

Gene-expression signature of benign monoclonal gammopathy evident in multiple myeloma is linked to good prognosis

Fenghuang Zhan,¹ Bart Barlogie,¹ Varant Arzoumanian,¹ Yongsheng Huang,¹ David R. Williams,¹ Klaus Hollmig,¹ Mauricio Pineda-Roman,¹ Guido Tricot,¹ Frits van Rhee,¹ Maurizio Zangari,¹ Madhav Dhodapkar,² and John D. Shaughnessy Jr¹

¹Donna D. and Donald M. Lambert Laboratory of Myeloma Genetics, Myeloma Institute for Research and Therapy, University of Arkansas for Medical Sciences, Little Rock; ²Laboratory of Tumor Immunology and Immunotherapy, Rockefeller University, New York, NY

Monoclonal gammopathy of undetermined significance (MGUS) can progress to multiple myeloma (MM). Although these diseases share many of the same genetic features, it is still unclear whether global gene-expression profiling might identify prior genomic signatures that distinguish them. Through significance analysis of microarrays, 52 genes involved in important pathways related to cancer were differentially expressed in the plasma cells of healthy subjects (normal plasma-cell [NPC]; n = 22) and patients with strin-

gently defined MGUS/smoldering MM (n = 24) and symptomatic MM (n = 351) (P < .001). Unsupervised hierarchical clustering of 351 patients with MM, 44 with MGUS (24 + 20), and 16 with MM from MGUS created 2 major cluster branches, one containing 82% of the MGUS patients and the other containing 28% of the MM patients, termed MGUS-like MM (MGUS-L MM). Using the same clustering approach on an independent cohort of 214 patients with MM, 27% were found to be MGUS-L. This molecular

signature, despite its association with a lower incidence of complete remission (P = .006), was associated with low-risk clinical and molecular features and superior survival (P < .01). The MGUS-L signature was also seen in plasma cells from 15 of 20 patients surviving more than 10 years after autotransplantation. These data provide insight into the molecular mechanisms of plasma-cell dyscrasias. (Blood. 2007;109:1692-1700)

© 2007 by The American Society of Hematology

Introduction

Multiple myeloma (MM) is a prototypical clonal B-cell malignancy with a terminally differentiated plasma-cell (PC) phenotype. According to longitudinal follow-up of residents of Olmsted County, Minnesota, the annual rate of progression to MM from monoclonal gammopathy of undetermined significance (MGUS) is 1%.¹ This clinically benign condition, with distinct neoplastic features such as aneuploidy, increases in frequency with advancing age and reaches 5.3% in persons 70 years of age and older.² Although the median patient age is approximately 70 years, MM has been diagnosed in teenagers with clinical features and clinical courses resembling those of elderly patients. Family clusters of MM have also been reported,³ but exposure to chemical and physical carcinogens is generally considered etiologic in MM;⁴ HHV8 infection of dendritic cells could not be confirmed.^{5,6} A long latency phase of 15 to 20 years passed before an increase in MM incidence was documented in Japan as a result of the nuclear fallout from the atomic bombs.⁷ Thus, it is plausible to assume that younger patients are likely to develop MM acutely, whereas a smoldering clinical course has been frequently documented to precede the onset of symptomatic MM in the elderly. Smoldering MM (SMM) can be considered an advanced phase of MGUS; even at the time of progression, SMM-evolved MM usually lacks osteolytic lesions or other cardinal features of symptomatic MM. MM remains hypoproliferative, with a long lifespan of malignant B cells, that

assumes high-grade proliferative features only in the terminal phase, from which all human MM cell lines have been derived.^{8,9}

Most genetic lesions typical of MM are already present at the MGUS stage.¹⁰⁻¹² Although these genetic abnormalities can be detected only in interphase cells in MGUS, their detection by metaphase karyotyping in one third of patients with MM reflects increased mitotic activity (possibly reflecting the ability of cells to proliferate outside the confines of the bone marrow milieu) and confers an adverse prognosis.¹³

Although gene-expression profiling (GEP) of CD138-selected cells from bone marrow aspirates was able to distinguish between NPC and MM, it was difficult to distinguish between MM and MGUS.¹⁴⁻¹⁶ GEP studies with whole genome microarrays on larger cohorts of healthy subjects and patients with MGUS, SMM, and MM permitted the discovery of genes that were differentially expressed in comparisons of NPC and of MGUS and MM. Unsupervised hierarchical clustering led to the identification of MGUS with features of MM and of MM with features of MGUS.

Patients, materials, and methods

International Myeloma Working Group criteria were used to classify patients as having MGUS, SMM, or symptomatic MM. For a diagnosis of

Submitted July 31, 2006; accepted September 23, 2006. Prepublished online as *Blood* First Edition Paper, October 5, 2006; DOI 10.1182/blood-2006-07-037077.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

The online version of this article contains a data supplement.

© 2007 by The American Society of Hematology

MGUS, levels of monoclonal protein could not exceed 30 g/L, and bone marrow infiltration with plasma cells could not be less than 10%. Also excluded for this diagnosis was any evidence of related organ or tissue impairment (ROTI) defined as hypercalcemia, renal impairment, anemia, or bone lesions attributed to plasma-cell proliferation. For SMM, ROTI had to be absent, but levels of bone marrow plasmacytosis could exceed 10% and monoclonal protein levels could be greater than 30 g/L.

The analysis described here made use of samples from 22 healthy donors, patients with MGUS (n = 44), SMM (n = 12), or MM with MGUS history (referred to as MM from MGUS [n = 16]), and 351 patients with newly diagnosed MM subsequently treated with total therapy 2 (TT2), a tandem transplantation trial for symptomatic or progressive MM.¹⁷ Test sets consisted of 214 patients with MM enrolled in TT3 and 20 patients surviving more than 10 years after treatment with TT1.¹⁸ Table 1 lists laboratory parameters for patients with MGUS and SMM (at diagnosis or progression to MM) and for patients with MM (before the initiation of therapy). For patients with MGUS/SMM, data were also retrieved from records of the referring institution. Thirty-two patients in the MGUS/SMM group had been enrolled in a prospective observational Southwest Oncology Group study (SWOG 0210) that called for clinical staging with bone marrow biopsy, skeletal survey, MRI at initial registration, and follow-up at 3- to 6-month intervals.

Samples were broken into groups. Group A (19 MGUS and 5 SMM) was composed of patients with documented stable disease parameters for at least 2.5 years (median, 4.3 years; mean, 5.5 years; range, 2.5-14.5 years). SMM patients included in group A had less than 20% PCs at their latest follow-up. For group B (25 MGUS and 7 SMM), the most recent follow-up was less than 2.5 years (median, 1.5 years; mean, 2.0 years; range, 0-7.3 years). Group C was composed of 16 patients with MM from MGUS (12 MGUS and 4 SMM); criteria were relevant for at least 1 year (median, 4.5 years; mean, 6.2 years; range, 1.1-19 years). In keeping with institutional and federal policies, written informed consent was obtained before bone marrow aspirate removal from the posterior iliac crest under local anesthesia.

The institutional review board of the University of Arkansas for Medical Sciences (Little Rock, AR) approved these studies.

Sample processing and molecular analyses

Plasma-cell isolation, total RNA extraction, cRNA synthesis, and hybridization to microarrays (U133 Plus 2.0; Affymetrix, Santa Clara, CA) were performed as described previously.¹⁹ Differences between MGUS (n = 24; group A) and NPC (n = 22) were determined by first filtering out all genes with an absent detection call in more than half these samples or by a χ^2 value of more than 3.84. Significance analysis of microarray (SAM),²⁰ with

a false discovery rate (FDR) of 1%, was then applied to 9935 probe sets. A total of 2864 probe sets was differentially expressed between the 2 groups. Genes making up a myeloid-cell or a normal plasma-cell contamination signature were further subtracted. The contamination signature, including 5351 probe sets, was defined by the comparison of 95 MM contaminated by myeloid cells or normal plasma cells to 256 MM without contamination (SAM FDR less than 1%).¹⁹ This led to the identification of 2181 probe sets, with 1736 overexpressed and 444 underexpressed in MGUS compared with normal plasma cells (Table S1, available on the *Blood* website; see the Supplemental Tables link at the top of the online article). By using the same strategy, 458 genes with 161 overexpressed and 297 underexpressed in MGUS were found differentially expressed in a comparison between MGUS (n = 24; group A) and MM (n = 351) (Table S2). With the use of SAM intersect analysis, 52 genes were found to be significantly differentially expressed in both comparisons. Unsupervised hierarchical cluster analysis²¹ and supervised colorgram analysis²² were used on log₂-transformed signal intensity values of the 52 SAM-defined genes. We recently reported on a validated molecular classification of MM into 7 disease subtypes based on commonalities of gene-expression signatures.¹⁹ In this classification scheme, “spiked” expression of *MMSET* (MS), *MAF/MAFB* (MF), and proliferation (PR) signatures together constituted high-risk disease, whereas hyperdiploidy (HY), low bone disease (LB), and *CCND1/CCND3* translocations (*CCND1/CCND3* also spike with *CD20/MS4A1* and *VPREB3* expression) represented low-risk MM.¹⁹ All patients in this study were also classified into 7 molecular subgroups based on a recently described 700-gene model.¹⁹ The classification of the training and test sets has been described.¹⁹ Approximately one third of CD138-enriched cells from patients with newly diagnosed disease in training and test sets and nearly all MGUS and SMM patients had myeloid-cell gene-expression signatures attributable to contamination of the selected fraction with cells of this lineage.¹⁹ Such patients were not included in a previous molecular classification of MM.¹⁹ No significant difference was observed in the proportions of MM patients with myeloid signature in MGUS-L and non-MGUS-L MM of the training set (24% vs 28%; *P* = .56), suggesting that the MGUS-L designation was not an artifact of the cell purification procedure. Because of the presence of strong myeloid signatures in the PCs of MGUS and SMM, these patients could not be accurately classified with the 7-subgroup model applied to MM. However, as with most patients with MM regardless of myeloid signature, translocation spikes can be observed. Thus, of the 56 patients with MGUS/SMM, 12 (21%) had *CCND1* spikes, 2 (3%) had *CCND3* spikes, 2 (3%) had *MAF* spikes, 4 had *MAFB* spikes, and none had *MMSET/FGFR3* spikes. When the 7-subgroup model was applied to the MGUS/SMM patients, all samples with *CCND1* or *CCND3* spikes were classified as the CD-2 subtype. Moreover, all these patients expressed *CD20 (MS4A1)* and *VPREB3* (data not shown), as previously reported for

Table 1. Characteristics of patients with MGUS, SMM, MM from MGUS, TT1, TT2, and TT3 at diagnosis

Characteristics*	MGUS	SMM	MM from MGUS	TT1 MM	TT2 MM	TT3 MM	P
N	44	12	16	20	351	214	—
Age 65 and older	36	25	63	5	23	30	.001
Female sex	45	58	25	45	43	35	NS
White race	84	92	25	80	89	89	< .001
IgA isotype	12	25	19	15	25	20	NS
B2M ≥ 3.0 mg/L	11	17	64	45	49	57	< .001*
C-reactive protein ≥ 8.0 mg/L	21	25	25	20	36	32	NS
Creatinine ≥ 2.0 mg/dL	2	0	7	5	11	8	NS
Lactate dehydrogenase ≥ 210 IU/L	5	17	15	5	22	19	.041
Albumin < 3.5 g/dL	2	8	6	30	14	14	.049
Hemoglobin < 10 g/dL	2	0	13	25	26	31	< .001
Cytogenetic abnormalities	0	0	6	20	34	31	< .001
Plasma cells (aspirate) > 10%	0	100	46	95	91	90	< .001
MR1 ≥ 1	0	25	25	62	72	73	< .001

P values were determined by Fisher exact test or χ^2 analysis. Values represent the percentages of patients with the specified variable. — indicates not applicable.

*Conversion factors for SI units: multiply B2M values by 85 for conversion to nM; multiply creatinine values by 88.4 for conversion to μ M; multiply albumin values by 10 for conversion to g/L; and multiply hemoglobin values by 10 for conversion to g/L.

this subclass.¹⁹ In addition, all samples with *MAF* and *MAFB* spikes were classified as the MF subtype. These data indicate that in spite of myeloid contamination, a gene-expression signature consistent with these translocations could be recognized. For triple-color interphase fluorescence in situ hybridization (FISH) analysis of abnormalities of chromosomes 1q21 and 13q14, we used procedures previously described.²³

Statistical analyses

The Kaplan-Meier Method²⁴ was used to estimate overall survival, with group comparisons made using the log-rank test.²⁵ Overall survival was defined from the date of registration until death from any cause; survivors were censored at the time of last contact. Univariate and

Table 2. Fifty-two SAM-defined genes are differentially expressed in NPC, MGUS, and MM

Probe set	Symbol	Function	MGUS vs NPC*		MM vs MGUS†	
			SAM score	Fold change	SAM score	Fold change
201486_at	<i>RCN2</i>	Unknown; calcium binding, ER lumen	2.48	1.55	4.13	1.63
202475_at	<i>NIFIE14</i>	Unknown	2.30	1.51	4.00	1.67
212846_at	<i>KIAA0179</i>	Unknown; nucleolar protein	2.07	1.30	3.80	1.48
225260_s_at	<i>MRPL32</i>	Ribosomal protein	2.13	1.32	3.77	1.43
222673_x_at	<i>TMEM57</i>	Unknown; transmembrane protein	3.19	1.65	3.60	1.63
228324_at	<i>C9orf41</i>	Unknown	2.86	1.51	3.59	1.57
225223_at	<i>SMAD5</i>	Gene transcription	3.12	1.70	3.43	1.41
212536_at	<i>ATP11B</i>	Cation transport ATPase	2.16	1.41	3.41	1.55
203216_s_at	<i>MYO6</i>	Recessive actin-based motor	1.98	1.57	3.35	1.91
238761_at	<i>MED28</i>	Gene transcription	3.43	1.80	3.32	1.77
200910_at	<i>CCT3</i>	Molecular chaperone	2.04	1.34	3.30	1.60
231530_s_at	<i>C11orf1</i>	Unknown	2.24	1.57	3.26	1.49
221207_s_at	<i>NBEA</i>	Protein kinase A regulator	3.42	1.83	3.22	2.49
200692_s_at	<i>HSPA9B</i>	Cell proliferation and cellular aging; chaperone.	2.05	1.55	3.18	1.36
212038_s_at	<i>VDAC1</i>	Ion channel for cytochrome c	2.62	1.56	3.17	1.42
208308_s_at	<i>GPI</i>	Energy metabolism	1.97	1.48	3.17	1.48
201013_s_at	<i>PA1CS</i>	DNA synthesis	2.81	1.43	3.17	1.45
202708_s_at	<i>HIST2H2BE</i>	Chromosome organization and biogenesis	2.35	1.71	3.14	2.11
215071_s_at	<i>HIST1H2AC</i>	Chromosome organization and biogenesis	3.47	3.62	3.13	1.97
225361_x_at	<i>LOC159090</i>	Unknown	3.43	1.68	3.11	1.53
219366_at	<i>AVEN</i>	Antiapoptosis	2.63	1.76	3.10	1.47
209398_at	<i>H1ST1H1C</i>	Chromosome organization and biogenesis	5.39	5.53	3.09	2.02
221652_s_at	<i>C12orf11</i>	Unknown; sarcoma antigen NY-SAR-95	2.07	1.37	3.06	1.57
225028_at	<i>LOC550643</i>	Unknown	3.83	2.01	3.03	1.45
214214_s_at	<i>C1QBP</i>	Immunity	2.64	1.45	3.03	1.47
201577_at	<i>NME1</i>	Nucleotide biosynthesis	3.57	1.62	2.99	1.58
218280_x_at	<i>H1ST2H2AA</i>	Chromosome organization and biogenesis	4.24	4.76	2.95	1.92
201479_at	<i>DKC1</i>	Telomere maintenance	2.24	1.42	2.92	1.43
208864_s_at	<i>TXN</i>	Redox reactions	2.90	1.68	2.91	1.51
212297_at	<i>ATP13A3</i>	Cation transport	2.64	1.42	2.88	1.54
222825_at	<i>OTUD6B</i>	Unknown	2.66	1.35	2.88	1.40
209267_s_at	<i>SLC39A8</i>	Ion transport	3.41	1.86	2.88	1.56
217898_at	<i>C15orf24</i>	Unknown	2.35	1.70	2.83	1.33
210275_s_at	<i>ZA20D2</i>	Unknown	2.27	1.36	2.81	1.37
213485_s_at	<i>ABCC10</i>	ATP-dependent efflux pump; multidrug resistance pump	3.17	1.51	2.76	1.34
200994_at	<i>IPO7</i>	Nuclear trafficking	2.08	1.46	2.72	1.37
222428_s_at	<i>LARS</i>	Protein synthesis	3.64	1.68	2.71	1.32
202591_s_at	<i>SSBP1</i>	Mitochondrial DNA replication	2.22	1.42	2.69	1.35
204244_s_at	<i>ASK</i>	Cell cycle	3.25	1.77	2.67	1.37
225916_at	<i>ZNF131</i>	Gene transcription	2.47	1.39	2.63	1.36
202396_at	<i>TCERG1</i>	Gene transcription	3.85	1.62	2.63	1.32
213340_s_at	<i>K1AA0495</i>	Unknown	-2.40	0.78	-2.52	0.75
206150_at	<i>TNFRSF7</i>	Immune cell signaling	-3.82	0.55	-2.56	0.75
228139_at	<i>RIPK3</i>	Cell signaling	-2.63	0.74	-2.59	0.77
220066_at	<i>CARD15</i>	Antiapoptosis	-4.02	0.50	-2.66	0.62
210347_s_at	<i>BCL11A</i>	Gene transcription	3.48	2.04	-2.67	0.64
232511_at	<i>RANBP2L1</i>	Nuclear import	2.07	1.54	-2.71	0.64
214041_x_at	<i>RPL37A</i>	Ribosomal protein	2.22	1.80	-3.09	0.60
219371_s_at	<i>KLF2</i>	Gene transcription	3.81	1.78	-3.11	0.66
202724_s_at	<i>FOXO1A</i>	Gene transcription	2.49	1.55	-3.27	0.64
215671_at	<i>PDE4B</i>	Drug metabolism	2.07	1.53	-3.30	0.59
219675_s_at	<i>UXS1</i>	Glycosaminoglycan biosynthesis	-2.33	0.76	2.97	1.35

Genes are ordered based on the SAM score in the MM vs MGUS comparison.

*Positive SAM score in the MGUS vs NPC column indicates the gene expression is higher in MGUS than in NPC.

†Positive SAM score in the MM vs MGUS column indicates the gene is higher in MM than in MGUS.

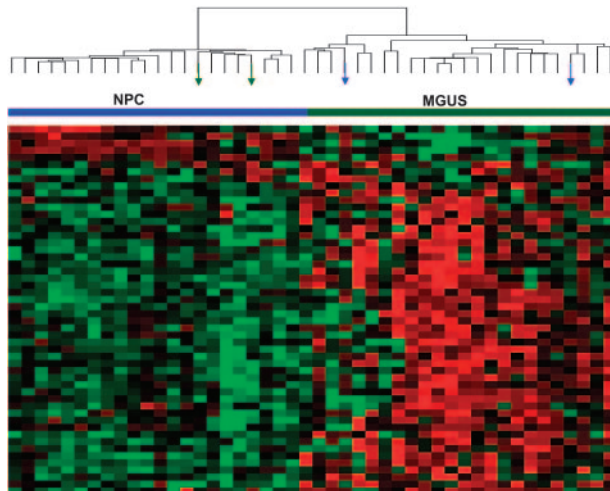


Figure 1. Expression patterns of 52 genes differentially expressed in PCs of healthy donors (NPC) and patients with MGUS. Two-dimensional unsupervised hierarchical cluster analysis of 52 genes (rows) in CD138-enriched plasma cells from 22 healthy donors and 24 MGUS patients (columns). Mean-centered gene expression is depicted by a normalized-signal pseudocolor scale, as described.²¹ Red and green indicate overexpressed and underexpressed genes, respectively. Sample dendrogram (top), reflecting relatedness among samples, consists of 2 major branches defined by overexpressed and underexpressed genes. The left branch consists of 22 NPC samples (horizontal blue bar) and 2 MGUS samples (green arrows), whereas the right branch contains all MGUS (horizontal green bar) and a subset of 2 NPC samples (blue arrows).

multivariate analyses of prognostic factors were performed with Cox regression.²⁶ The cumulative incidence of complete remission was estimated using the method outlined in Gooley et al²⁷ and was compared with the log-rank test.

Gene-expression data

All microarray data have been deposited in the NIH Gene Expression Omnibus (GEO; National Center for Biotechnology Information [NCBI], <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE2658 for 351 patients with MM in the training set and 214 with MM in the test set and under accession number GSE5900 for 22 patients with NPC, 44 with MGUS, and 12 with SMM.

Results

Clinical, laboratory, and molecular features in patients with MGUS, SMM, and MM

As expected, patients with MM had features of greater tumor burden and aggressiveness than subjects with MGUS or SMM (Table 1). Thus, higher proportions of patients in the MM group had elevations of β 2-microglobulin (B2M), C-reactive protein (CRP), lactate dehydrogenase (LDH), creatinine, and bone marrow plasmacytosis and lower levels of hemoglobin and albumin. Focal lesions were absent in all subjects with MGUS but present in 17% of patients with SMM and 80% of patients with MM, 59% of whom had at least 3 focal lesions. Cytogenetic abnormalities were absent in all patients with MGUS and SMM and present in one third of patients with MM. Patients with MM from MGUS tended to be older and, as a group, to have higher hemoglobin levels, fewer cytogenetic abnormalities, lower levels of marrow plasmacytosis, and fewer MRI lesions than patients with MM in whom a history of MGUS/SMM was not documented.

Identifying genes uniquely dysregulated in PCs of MGUS in the context of PCs of healthy subjects and patients with MM

SAM intersection analyses identified 52 genes with differential expression levels across NPC, MGUS, and MM; these were involved in cell-cycle control, DNA synthesis, chromosome assembly, nuclear protein import, gene transcription, cell aging, cell signaling, metabolism, energy production, ion transport, reactive oxygen metabolism, drug resistance, and programmed cell death/apoptosis (Table 2). Of the 52 genes, 41 exhibited a progressive increase in expression levels along the transition from NPC to MGUS to MM, and 4 exhibited a progressive reduction in expression from NPC to MGUS to MM; 6 genes had higher and 1 gene had lower expression levels in MGUS compared with NPC and MM.

The differential expression of the 52 genes in the 22 NPC and 24 MGUS patients was visualized through unsupervised hierarchical clustering (Figure 1): 2 major branches were identified, one containing all but 2 of the MGUS patients and the other containing all but 2 of the NPC samples. When applied to the 56 patients with MGUS, 16 patients with MM from MGUS, and 351 patients with MM in the training group (Figure 2), the sample dendrogram produced 2 major branches, one containing 49 of the 56 (88%) MGUS patients (including 8 of the 12 SMM), 7 of the 16 (44%) MM from MGUS patients, and 99 of the 351 (28%) MM patients; the second branch consisted of 7 of the 56 (12%) MGUS patients (including 4 of the 12 with SMM), 9 of the 16 (56%) with MM from MGUS, and 252 of the 351 (72%) with MM. MM in the first branch was designated MGUS-like (MGUS-L MM), and that in the second branch was termed non-MGUS-like (non-MGUS-L MM); MGUS in the second branch was designated MM-like (MM-L). Three of the 24 (12%) MGUS/SMM in group A and 4 of the 32 (12%) MGUS/SMM in group B were classified as non-MGUS-L.

Supervised visualization of the expression of the 52 genes across the groups described, along with NPC and human myeloma cell lines (MMCL), is provided (Figure 3). NPC and MMCL represent the extremes of benign and malignant PCs, and their PC GEP signatures are consistent with this extreme divergence.

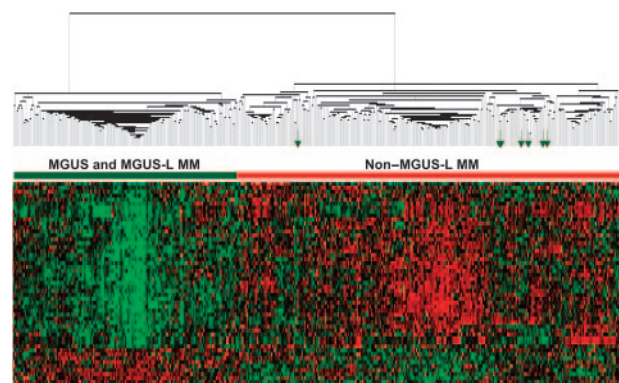


Figure 2. Expression patterns of 52 genes segregate MGUS and MGUS-L MM from non-MGUS-L MM. Two-dimensional unsupervised hierarchical cluster analysis of 52 MGUS genes (rows) in CD138-enriched plasma cells of patients with MGUS ($n = 56$), MM from MGUS ($n = 16$), and newly diagnosed MM ($n = 351$) (columns). The left branch consists of MGUS and MGUS-like MM samples (horizontal green bar), and the right branch contains the non-MGUS-L MM (horizontal red bar). Green arrows represent MGUS patients (MM-L MGUS).

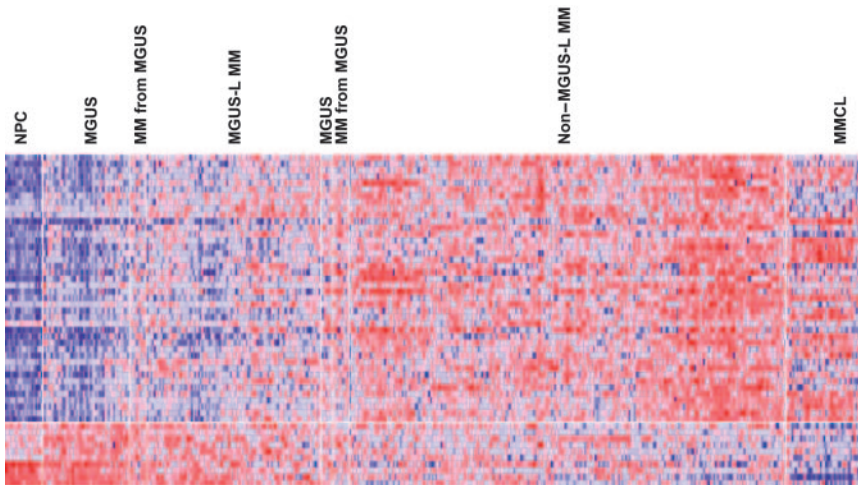


Figure 3. Expression levels of the 52 MGUS genes in PC of healthy donors and patients with MGUS and MM. A colorgram of the expression of the 52 genes in NPC (n = 22), MGUS, MM from MGUS (n = 72), and MM (n = 351) (based on their location in either of the 2 major branches of the dendrogram in Figure 2) and multiple myeloma cell lines (MMCLs) (n = 22). MGUS and MM from MGUS on the left side of the figure represent clusters in the MGUS-L MM branch of Figure 2, and those on the right side represent clusters in the non-MGUS-L MM branch. Genes are indicated along the vertical axis and samples on the horizontal axis. The normalized expression value for each gene is indicated by a color, with red representing high expression and blue representing low expression. Note that NPCs have a distinct pattern of overexpressed and underexpressed genes that progressively inverts with transition to MGUS, MGUS-L MM, non-MGUS-L MM, and finally to MMCL.

Box plots of the expression of select genes are shown in Figure 4. *TNFSF7/CD27*, *KIAA0495*, and *CARD15* genes were representative of those progressively down-regulated, whereas *CCT3*, *VDAC1*, and *DKC1* genes were representative of those progressively up-regulated in the transition from NPC to MMCL. *HIST1H2AC*, *HIST1H2AC*, and *NBEA* were representative of genes showing an increase from NPC to the MGUS-L MM with a significant reduction in expression seen in the non-MGUS-L MM and especially in MMCL.

Relating MGUS-L and non-MGUS-L signatures to previously identified molecular classes of MM

We next determined whether there were differences in the distribution of molecular subgroups of MM in the MGUS-L MM and non-MGUS-L MM patients. Of the 351 patients with MM used in

the current analysis, 95 had been excluded, because of a myeloid expression signature, from the molecular classification schema described previously,¹⁹ thus leaving 76 MGUS-L MM and 180 non-MGUS-L MM classified (Table 3). A PR signature was absent in all MGUS-L MM and present in 29 (16%) of the non-MGUS-L MM ($P < .001$). HY was less frequent in MGUS-L (5% vs 34%, $P < .001$); CD-1 and CD-2 groups were more common in MGUS-L than in non-MGUS-L MM (15% vs 6% [$P = .038$] for CD-1 and 47% vs 4% [$P < .001$] for CD-2).

MGUS-L signature is associated with favorable clinical characteristics and superior survival in spite of lower incidence of complete remission

Compared with non-MGUS-L MM, MGUS-L MM was characterized by lower frequencies of elevated B2M (more than 3 mg/L)

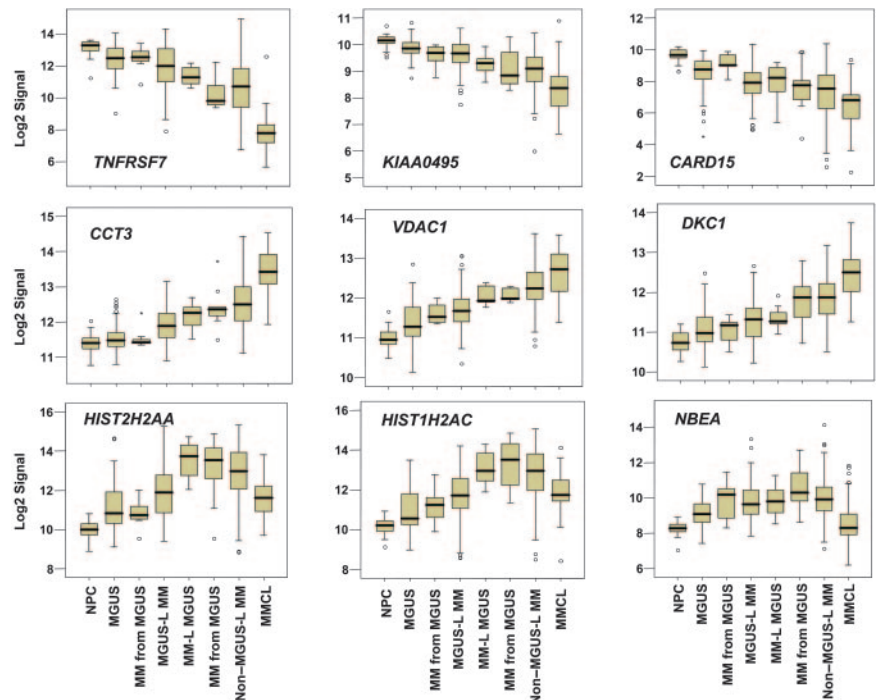


Figure 4. Box plots of expression profiles of genes exhibiting common patterns. Expression levels of select genes exhibiting progressive loss (top 3 panels), progressive increase (middle 3 panels), or increased followed by decreased expression (bottom 3 panels) across the sample groups, as ordered in Figure 2. Sample groups are plotted along the x-axis, and the natural log-transformed Affymetrix-derived signal is plotted on the y-axis. Top, bottom, and middle lines of each box correspond to the 75th percentile (top quartile), 25th percentile (bottom quartile), and 50th percentile (median), respectively. The whiskers extend from the 10th percentile (bottom decile) and top 90th percentile (top decile). Open circles denote outliers within each group. MGUS and MM from MGUS on the left side of the figure represent clusters in the MGUS-L MM branch of Figure 2, and those on the right side represent clusters in the non-MGUS-L MM branch.

Table 3. Molecular subgroup distribution in MGUS-L MM and non-MGUS-L MM in the training set

Type	MGUS-L, %	Non-MGUS-L, %	P
n	74	182	—
PR	0	16	< .001
LB	12	12	NS
MS	12	19	NS
HY	5	34	< .001
CD-1	15	6	.038
CD-2	46	5	< .001
MF	10	8	NS

(33% vs 56%; $P < .001$), cytogenetic abnormalities (17% vs 42%; $P < .001$), high-risk molecular subgroups (PR, MS, MF) (22% vs 43%, $P = .001$), elevated lactate dehydrogenase (LDH) (more than upper limit of normal [ULN]) (12% vs 26%; $P = .005$), and bone marrow plasmacytosis (more than 30%) (56% vs 70%; $P = .001$) (Table 4). Despite a lower frequency of complete and near-complete remission in patients with MGUS-L MM ($P = .006$), these patients had better 5-year survival rates than those with non-MGUS-L MM (76% vs 59%; $P = .009$) (Figure 5A). Interphase FISH–defined gain/amplification of 1q21 (amp1q21) in MM tumor cells is associated with inferior survival.²³ When examined in 253 of the 351 patients, amp1q21 was less frequent in patients with MGUS-L MM than in those with non-MGUS-L MM (35% vs 49%; $P = .04$). The negative prognostic impact of amp1q21 was seen only in the non-MGUS-L MM group, whose 5-year survival rates were 44% in the presence and 73% in the absence of amp1q21 ($P = .001$), similar to those of MGUS-L MM (Figure 5B). Chromosome 13 deletion, tested in 325 patients, was present with similar frequencies in MGUS-L and non-MGUS-L MM (52% vs 49%; $P = .5$) and was not linked to survival in either group (data not shown).

On multivariate analysis, the non-MGUS-L designation was an independent high-risk feature in addition to high-risk molecular subgroup designation, low albumin, high LDH, and presence of focal lesions on MRI examination (Table 5).

MGUS-L MM signature is linked to low-risk clinical and molecular characteristics in a separate test cohort of newly diagnosed MM

When unsupervised hierarchical clustering with the 52 genes was applied to profiles of PC from a separate cohort of 214 patients with newly diagnosed MM enrolled in TT3 and 56 MGUS and 16 MM-from-MGUS patients, 2 major branches in the sample dendrogram were noted in this test set: 45 of 56 (80%) MGUS and 5 of 16 (31%) of MM-from-MGUS patients clustered together with 55 (26%) of 214 MM patients (Figure 6). As with the 28% of MGUS-L MM patients in the training group, the MGUS-L MM group of the test set consisted of fewer patients with elevated B2M (42% vs 63%; $P = .006$), elevated LDH (7% vs 23%; $P = .012$), cytogenetic abnormalities (16% vs 36%; $P = .006$), or high-risk genetic subgroups (19% vs 41%; $P = .018$) (Table 6). None of the MGUS-L MM patients in the test set had a proliferation signature, few belonged to the MMSET group, and CCND1-1 and CCND1-2 designations predominated (Table 7).

Long-term survivors have an MGUS-L signature

Finally, we performed unsupervised hierarchical cluster analysis of CD138-selected PCs of 20 patients with MM surviving more than 10 years after the initiation of TT1,¹⁸ together with the 72 MGUS patients and the 351 patients with newly diagnosed MM in the training set (Figure 7). The sample dendrogram had 2 main branches, one containing all MGUS and MGUS-L MM and 15 of 20 (75%) of the long-term survivors ($P < .001$). Expression spikes for *MAF* and *MAFB* (with *CCND2* overexpression), *CCND1*, and *CCND3* were observed in the TT1 plasma cells (data not shown). The presence of spikes had no influence on whether the sample was defined as g MGUS-L or non-MGUS-L (data not shown).

Discussion

This is the first report on genomic differences recognized by global GEP in a comparative analysis between MM and its precursor conditions, MGUS and SMM. In an earlier report with a smaller number of patients and using a first-generation microarray, normal PCs could be distinguished from PCs of MM and MGUS combined, but PCs of MM and MGUS were indistinguishable.¹⁴⁻¹⁶ The similarity in the transcriptome between MGUS and MM was puzzling because MGUS usually remains clinically benign.^{1,2} By applying more sophisticated data mining approaches to a larger number of samples and a third-generation microarray with more than 54 000 gene features, we identified, in the current investigation, genes with roles in pathways related to cancer that were differentially expressed in PCs of healthy subjects, subjects with MGUS, and patients with MM.

Although MGUS progresses to MM at a low annual frequency of 1%, little is known about the proportion of patients whose MM has evolved from this precursor condition. With the use of unsupervised hierarchical clustering of the 52 genes differentially expressed in MGUS and MM, we identified and validated a subset of MGUS-L MM with favorable clinical features and longer survival; the lower complete response rate may be consistent with the reestablishment of an MGUS condition, assuming its PC to be highly resistant to cytotoxic therapies (similar to normal PCs present at the time of bone

Table 4. Patient characteristics in MGUS-L and non-MGUS-L MM in the training set

Characteristics	MGUS-L, %	Non-MGUS-L, %	P
n	101	250	—
B2M \geq 3.0 mg/L†	33	56	< .001
Cytogenetic abnormalities	17	42	< .001
Subgroups with poor prognosis*	22	43	.001
LDH \geq upper limit of normal	12	26	.005
Plasma-cell aspirate > 30%	56	70	.018

Variables with $P \geq .02$: age, race, sex, isotype, creatinine, hemoglobin level, MRI lesions, C-reactive protein, and albumin.

— indicates not applicable.

*Proliferation/MMSET/MAF.

†Conversion factor for SI units: multiply B2M values by 85 for conversion to nM.

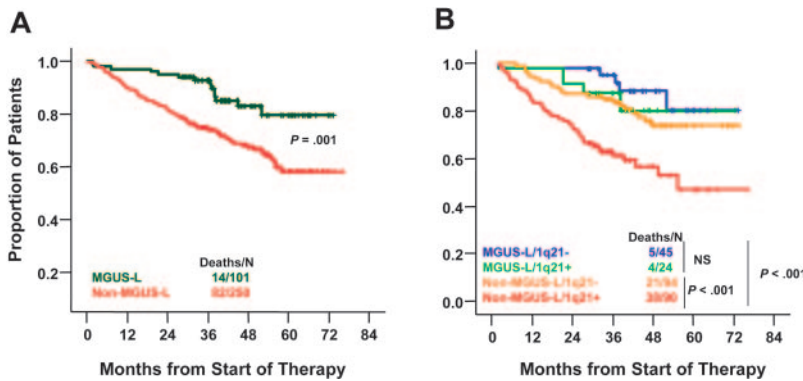


Figure 5. Superior overall survival in MGUS-L MM and non-MGUS-L MM lacking amp1q21. (A) Kaplan-Meier estimates of overall survival in MGUS-L MM and non-MGUS-L MM showed superior 5-year actuarial probabilities of event-free survival (64% vs 44%; $P = .001$) and overall survival (76% vs 59%; $P = .009$) in patients with MGUS signature. (B) Kaplan-Meier estimates of overall survival in MGUS-L and non-MGUS-L MM according to presence of amp1q21 by interphase FISH. Amp1q21 was not a significant adverse parameter in MGUS-L MM, but it identified a group at high risk among patients with non-MGUS-L MM.

marrow aplasia after induction therapy for acute leukemia). These results are supported by the observation of an MGUS-L signature in most patients surviving more than 10 years after the initiation of total therapy.¹⁸ It is important to point out that only 5% of the TT1 long-term survivors were older than 65 years of age at diagnosis.

The difference in the distribution of molecular subgroups between the MGUS-L and non-MGUS-L MM tumors in the training (Table 4) and test (Table 7) sets is perplexing. Although some similarities hold, there was a particularly striking difference between the training and test sets for the HY group and somewhat less so for the MMSET group, perhaps because of the difference in the number of patients in the 2 trials or differences in the percentages of patients within the 7 subgroups enrolled in the 2 trials. Indeed, the test cohort had more HY patients than the training cohort.¹⁹ Larger numbers of patients may be required to resolve this question. One possibility is that it is related to the stage of disease when profiled. Perhaps MM with an MGUS-L signature may eventually acquire a non-MGUS-L signature over time. Once enough patients have been profiled and larger numbers within the 7 molecular classes with MGUS-L and non-MGUS-L designations are identified, Kaplan-Meier survival analyses can be performed to evaluate the interactions between the various classification schemas. Longitudinal studies on a case-by-case basis will reveal insight into the molecular changes accompanying progression in an individual patient.

In addition to identifying a low-risk MM entity with features of MGUS, clustering of MGUS patients with non-MGUS-L

MM patients may relate to a subset of MGUS patients at high risk for conversion to MM. Subjects with MM-L MGUS may benefit from early therapeutic intervention in the absence of symptoms, whereas treatment may be held or deferred in patients with MGUS-L MM.

When examined in the context of a recently published molecular classification of MM,¹⁹ nearly 50% of MGUS-L MM patients were characterized as having CD-2–type disease as opposed to only 4% of patients with non-MGUS-L MM. In contrast, the other molecular class of t(11;14)(q13;q32)–positive disease, termed CD-1 (characterized by CCND1 or CCND3 spikes and lacking CD20 expression) was significantly underrepresented. As expected, the PR class was extremely rare in MGUS-L MM. The infrequent presence of high-risk PR and MS genetic subtypes in MGUS-L MM may explain the superior survival of such patients in comparison with those exhibiting non-MGUS-L MM. The possibility, however, of the eventual acquisition of a non-MGUS-L and PR signature with progression must be considered.¹⁹

Because of technical limitations, it is unclear whether the GEP differences among MGUS, MGUS-L MM, and non-MGUS-L MM are the result of a dilution effect caused by copurification of normal plasma cells or heterogeneity among clonally related tumor cells. Support for the latter possibility comes from the observation that *TNFSF7/CD27*, one of the

Table 5. Multivariate proportional hazards analysis for overall survival in test set (N = 234)

	%	Overall survival		
		HR	95% CI	P
MRI \geq 1 focal lesion	71	3.31	1.64, 6.68	.001
LDH \geq upper limit of normal	21	2.27	1.35, 3.82	.002
Subgroups with poor prognosis*	37	2.13	1.30, 3.48	.003
Albumin < 3.5 g/dL†	16	1.92	1.09, 3.37	.024
Non-MGUS-like	71	2.23	1.08, 4.61	.030

Two hundred thirty-four patients with complete data on all variables were available for the analysis. Only significant variables are shown. Seventy-four patients died.

*Proliferation/MMSET/MAF.

†Conversion factor for SI units: multiply albumin values by 10 for conversion to g/L.

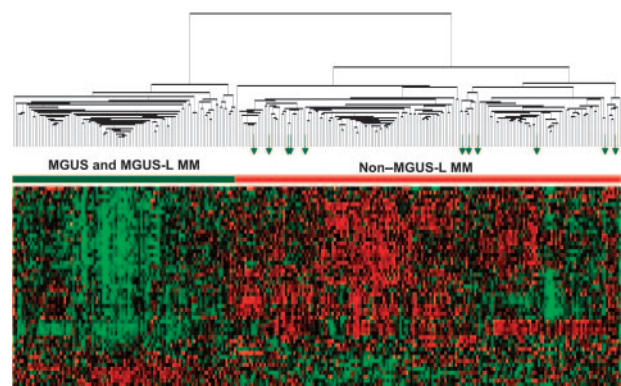


Figure 6. MGUS-L signature is discernible in a test cohort of patients with newly diagnosed MM enrolled in TT3. As in Figure 2, a 2-dimensional unsupervised hierarchical cluster analysis of 52 genes (rows) in CD138-enriched plasma cells from MGUS (n = 56), MM from MGUS (n = 16), and newly diagnosed MM (n = 214). Green arrows represent MGUS clusters with so-called non-MGUS-L MM.

Table 6. Characteristics of patients with MGUS-L and non-MGUS-L MM in the test set

Characteristics	MGUS-L, %	Non-MGUS-L, %	P
n	55	159	—
B2M ≥ 3.0 mg/L*	42	63	.006
Cytogenetic abnormalities	16	36	.006
LDH ≥ upper limit of normal	7	23	.012
Subgroups with poor prognosis†	19	41	.018

Variables with P ≥ .02: age, race, sex, isotype, creatinine, hemoglobin level, MRI lesions, C-reactive protein, and albumin.

— indicates not applicable.

*Conversion factor for SI units: multiply B2M values by 85 for conversion to nM.

†Proliferation/MMSET/MAF.

genes in the current list of 52, is progressively down-regulated in the transition from NPC to MGUS to MM.²⁸ With the use of flow cytometry, Moreau et al²⁹ noted CD27 expression in PCs of all healthy donors, its absence in 36% of patients with MM at diagnosis and in 47% at relapse, and in 92% of human myeloma cell lines; survival was superior in CD27⁺ compared with CD27⁻ MM. The agreement between the GEP data presented here and previously published protein expression studies supports our contention that the observed differences in PC among healthy donors, patients with MGUS, and patients with MM are specific to the disease process rather than a reflection of normal PC contamination.

We recently reported on a high-resolution map of recurrent, minimal common regions (MCRs) of gain/amplification and loss/deletion in newly diagnosed MM along with the genes residing in these MCRs whose expression was strongly correlated with changes in copy number.³⁰ Of the genes identified in this study, the overexpression of *CCT3* and *HIST2H2AA*, and the mapping to MCRs at 1q21-1q22, in MM relative to MGUS was observed. Consistent with this, we recently showed that FISH-defined amplification of 1q21 was absent in MGUS; its presence in some patients with SMM was associated with a higher risk for conversion to MM,^{23,31} and its presence in MM conferred short survival.²³ Other genes mapping to MCRs of gain/amplification included *SLC39A8* mapping to 4q22.3-4q24 and *ASK/DBF4* mapping to 7q21.12. Genes mapping to MCRs of loss/deletion and exhibiting reduced expression in MM relative to MGUS included the caspase recruitment domain-containing protein 15 (*CARD15*) mapping at 16q11.2 and the fork-head box O1A (*FOXO1A*) transcription factor mapping near the peak of an MCR at 13q14.1. The identification of genes whose expression level is copy number sensitive in MM as differentially expressed in a comparison of MGUS with MM again suggests that

Table 7. Comparison of the distribution of molecular subgroups in MGUS-L MM and non-MGUS-L MM of test set

Test set	MGUS-L, %	Non-MGUS-L, %	P
N	32	126	—
Proliferation	0	14	< .001
Low bone disease	13	18	NS
MMSET	3	18	< .001
Hyperdiploidy	28	33	NS
CCND1-1	13	2	.003
CCND1-2	28	6	< .001
MAF	15	9	NS

— indicates not applicable.

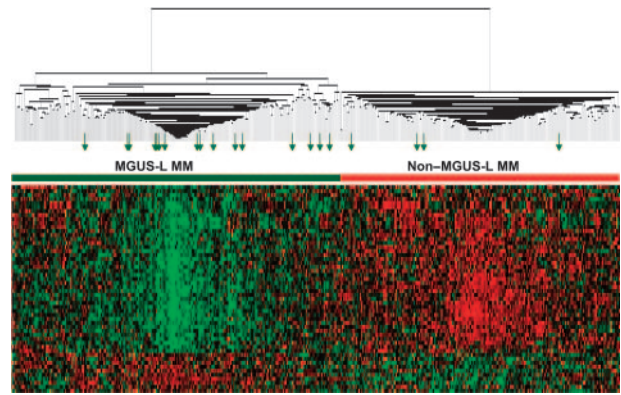


Figure 7. MGUS-L signature is present in most PCs of patients still alive more than 10 years after TT1. As in Figure 2, a 2-dimensional unsupervised hierarchical cluster analysis of 52 genes (rows) in CD138-enriched plasma cells from patients with MGUS (n = 56), MM from MGUS (n = 16), and newly diagnosed MM (n = 351, training cohort) and from 20 long-term survivors (columns). Long-term survivor samples are indicated by green arrows.

differences are not likely to reflect the degree of contamination of normal plasma cells in the CD138-selected fractions, but rather the altered expression of this small subset of genes might be important in disease progression.

In conclusion, we used genomic profiling to identify a subset of genes whose expression patterns differentiated among PCs from healthy donors, patients with MGUS, and patients with MM. Patients with MGUS, exhibiting molecular features of MM and deemed at higher risk for conversion to overt MM, could be selected for secondary prevention trials. The prevalence of a MGUS-L signature in PCs of long-term survivors of TT1 raises the question whether these superior results could have been achieved with less aggressive treatment strategies. Investigation of the functional pathways of genes with differential expression levels in the various plasma cell dyscrasias may provide valuable insight into the enigmatic mechanisms of the multistep molecular pathogenesis of MM.

Acknowledgments

This work was supported by National Institutes of Health grants CA55819 (J.D.S., F.Z., G.T., and B.B.) and CA97513 (J.D.S.), the Fund to Cure Myeloma and Peninsula Community Foundation, and the Southwest Oncology Group.

Authorship

Contribution: F.Z., B.B., and J.D.S. aided in conceptualization of the study and wrote the paper. J.D.S. supervised the work and analyzed the data. F.Z., Y.H., and D.R.W. performed microarray and data analyses. B.B., V.A., K.H., M.P.-R., G.T., F.v.R., M.Z., and M.D. provided clinical data.

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

Correspondence: John D. Shaughnessy, Donna D. and Donald M. Lambert Laboratory of Myeloma Genetics, Myeloma Institute for Research and Therapy, University of Arkansas for Medical Sciences, 4301 W. Markham Street, Little Rock, AR 72205.

References

- Kyle RA, Therneau TM, Rajkumar SV, et al. A long-term study of prognosis in monoclonal gammopathy of undetermined significance. *N Engl J Med*. 2002;346:564-569.
- Kyle RA, Therneau TM, Rajkumar SV, et al. Prevalence of monoclonal gammopathy of undetermined significance. *N Engl J Med*. 2006;354:1362-1369.
- Lynch HT, Watson P, Tarantolo S, et al. Phenotypic heterogeneity in multiple myeloma families. *J Clin Oncol*. 2005;23:685-693.
- Durie BG. The epidemiology of multiple myeloma. *Semin Hematol*. 2001;38:1-5.
- Tilsdale JF, Stewart AK, Dickstein B, et al. Molecular and serological examination of the relationship of human herpesvirus 8 to multiple myeloma: orf 26 sequences in bone marrow stroma are not restricted to myeloma patients and other regions of the genome are not detected. *Blood*. 1998;92:2681-2687.
- Ablashi DV, Chatlynne L, Thomas D, et al. Lack of serologic association of human herpesvirus-8 (KSHV) in patients with monoclonal gammopathy of undetermined significance with and without progression to multiple myeloma. *Blood*. 2000;96:2304-2306.
- Neriishi K, Nakashima E, Suzuki G. Monoclonal gammopathy of undetermined significance in atomic bomb survivors: incidence and transformation to multiple myeloma. *Br J Haematol*. 2003;121:405-410.
- Rajkumar SV. MGUS and smoldering multiple myeloma: update on pathogenesis, natural history, and management. *Hematology (Am Soc Hematol Educ Program)*. 2005;340-345.
- Blade J, Rosinol L. Smoldering multiple myeloma and monoclonal gammopathy of undetermined significance. *Curr Treatment Options Oncol*. 2006;7:237-245.
- Fonseca R, Bailey RJ, Ahmann GJ, et al. Genomic abnormalities in monoclonal gammopathy of undetermined significance. *Blood*. 2002;100:1417-1424.
- Kaufmann H, Ackermann J, Baldia C, et al. Both IGH translocations and chromosome 13q deletions are early events in monoclonal gammopathy of undetermined significance and do not evolve during transition to multiple myeloma. *Leukemia*. 2004;18:1879-1882.
- Chang WJ, Van Wier SA, Ahmann GJ, et al. A validated FISH trisomy index demonstrates the hyperdiploid and non-hyperdiploid dichotomy in MGUS. *Blood*. 2005;106:2156-2161.
- Shaughnessy J, Jacobson J, Sawyer J, et al. Continuous absence of metaphase-defined cytogenetic abnormalities, especially of chromosome 13 and hypodiploidy, ensures long-term survival in multiple myeloma treated with total therapy I: interpretation in the context of global gene expression. *Blood*. 2003;101:3849-3856.
- Zhan, F, Hardin, J, Bumm K, et al. Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance and normal bone marrow plasma cells. *Blood*. 2002;99:1745-1757.
- Davies FE, Dring AM, Li C, et al. Insights into the multistep transformation of MGUS to myeloma using microarray expression analysis. *Blood*. 2003;102:4504-4511.
- Hardin J, Waddell M, Cheng J, et al. Evaluation of multiple models to distinguish closely related forms of disease using DNA microarray data: an application to multiple myeloma. *Stat Appl Genet Mol Biol*. 2004;3:Article10.
- Barlogie B, Tricot G, Anaissie E, et al. Thalidomide and hematopoietic-cell transplantation for multiple myeloma. *N Engl J Med*. 2006;354:1021-1030.
- Barlogie B, Tricot G, Rasmussen E, et al. Total therapy 2 without thalidomide: comparison with total therapy 1: role of intensified induction and post-transplant consolidation therapies. *Blood*. 2006;107:2633-2638.
- Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. *Blood*. 2006;108:2020-2028.
- Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98:5116-5121.
- Eisen M, Spellman P, Brown P, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 1998;95:14863-14888.
- Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;286:531-537.
- Hanamura I, Stewart JP, Huang Y, et al. Frequent gain of chromosome band 1q21 in plasma cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem cell transplantation. *Blood*. 2006;108:1724-1732.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457-481.
- Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep*. 1966;50:163-170.
- Cox DR. Regression tables and life tables. *J R Stat Soc B*. 1972;34:187-202.
- Gooley TA, Leisenring W, Crowley J, et al. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med*. 1999;18:695-706.
- Guikema JE, Hovenga S, Vellenga E, et al. CD27 is heterogeneously expressed in multiple myeloma: low CD27 expression in patients with high-risk disease. *Br J Haematol*. 2003;121:36-43.
- Moreau P, Robillard N, Jego G, et al. Lack of CD27 in myeloma delineates different presentation and outcome. *Br J Haematol*. 2006;132:168-170.
- Carrasco D, Tonon G, Huang Y, et al. High-resolution genomic profiles define distinct clinicopathogenetic subgroups of multiple myeloma patients. *Cancer Cell*. 2006;9:313-325.
- Rosinol L, Carrio A, Blade J, et al. Comparative genomic hybridisation identifies two variants of smoldering multiple myeloma. *Br J Haematol*. 2005;130:729-732.