Molecular Dissection of Hyperdiploid Multiple Myeloma by Gene Expression Profiling

Wee J. Chng,[1] Shaji Kumar,[2] Scott VanWier,[1] Greg Ahmann,[1] Tammy Price-Troska,[2] Kim Henderson,[2] Tae-Hoon Chung,[3] Seungchan Kim,[3,4] George Mulligan,[5] Barbara Bryant,[5] John Carpten,[3] Morie Gertz,[2] S. Vincent Rajkumar,[2] Martha Lacy,[2] Angela Dispenzieri,[2] Robert Kyle,[2] Philip Greipp,[2] P. Leif Bergsagel,[1] and Rafael Fonseca[1]

[1]Department of Hematology-Oncology, Mayo Clinic, Scottsdale, Arizona; [2]Division of Hematology, Mayo Clinic, Rochester, Minnesota; [3]Translational Genomics Research Institute, Phoenix, Arizona; [4]Department of Computer Science and Engineering, Arizona State University, Tempe, Arizona; and [5]Millenium Pharmaceuticals, Inc., Cambridge, Massachusetts

Abstract

Hyperdiploid multiple myeloma (H-MM) is the most common form of myeloma. In this gene expression profiling study, we show that H-MM is defined by a protein biosynthesis signature that is primarily driven by a gene dosage mechanism as a result of trisomic chromosomes. Within H-MM, four independently validated patient clusters overexpressing nonoverlapping sets of genes that form cognate pathways/networks that have potential biological importance in multiple myeloma were identified. One prominent cluster, cluster 1, is characterized by high expression of cancer testis antigen and proliferation-associated genes. Tumors from these patients were more proliferative than tumors in other clusters (median plasma cell labeling index, 3.8; $P < 0.05$). Another cluster, cluster 3, is characterized by genes involved in tumor necrosis factor/nuclear factor-κB signaling and antiapoptosis. These patients have better response to bortezomib as compared with patients within other clusters (70% versus 29%; $P = 0.02$). Furthermore, for a group of patients generally thought to have better prognosis, a cluster of patients with short survival (cluster 1; median survival, 27 months) could be identified. This analysis illustrates the heterogeneity within H-MM and the importance of defining specific cytogenetic prognostic factors. Furthermore, the signatures that defined these clusters may provide a basis for tailoring treatment to individual patients. [Cancer Res 2007;67(7):2982–9]

Introduction

Multiple myeloma is an incurable plasma cell malignancy characterized by marked genetic heterogeneity that has been gradually unraveled in recent years. It is now recognized that there are two broad genetic subtypes of multiple myeloma as defined by chromosome number: hyperdiploid multiple myeloma (H-MM, 48–74 chromosomes), which is characterized by trisomies of chromosomes 3, 5, 7, 9, 11, 15, 19, and 21 and a lower prevalence of primary translocations involving the immunoglobulin heavy chain (IgH) locus at 14q32, and nonhyperdiploid multiple myeloma (NH-MM; <48 or >75 chromosomes), which is associated with the presence of primary IgH translocations such as t(4;14), t(11;14), and t(14;16) (refs. 1, 2). The current model of molecular pathogenesis of multiple myeloma proposes a dichotomous pathway for the two main genetic subtypes of multiple myeloma because characteristic changes are already observed at the premalignant monoclonal gammopathy of undetermined significance stage (3).

H-MM constitutes ∼50% to 60% of multiple myeloma (1, 4), yet little of its biology is known. Overall survival is favorable for patients with H-MM, but some H-MM patients clearly have a more aggressive variant of the disease (5, 6). In view of this, we hypothesize that H-MM is likely to be biologically heterogeneous, and understanding this heterogeneity may provide a better understanding of the underlying biology, identify possible molecular therapeutic targets, and allow better risk stratification of patients.

In this study, we use gene expression profiling (GEP) to characterize the molecular profile of H-MM, with a view of gaining some insights about its biology, and to study the heterogeneity within H-MM.

Materials and Methods

Patients

Mayo Clinic (group 1). Fifty-three H-MM and 37 NH-MM with adequately good quality RNA for gene expression study were included. Of the NH-MM patients, 29 were newly diagnosed and 8 had relapsed disease. Among the 53 H-MM patients, 37 were newly diagnosed and 16 had relapsed multiple myeloma. Bone marrow aspirate samples were obtained after informed consent in accordance with the Helsinki Declaration and enriched for CD138+ cells (plasma cells) using immunomagnetic bead separation (AutoMACS, Miltenyi-Biotec, Auburn, CA). In addition, IgH translocations and chromosome 13 deletion (Δ13) were assessed by fluorescence *in situ* hybridization (FISH) with concurrent immunofluorescent staining for clonotypic cytoplasmic immunoglobulin (7) using probes previously described (8). Plasma cell labeling index (PCLI) was assessed using a slide-based method as previously described (9). Where data were available, patients were also staged according to the International Staging System (10). A summary of clinical information of H-MM patients is presented in Supplementary Table S1. Survival data were available from 50 of the H-MM patients (37 newly diagnosed and 13 relapsed). Twenty-seven of the 50 patients have died. The median and minimum durations of follow-up from time of sampling for the surviving patients are 30 and 13 months, respectively. The study was approved by the Mayo Clinic Institutional Review Board.

Patients entered into bortezomib trials (group 2). For validation of subtype signatures detected in group 1 patients, we used a cohort of 86 H-MM entered into bortezomib trials (11). For the assessment of response and progression-free survival to bortezomib, we analyzed a subset of 63 H-MM patients entered into the international phase 3 clinical trial (APEX) of

bortezomib versus dexamethasone. In total, 29 patients received single-agent bortezomib whereas the other 34 patients were randomized to dexamethasone. These patients all had relapsed disease. The European Blood and Marrow Transplant group criteria were used to defined response and progression in these patients (12). In this analysis, we considered patients with minimal response or better as having a response. For response analysis, only patients treated with bortezomib were included.

University of Arkansas patients (group 3). As an additional validation cohort, 91 newly diagnosed patients with H-MM from the University of Arkansas with publicly available gene expression data were assessed (13).

Assignment of Hyperdiploid Cases

For group 1 patients, hyperdiploidy was determined by a FISH-based trisomy index that is highly specific for H-MM (14). For group 2 and 3 patients, the H-MM subset is identified using the TC (*T*ranslocation and *C*yclin D) classification (13) based on the expression of only seven genes (*FGFR3*, *MMSET*, *ITGB7*, *MAF*, *CCND3*, *CCND1*, and *CCND2*). Only those patients assigned the D1 TC group (aberrant low level of *CCND1* expression) that has been shown to unambiguously correspond to H-MM (13, 15) were selected for further analysis. (For details, refer to Supplementary methods).

Gene Expression Profiling

RNA isolation, purification, and microarray hybridization have previously been reported for group 1 (16), group 2 (Mulligan et al. Blood, in press), and group 3 (13). Gene expression analysis was done on CD138[+] selected (groups 1 and 3) and negative-selection enriched (group 2) plasma cell RNA using the Affymetrix U133A (group 1), the U133A and B (group 2), and U95Av2 chips (group 3; Affymetrix, Santa Clara, CA). For the Mayo cohort, gene expression intensity values were generated using the Affymetrix MAS 5.0 software. For normalization, the data were scaled by global methods to a target value of 500, log transformed, median centered, and analyzed using GeneSpring 7 (Agilent Technologies, Palo Alto, CA). The gene expression data are deposited on the Gene Expression Omnibus[6] (accession no. GSE 6477).

Gene Expression Analysis

To define the molecular signature of H-MM, we did a Student's *t* test with Benjamini and Hochberg multiple testing corrections between hyperdiploid and nonhyperdiploid cases in group 1. Differentially expressed genes with adjusted $P < 0.01$ were selected for further Gene Ontology analysis using the Gene Ontology mining tool available in GeneSpring and pathway/network analysis using a web-based software tool, MetaCore (GeneGo, Inc., St Joseph, MI). MetaCore contains an interactive, manually annotated database derived from literature publications on proteins and small molecules that allows for representation of biological functionality and integration of functional, molecular, or clinical information. Several algorithms to enable both the construction and analysis of gene networks are integrated as previously described (17). The output *P* values reflect scoring, prioritization, and statistical significance of networks according to relevance of input data.

To investigate possible molecular heterogeneity within H-MM, we did a recursive analysis using a combination of supervised and unsupervised methods that allows identification of patient clusters defined by nonoverlapping genes. Genes that varied across the individual H-MM samples were first identified by Welch ANOVA using variance computed by applying the Cross-Gene Error Model based on Deviation from 1 available within GeneSpring. This method overcomes the lack of replicates and variance associated with the individual samples and is similar in principle to variance filtering. The filtered gene list (see Supplementary Table S2 for gene list used for each recursive iteration) was then used to cluster group 1 H-MM samples using the hierarchical agglomerative algorithm (18). Pearson's correlation coefficient and centroid linkage were used as similarity and linkage methods, respectively. The most highly correlated
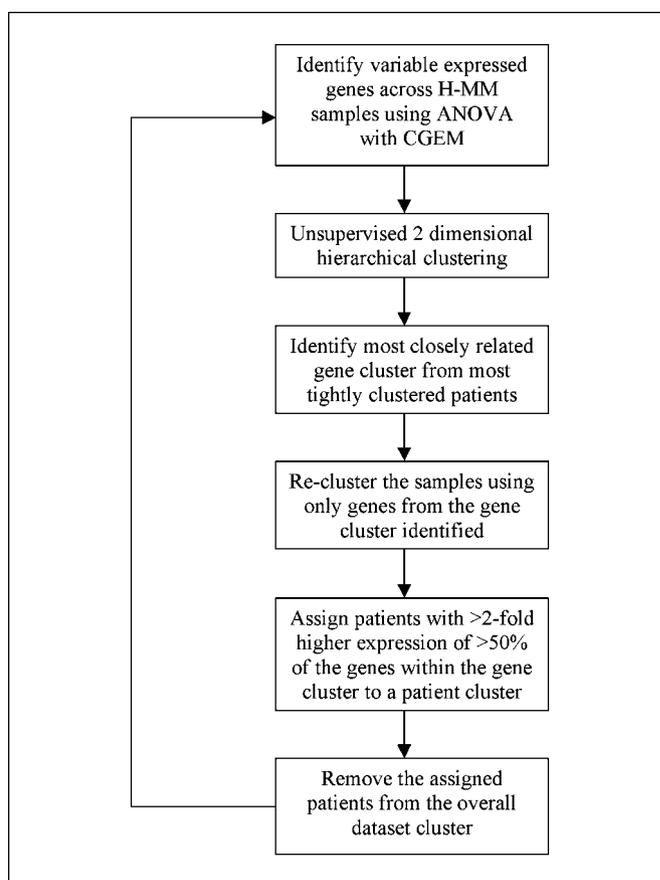

Figure 1. Algorithm for recursive clustering analysis that allows the identification of subclasses within H-MM that are characterized by overexpression of mutually exclusive genes.

gene cluster (at a distance of 0.7) that characterized the most tightly clustered patients (at a distance of 0.7) was selected and used to recluster the patients. Patients with 2-fold higher (than median) expression of 50% of genes within the selected gene cluster were then assigned to a cluster. These patients were removed from the overall cohort and the analysis repeated on the remaining patients until no obvious large clusters could be identified (Fig. 1). Subsequently, genes that distinguished these clusters, requiring that the distinguishing genes have 2-fold higher expression than in patients not in the cluster, and significantly different expression by *t* test with Benjamini and Hochberg multiple testing corrections, were generated. For validation, the same probe sets that distinguished these clusters were used to cluster group 2 samples. For group 3 samples, the best match probe sets on the Affymetrix U95Av2 chip compared with U133A chip were identified using Affymetrix mapping information[7] and used for clustering.

To examine the biology of the differentially expressed genes, we did gene set enrichment analysis, which has been described elsewhere (19). Briefly, the method requires two inputs: (*a*) a list of genes (*L*) ranked based on the correlation between their expression and the class distinction by using a suitable metric and (*b*) a priori defined gene sets (*S*; e.g., pathways or promoter motif sequences extracted from published experimental data or curated databases). The goal of gene set enrichment analysis is to determine whether the members of *S* are randomly distributed throughout *L* or primarily found at the top or bottom, in which case the gene set is correlated with the phenotypic class distinction (in our case, the different H-MM clusters). The ranking metric used was Signal2Noise, and the

---

[6] http://www.ncbi.nlm.nih.gov/geo/

[7] http://www.affymetrix.com/support/technical/comparison_spreadsheets.affx

phenotype was permutated with 1,000 permutations to estimate the statistical significance of enrichment. False discovery rate was controlled at 5% (see Supplementary methods for details of statistics).

Statistical Analysis

Nonparametric Mann-Whitney U test was used for comparison of continuous variables. Categorical variables were compared using the χ^2 test. The distribution for overall survival was estimated using the method of Kaplan and Meier. The log-rank test was used to test for differences in survival between groups. Overall survival was calculated from diagnosis to death for group 1 patients. For group 2 patients, progression-free survival (PFS) was calculated from day of commencing bortezomib until disease progression. $P < 0.05$ was considered significant.

Results

H-MM is characterized by overexpression of genes involved in protein biosynthesis. Five hundred and twenty-two probesets representing 413 genes were differentially expressed between H-MM and NH-MM in group 1 patients. On Gene Ontology analysis, the Gene Ontology categories most associated with these differentially expressed genes were protein biosynthesis and metabolism and ribosomal proteins (Table 1). Seventy-five of the 300 (25%) genes with known function were involved in protein biosynthesis, protein degradation, or protein transport. Twenty-eight of these genes are in the top 100 most differentially expressed genes between H-MM and NH-MM. Fifty-five genes were involved in ribosome/protein biosynthesis, the majority of which were overexpressed in H-MM. Of the seven genes in this category that were underexpressed relative to NH-MM, *EIF4EBP2* is a negative regulator of protein biosynthesis. Thirteen and seven genes were involved in protein degradation and protein transport, respectively. Most of these genes were underexpressed in H-MM (Supplementary Tables S3 and S4). Overall, these data seem to indicate that H-MM, compared with NH-MM, is characterized by increased protein synthesis with defective protein catabolism and transport. Indeed, when the 522 gene probes were analyzed by MetaCore, the most significant and relevant processes/networks were involved in protein biosynthesis. Of interest, several of these networks contained elements important in myeloma biology [nuclear factor-κB (NF-κB), c-myc, and mitogen-activated protein kinase (MAPK)], raising the possibility that one consequence of this protein biosynthesis profile in H-MM is deregulation and overproduction of oncogenic proteins (Supplementary Fig. S1*A*).

Proliferation of H-MM was not statistically different from NH-MM [median PCLI, 0.6 (range, 0–9.4) versus 0.5 (range, 0–6.2); $P = 0.34$], suggesting that the up-regulation of protein biosynthesis genes in H-MM was unlikely to be proliferation related.

Of the 48 overexpressed genes involved in protein biosynthesis, only 6 were not located on the common trisomic chromosomes (chromosomes 3, 5, 7, 9, 11, 15, 19, and 21), suggesting that chromosomal copy number amplification is the most common mechanism for up-regulation of these genes. The overexpression of ribosomal protein genes, which constitute a subset of the protein biosynthesis genes, is particularly striking. Twenty-nine of 74 ribosomal proteins genes whose probes were present in the U133A chip were overexpressed in H-MM at a false discovery rate of 1% by the Benjamini and Hochberg method. Sixty ribosomal protein genes were overexpressed if a false discovery rate of 5% was used. Interestingly, most of these overexpressed ribosomal protein genes are located on the commonly trisomic chromosomes. When the ratio of normalized expression of each ribosomal protein gene in

H-MM compared with NH-MM was calculated, these were higher for those genes present on the commonly trisomic chromosomes with a ratio close to 1.5 (Supplementary Fig. S1*B*). These findings suggest the overexpression of ribosomal protein/protein biosynthesis genes in H-MM relative to NH-MM may be primarily due to a gene dosage effect.

Recursive analysis reveals four putative clusters of H-MM. Our recursive analysis identified four putative clusters of H-MM characterized by the expression of largely mutually exclusive gene sets (Fig. 2*A*). When the same 73 probe sets were used to cluster the 63 relapsed H-MM patients entered into bortezomib trials, four clusters were again apparent (Fig. 2*B*). When the 42 best matching probe sets for the U95Av2 chip were used to cluster group 3 patients, the same four clusters can be identified (Fig. 2*C*), suggesting the validity and reproducibility of the four clusters. When the same gene list was used to cluster the 37 NH-MM in the Mayo data set or 13 asymptomatic hyperdiploid tumors (monoclonal gammopathy of undetermined significance and smoldering multiple myeloma), clusters expressing similar set of genes could not be clearly identified (Supplementary Fig. S2), suggesting that these clusters identified by the specific gene expression signatures

Table 1. Gene Ontology analysis of differentially expressed genes between H-MM and NH-MM

Category*	Genes in list in category	P
Gene Ontology biological process		
GO:6412: protein biosynthesis	32	9.26e−10
GO:9059: macromolecule biosynthesis	32	2.62e−08
GO:44249: cellular biosynthesis	38	7.57e−07
GO:9058: biosynthesis	41	1.00e−06
GO:44267: cellular protein metabolism	68	6.14e−06
GO:44260: cellular macromolecule metabolism	68	9.13e−06
GO:19538: protein metabolism	68	1.01e−05
GO:43170: macromolecule metabolism	87	6.90e−05
GO:43037: translation	12	0.000111
GO:42981: regulation of apoptosis	15	0.00485
GO:43067: regulation of programmed cell death	15	0.00498
Gene Ontology molecular function		
GO:3735: structural constituent of ribosome	20	5.90e−11
GO:3723: RNA binding	27	7.11e−07
GO:5198: structural molecule activity	23	0.00644
Gene Ontology cellular component		
GO:5840: ribosome	16	2.28e−08
GO:30529: ribonucleoprotein complex	21	1.33e−07
GO:5737: cytoplasm	87	4.48e−06
GO:43234: protein complex	49	4.58e−06
GO:43229: intracellular organelle	131	0.000469
GO:43226: organelle	131	0.000469
GO:5622: intracellular	150	0.000816
GO:5739: mitochondrion	23	0.00186

*Only GO categories that are significantly associated with the differentially expressed genes at $P < 0.01$ and contain more than 10 of these differentially expressed genes are included in the table.
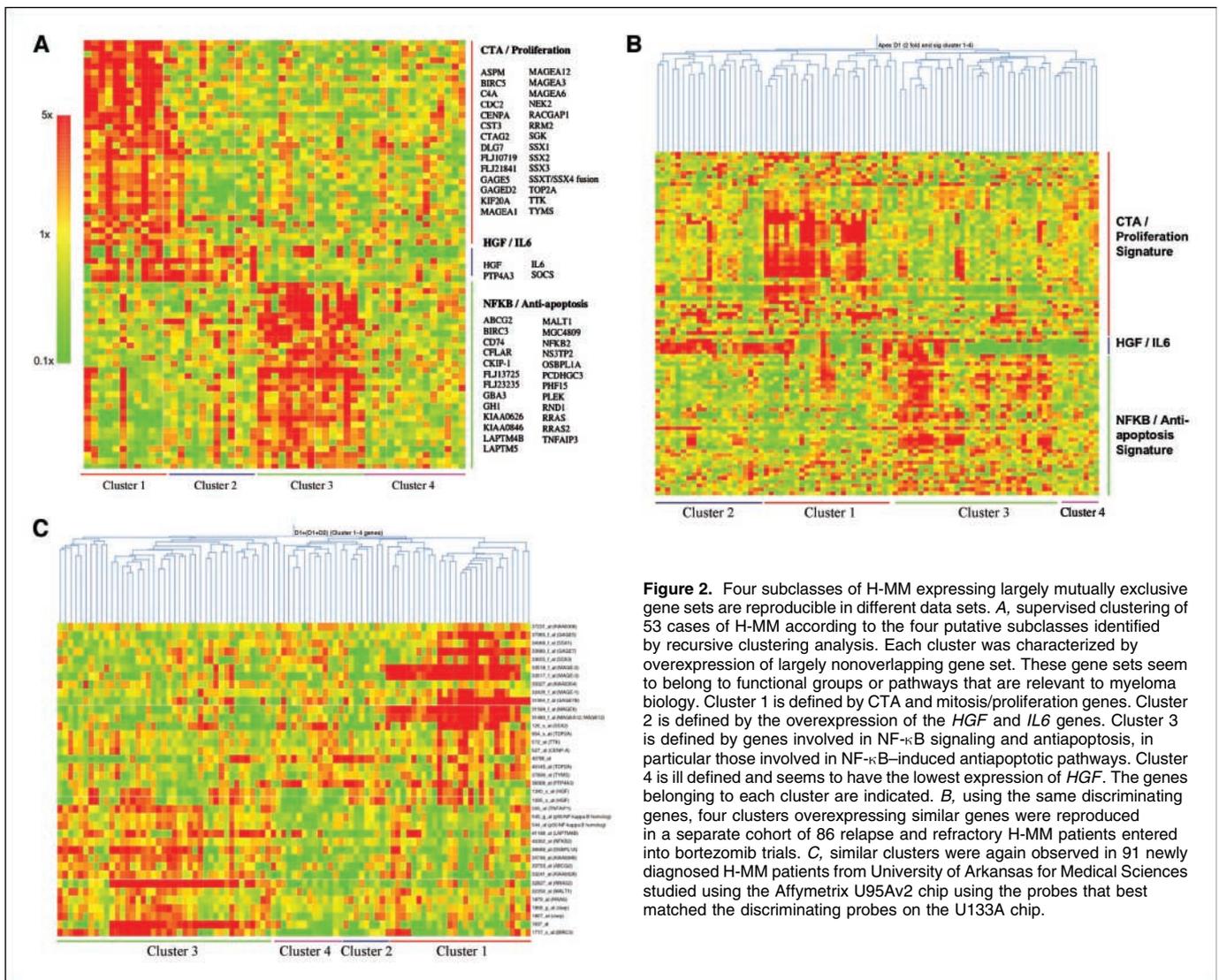
Figure 2. Four subclasses of H-MM expressing largely mutually exclusive gene sets are reproducible in different data sets. *A,* supervised clustering of 53 cases of H-MM according to the four putative subclasses identified by recursive clustering analysis. Each cluster was characterized by overexpression of largely nonoverlapping gene set. These gene sets seem to belong to functional groups or pathways that are relevant to myeloma biology. Cluster 1 is defined by CTA and mitosis/proliferation genes. Cluster 2 is defined by the overexpression of the *HGF* and *IL6* genes. Cluster 3 is defined by genes involved in NF-κB signaling and antiapoptosis, in particular those involved in NF-κB–induced antiapoptotic pathways. Cluster 4 is ill defined and seems to have the lowest expression of *HGF*. The genes belonging to each cluster are indicated. *B,* using the same discriminating genes, four clusters overexpressing similar genes were reproduced in a separate cohort of 86 relapse and refractory H-MM patients entered into bortezomib trials. *C,* similar clusters were again observed in 91 newly diagnosed H-MM patients from University of Arkansas for Medical Sciences studied using the Affymetrix U95Av2 chip using the probes that best matched the discriminating probes on the U133A chip.

seem to be unique to H-MM. Furthermore, a similar recursive analysis of t(11;14) samples identifies subsets defined by genes that are completely different from the genes that define the different H-MM clusters.[8]

Cluster 1 was characterized by high expression of a large group of cancer testis antigen (CTA) genes (*CTAG, SSX, GAGE,* and *MAGE* families) and mitosis/proliferation-related genes (*TOP2A, NEK, ASPM,* and *CENPA*). Cluster 2 was characterized by high expression of hepatocyte growth factor (*HGF*) and interleukin-6 (*IL-6*) genes. Cluster 3 was characterized by high expression of genes involved in NF-κB signaling and antiapoptosis (*NFKB2, CFLAR, BIRC3, CKIP-1, TNFAIP3,* and *RRAS2*). Cluster 4 was less well defined except for lack of high expression of the above genes with low expression of *HGF* in particular. This was not the result of differences in trisomies because there is significant overlap in the composite of trisomies between these two clusters of patients.

Gene set enrichment analysis was done to derived greater biological insights into the different clusters. Cluster 1 was

enriched for genes involved in cell cycle and proliferation as well as genes with E2F binding motifs; cluster 2 was enriched for genes involved in receptor-ligand signaling [e.g., insulin-like growth factor I (*IGF-I*), vascular endothelial growth factor (*VEGF*), and platelet-derived growth factor (*PDGF*)]; cluster 3 was enriched for genes involved in activation of NF-κB pathways [e.g., tumor necrosis factor (TNF) receptor 2, TNFα, NF-κB, and inflammatory response pathways] and regulation of apoptosis (e.g., apoptosis and death pathways), as well as genes with NF-κB binding motif; and cluster 4 was enriched by genes involved in Wnt signaling pathways (e.g., activin receptor–like kinase pathway and Wnt targets; Supplementary Tables S5–S8). Of interest, the enriched gene sets for the different clusters also seem to be mutually exclusive, further suggesting that these different GEP-defined clusters are likely to be biologically distinct.

GEP-identified H-MM subclasses have unique clinical associations. Cluster 1 tumors were more proliferative (median PCLI, 3.8; $P < 0.05$). Cluster 3 consisted of predominantly newly diagnosed cases with none of the patients having International Staging System stage III disease. There was no obvious association between any of the clusters with bone disease and Δ13 (Table 2).

---

[8] Manuscript in preparation.

To assess if these signatures represented transcriptional switches at progression, we compared the prevalence of these signatures in the pooled newly diagnosed (group 1 and group 3 patients) and relapsed cases (group 1 and group 2 patients). The prevalence of all four clusters was similar in newly diagnosed compared with relapsed cases (P = 0.16), suggesting that these profiles do not represent progression signatures.

As cluster 3 was defined by a NF-κB signature, we hypothesized that patients belonging to this cluster should have better response to bortezomib. As predicted, cluster 3 patients had better response to bortezomib compared with patients belonging to other clusters (70% versus 29%; P = 0.02). On the other hand, there was no statistically significant difference in response to dexamethasone between cluster 3 patients and the others (36% versus 50%; P = 0.24). The PFS of cluster 3 patients treated with bortezomib was significantly longer (median PFS, 253 versus 90 days; log-rank P = 0.0002), whereas if they were treated with dexamethasone, the survival is not significantly different (median PFS, 94 versus 168 days; log-rank P = 0.17). When considering only cluster 3 patients, the difference in PFS between bortezomib- and dexamethasone-treated patients was highly significant (median PFS, 253 versus 94 days; log-rank P = 0.0002). Furthermore, the PFS of cluster 3 H-MM patients was almost twice as long as the PFS of all other patients, including the NH-MM patients (median PFS, 253 versus 127 days; log-rank P = 0.13), although this did not reach statistical significance probably due to the relatively small number of cluster 3 patients (Fig. 3).

In addition, the four distinct patient clusters in group 1 had significantly different survival. The median survival is 27 months for cluster 1, not yet reached for cluster 2, 122 months for cluster 3, and 49 months for cluster 4 (Fig. 4).

A recent study showed that H-MM patients with higher prevalence of Δ13 and chromosome 1q amplification and lower prevalence of trisomy 11 have poorer survival (6). Among our GEP-defined clusters, the prevalence of Δ13 and trisomy 11 were similar. However, 1q amplification (assessed by gene expression–derived pseudo-CGH; see Supplementary methods) was significantly more common in cluster 1 (cluster with worst prognosis). It was therefore possible that the poor prognosis H-MM group identified by Carrasco et al. is mainly mediated through association with 1q amplification (and not Δ 13 or lack of trisomy 11) and may correspond to an underlying cluster 1 molecular

signature. The lack of association between Δ13 and the different H-MM clusters with different survival in our data set is consistent with recent findings that Δ13 is not associated with poorer outcome in H-MM (20).

Discussion

Myeloma is composed of roughly equal fractions of hyperdiploid and nonhyperdiploid disease. Whereas the heterogeneity of NH-MM is driven primarily by oncogenic translocations of IgH enhancer sequences, the diversity of H-MM is poorly understood. Although generally associated with a better prognosis compared with NH-MM, the median survival of H-MM is still only ~4 years with no plateau in the survival curve (20). Recent studies have suggested possible clinical and molecular heterogeneity within H-MM (6, 15). However, there has been no comprehensive study that characterizes H-MM in terms of its molecular phenotype and heterogeneity. In this study, we use global GEP to first characterize the molecular signature of H-MM and then to dissect the heterogeneity within H-MM. We found that a protein biosynthesis signature that is primarily driven by copy number changes characterizes H-MM. Furthermore, four biologically relevant subtypes with prognostic implications were identified. More importantly, the results of our study provide insights into the biology of H-MM and the deregulated pathways underpinning these subtypes that can be used to direct therapy.

The protein biosynthesis signature in H-MM is striking, comprising a third of genes differentially expressed between H-MM and NH-MM and constituting most of the genes overexpressed in H-MM relative to NH-MM. Importantly, this signature has now been reproduced in two other studies by different groups using different analysis methods. In a recently published study, in an attempt to find gene signatures that define the different TC groups, a similar signature was found in nine patients with low (but abnormal) CCND1 expression and lacking IgH translocations, which has been shown to correspond to H-MM patients. This signature was reproduced by meta-analysis of two other data sets (21). In another meeting report, a German group, in an attempt to find genes that discriminate between multiple myeloma tumors with and without chromosome 13 deletion, found that protein biosynthesis genes predominate in tumors lacking chromosome

Table 2. Comparison of clinical characteristics across the four GEP-defined subclasses of H-MM

	H-MM GEP clusters				P, χ^2 test
	1	2	3	4	
Relapse, n/N (%)	6/12 (50)	2/12 (17)	1/14 (7)*	7/15 (47)	0.03
Bone disease, n/N (%)	10/11 (91)	8/12 (67)	10/14 (71)	14/15 (93)	0.21
13 Del, n/N (%)	4/12 (33)	6/12 (50)	3/14 (21)	2/15 (13)	0.18
ISS stage III, n/N (%)	4/10 (40)	4/11 (36)	0/14 (0)	5/14 (36)	0.07
PCLI $\geq$ 1, n/N (%)	10/12 (83)*	2/12 (17)*	3/14 (21)	9/15 (60)	0.001
Trisomy 11 by FISH	9/12 (75)	8/12 (67)	13/14 (93)	12/15 (80)	0.4
1q Amp by pCGH	5/12 (42)*	2/12 (17)	0/14 (0)*	2/15 (13)	0.04

Abbreviation: ISS, International Staging System.
*Bonferroni corrected P < 0.05 when compared with others.

Figure 3. Cluster 3 patients have better PFS when treated with bortezomib compared with dexamethasone. PFS of cluster 3 patients were significantly longer than the other H-MM patients when treated with bortezomib [*A*; median PFS, 253 d (*n* = 11) versus 90 d (*n* = 18); log-rank *P* = 0.04] whereas it was not significantly different when treated with dexamethasone [*B*; median PFS, 94 d (*n* = 14) versus 168 d (*n* = 20); log-rank *P* = 017]. *C*, for cluster 3 patients, those randomized to bortezomib had significantly longer PFS than those randomized to dexamethasone [median PFS, 253 d (*n* = 11) versus 94 d (*n* = 14); log-rank *P* = 0.0002]. *D*, furthermore, PFS of cluster 3 H-MM patients (*n* = 11) was double that of all the other patients (*n* = 69), including the nonhyperdiploid patients, treated with bortezomib in this data set (median PFS, 253 versus 127 dy log-rank *P* = 0.13).
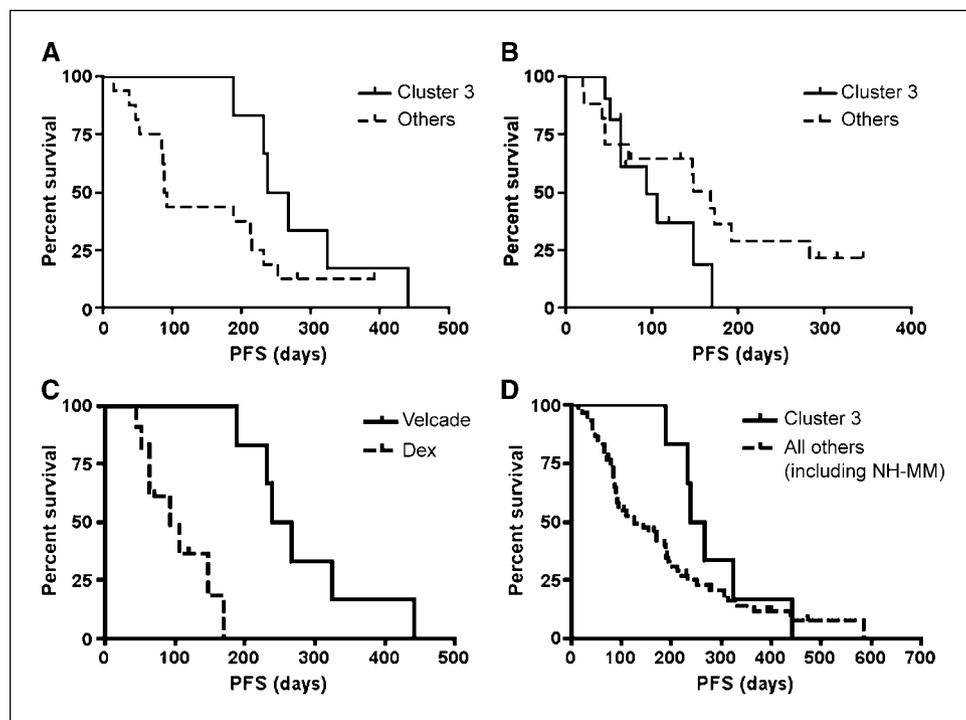
13 deletion, which are mainly H-MM (22). In addition, there is significant overlap between our H-MM signature and the top 50 genes overexpressed and underexpressed in the H-MM cluster in a recent study from the University of Arkansas for Medical Sciences (22 of 50 overexpressed and 19 of 50 underexpressed genes; ref. 15).

Overexpression of protein biosynthesis genes, especially ribosome protein genes, may be the result of rapid cellular proliferation and an indirect consequence of cellular transformation. This is unlikely to be the case in H-MM because these tumors are not more proliferative than NH-MM. Instead, overexpression of these genes seemed to be primarily driven by gene copy number. It has been proposed that cells that contain more ribosomes and translational initiation factors would have an increased translation rate and promote transformation by actually promoting the translation of proteins involved in cellular growth control (e.g., proto-oncogenes) or control of apoptosis (23). Consistent with this view is our network analysis showing that many of the overexpressed genes in H-MM involved in ribosome/protein biosynthesis represented end points of c-myc, NF-κB, and MAPK signaling. Proteomic analysis will provide more insights into the functional consequence of this protein biosynthesis gene expression signature, which may be important for maintenance of the tumor phenotype and survival. As a result, this may represent an avenue for therapeutic intervention of this subtype of multiple myeloma. In our analysis, several components of the mammalian target of rapamycin/P70S6 kinase pathways are deregulated in H-MM compared with NH-MM (*4EBP1, EIF4A, S6K1, RPS6*, and *EEF2*), suggesting that this pathway may be exploited therapeutically with mammalian target of rapamycin inhibitors (24).

Within H-MM, we identified four clusters that are reproducible in two independent cohorts of patients with GEP done on different generation of Affymetrix chips. In addition, these clusters are highly reproducible despite using different patient populations (newly diagnosed and relapsed) and different methods of defining

hyperdiploidy (although these methods are complementary). Furthermore, these clusters are not related to differences in genetic abnormalities or composite of trisomies or to changes relating to progression.

The signature for each cluster provides insights into its biology. Cluster 1 is defined by overexpression of various CTA genes and mitosis/proliferation-related genes. Consistent with the latter, the tumors seem to be more proliferative when assessed by PCLI. Epigenetic mechanisms (hypomethylation) are implicated in the expression of CTAs in tumors (25). It is therefore possible that the expression of these genes is a surrogate for a deregulated epigenetic mechanism that underlies cluster 1 tumors. Of interest, hypomethylation and decondensation of centromeric heterochromatin
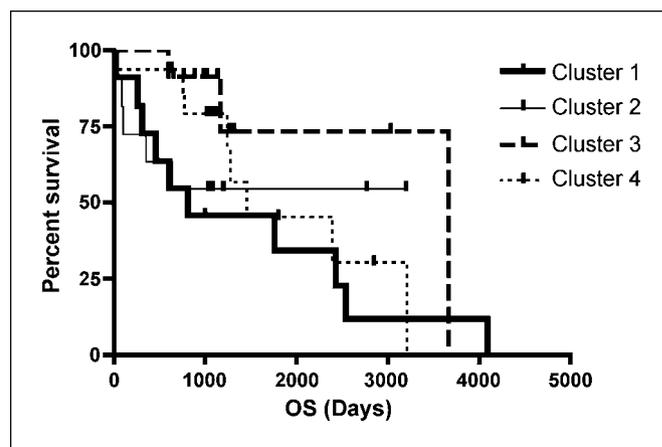

Figure 4. The four GEP-defined clusters of H-MM have differing prognosis. Kaplan-Meier survival curves of the 50 group 1 patients with available survival data, stratified according to their GEP-defined clusters. Median survival of cluster 1 patients (*n* = 11) is 27 mo, whereas median survival for cluster 2 (*n* = 11), cluster 3 (*n* = 13), and cluster 4 (*n* = 15) patients are not yet reached, 122 and 49 mo, respectively. On pairwise comparisons, only survival between clusters 1 and 3 (*P* = 0.04) was significantly different.

have been implicated in the development of 1q amplification in multiple myeloma (26), and 1q amplification is very common in cluster 1 tumors. Indeed, epigenetic silencing of cell cycle inhibitors (27–29) or increased copy number of critical cell cycle gene on chromosome 1q such as *CKS1B* (30) may be responsible for the proliferative phenotype. Cluster 2 is defined by the overexpression of the *HGF* and *IL6* genes. Deregulated HGF signaling induced migration, survival, growth, and invasion of cancer cells via multiple downstream pathways (31). Stimulation of multiple myeloma cells with HGF led to the activation of the Ras/ MAPK and phosphatidylinositol 3-kinase/Akt pathways (32). HGF also induces angiogenesis via VEGF. IL-6 mediates similar effects on cellular phenotype and signals through similar pathways (33), all of which are critical in multiple myeloma. Cluster 3 is defined by the overexpression of genes involved in NF-κB signaling, especially genes involved in NF-κB–induced antiapoptosis. NF-κB signaling is constitutively active in multiple myeloma and is important in mediating many downstream effects of the multiple myeloma-bone marrow interaction (33). One of the important consequences is the up-regulation of antiapoptotic molecules resulting in drug resistance and survival of myeloma cells. Therefore, these signatures may represent divergent pathways in H-MM.

More importantly, these disparate and distinct deregulated pathways provide a useful means for targeting subset of patients with specific therapy. A proof of this principle was provided by the analysis of response and progression of bortezomib-treated patients (group 2) belonging to the different clusters. The NF-κB pathway is one of the key therapeutic targets of bortezomib (34), and, as predicted, cluster 3 patients have better response and PFS when treated with bortezomib compared with dexamethasone, whereas this difference is not observed for the other clusters. Although the number of patients included in this analysis is small, the results are rather striking and are probably real. Our results will have to be confirmed in larger cohorts of H-MM patients in the future.

Many CTAs contain epitopes recognized by CTLs and are potential targets for active immunotherapy (24). Several of these antigens have been identified as potential targets for immunotherapy in multiple myeloma (35, 36). A possible treatment strategy for cluster 1 patients may therefore involve immunotherapy to eradicate residual clone following standard induction with stem cell transplantation. For cluster 2 patients, critical activated pathways (Ras/MAPK and phosphatidylinositol 3-kinase/Akt) could be targeted (37, 38).

The clinical outcome of these clusters is different, adding credence to the biological validity and distinction of these subtypes. Cluster 1 patients have the worst prognosis. Expression of CTA has been associated with poor prognosis in other cancer, and the more proliferative nature of these tumors may contribute to the poor prognosis. However, when all patients with proliferative tumors (PCLI ≥ 1) are considered, cluster 1 patients still have significantly worse survival (data not shown). Importantly, although H-MM patients are generally thought to have good prognosis, a cohort of patients (cluster 1) with survival of only ∼2 years could be identified.

Recently, Zhan et al. (15) have published an unsupervised molecular classification of myeloma. They identified a hyperdiploid cluster that contained most of the hyperdiploid cases and a proliferation cluster that also contained some hyperdiploid cases. This proliferation cluster (also express some CTAs) is associated with 1q amplification and poor prognosis and, hence, is similar to the cluster 1 identified in the current study. As their study aimed to derive a molecular classification of all multiple myeloma, they did not identify additional subtypes of H-MM. In the current study, we specifically wanted to molecularly dissect H-MM and have identified additional clusters including one with NF-κB signature that respond very well to bortezomib.

In summary, H-MM is characterized by a protein biosynthesis signature. In addition, four subclasses of H-MM with distinct clinical and biological associations could be identified, including a subgroup that have very bad prognosis and a subgroup that response particularly well to bortezomib. The gene signature associated with each subclass provides a molecular basis for the different clinical behavior and further insights into the deregulated pathways in H-MM that could potentially be exploited therapeutically.

References

1. Fonseca R, Debes-Marun CS, Picken EB, et al. The recurrent IgH translocations are highly associated with nonhyperdiploid variant multiple myeloma. Blood 2003; 102:2562–7.
2. Smadja NV, Leroux D, Soulier J, et al. Further cytogenetic characterization of multiple myeloma confirms that 14q32 translocations are a very rare event in hyperdiploid cases. Genes Chromosomes Cancer 2003; 38:234–9.
3. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. J Clin Oncol 2005;23:6333–8.
4. Smadja NV, Bastard C, Brigaudeau C, Leroux D, Fruchart C. Hypodiploidy is a major prognostic factor in multiple myeloma. Blood 2001;98:2229–38.
5. Chng WJ, Santana-Davila R, Van Wier SA, et al. Prognostic factors for hyperdiploid-myeloma: effects of chromosome 13 deletions and IgH translocations. Leukemia 2006;20:807–13.
6. Carrasco DR, Tonon G, Huang Y, et al. High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. Cancer Cell 2006;9:313–25.
7. Ahmann GJ, Jalal SM, Juneau AL, et al. A novel three-color, clone-specific fluorescence *in situ* hybridization procedure for monoclonal gammopathies. Cancer Genet Cytogenet 1998;101:7–11.
8. Fonseca R, Blood E, Rue M, et al. Clinical and biologic implications of recurrent genomic aberrations in myeloma. Blood 2003;101:4569–75.
9. Greipp PR, Lust JA, O'Fallon WM, Katzmann JA, Witzig TE, Kyle RA. Plasma cell labeling index and β2-microglobulin predict survival independent of thymidine kinase and C-reactive protein in multiple myeloma. Blood 1993;81:3382–7.
10. Greipp PR, San Miguel J, Durie BG, et al. International staging system for multiple myeloma. J Clin Oncol 2005;23:3412–20.
11. Mulligan G, Mitsiades C, Bryant B, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood 2006 Dec 21; [Epub ahead of print].
12. Blade J, Samson D, Reece D, et al. Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma Subcommittee of the EBMT. European Group

for Blood and Marrow Transplant. Br J Haematol 1998; 102:1115–23.

13. Bergsagel PL, Kuehl WM, Zhan F, Sawyer J, Barlogie B, Shaughnessy J, Jr. Cyclin D dysregulation: an early and unifying pathogenic event in multiple myeloma. Blood 2005;106:296–303.

14. Chng WJ, Van Wier SA, Ahmann GJ, et al. A validated FISH trisomy index demonstrates the hyperdiploid and nonhyperdiploid dichotomy in MGUS. Blood 2005;106: 2156–61.

15. Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. Blood 2006;108: 2020–8.

16. Abraham RS, Ballman KV, Dispenzieri A, et al. Functional gene expression analysis of clonal plasma cells identifies a unique molecular profile for light chain amyloidosis. Blood 2005;105:794–803.

17. Nikolsky Y, Ekins S, Nikolskaya T, Bugrim A. A novel method for generation of signature networks as biomarkers from complex high throughput data. Toxicol Lett 2005;158:20–9.

18. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.

19. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.

20. Chng WJ, Santana-Davila R, Van Wier SA, et al. Prognostic factors for hyperdiploid-myeloma: effects of chromosome 13 deletions and IgH translocations. Leukemia 2006;20:807–13.

21. Agnelli L, Bicciato S, Mattioli M, et al. Molecular classification of multiple myeloma: a distinct transcriptional profile characterizes patients expressing CCND1 and negative for 14q32 translocations. J Clin Oncol 2005; 23:7296–306.

22. Cremer FW, De Vos J, Hose D, et al. Microarray expression profiling indicates up-regulation of the ribosomal machinery in del(13)-negative clones. Haematologica 2005;90:36.

23. Ruggero D, Pandolfi PP. Does the ribosome translate cancer? Nat Rev Cancer 2003;3:179–92.

24. Holland EC, Sonenberg N, Pandolfi PP, Thomas G. Signaling control of mRNA translation in cancer pathogenesis. Oncogene 2004;23:3138–44.

25. Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ. Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer 2005;5:615–25.

26. Sawyer JR, Tricot G, Mattox S, Jagannath S, Barlogie B. Jumping translocations of chromosome 1q in multiple myeloma: evidence for a mechanism involving decondensation of pericentromeric heterochromatin. Blood 1998;91:1732–41.

27. Chim CS, Fung TK, Liang R. Disruption of INK4/CDK/Rb cell cycle pathway by gene hypermethylation in multiple myeloma and MGUS. Leukemia 2003;17:2533–5.

28. Mateos MV, Garcia-Sanz R, Lopez-Perez R, et al. p16/INK4a gene inactivation by hypermethylation is associated with aggressive variants of monoclonal gammopathies. Hematol J 2001;2:146–9.

29. Ng MH, Chung YF, Lo KW, Wickham NW, Lee JC, Huang DP. Frequent hypermethylation of p16 and p15 genes in multiple myeloma. Blood 1997;89:2500–6.

30. Shaughnessy J. Amplification and over-expression of CKS1B at chromosome band 1q21 is associated with reduced levels of p27Kip1 and an aggressive clinical course in multiple myeloma. Hematology 2005;10 Suppl 1:117–26.

31. Jiang WG, Martin TA, Parr C, Davies G, Matsumoto K, Nakamura T. Hepatocyte growth factor, its receptor, and their potential value in cancer therapies. Crit Rev Oncol Hematol 2005;53:35–69.

32. Derksen PW, de Gorter DJ, Meijer HP, et al. The hepatocyte growth factor/Met pathway controls proliferation and apoptosis in multiple myeloma. Leukemia 2003;17:764–74.

33. Hideshima T, Bergsagel PL, Kuehl WM, Anderson KC. Advances in biology of multiple myeloma: clinical applications. Blood 2004;104:607–18.

34. Adams J. Proteasome inhibition in cancer: development of PS-341. Semin Oncol 2001;28:613–9.

35. Chiriva-Internati M, Wang Z, Salati E, Bumm K, Barlogie B, Lim SH. Sperm protein 17 (Sp17) is a suitable target for immunotherapy of multiple myeloma. Blood 2002;100:961–5.

36. van Rhee F, Szmania SM, Zhan F, et al. NY-ESO-1 is highly expressed in poor-prognosis multiple myeloma and induces spontaneous humoral and cellular immune responses. Blood 2005;105:3939–44.

37. Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: its functions and alterations in human cancer. Apoptosis 2004;9:667–76.

38. Sebolt-Leopold JS, Herrera R. Targeting the mitogen-activated protein kinase cascade to treat cancer. Nat Rev Cancer 2004;4:937–47.

Molecular Dissection of Hyperdiploid Multiple Myeloma by Gene Expression Profiling

Wee J. Chng, Shaji Kumar, Scott VanWier, et al.

Updated version	Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/67/7/2982
Supplementary Material	Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2007/03/30/67.7.2982.DC1

Cited articles	This article cites 37 articles, 18 of which you can access for free at: http://cancerres.aacrjournals.org/content/67/7/2982.full#ref-list-1
Citing articles	This article has been cited by 48 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/67/7/2982.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/67/7/2982. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.