon T. (1999). *Blood*, **94,** 1156–1164.
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 and Shaughnessy Jr J. (2002). *Blood*, **99,** 1745–1757.
Zhan F, Tian E, Bumm K, Smith R, Barlogie B and Shaughnessy Jr J. (2003). *Blood*, **101,** 1128–1140.

Supplementary Information accompanies the paper on Oncogene website (http://www.nature.com/onc)