Data Supplement Text

Patients

MM patients received an induction therapy with four courses of vincristine, adriamycin, and dexamethasone (VAD), followed by double intensive therapy. The IFM99-02 trial was focused on patients with fewer than two adverse prognosis factors Sβ2M levels above 3 mg/L and del(13) by FISH. After induction, patients received two courses of high-dose melphalan (140 mg/m^2 and 200 mg/m^2), and were then randomly assigned for maintenance therapy—none (arm A), pamidronate (arm B), or pamidronate plus thalidomide (arm C)—until relapse. This trial recruited 780 patients. The IFM99-03 trial enrolled 65 patients with two poor prognosis factors and with an HLA-identical familial donor. After induction, patients received one high-dose melphalan course (200 mg/m^2), followed by a reduced intensity-conditioned allogeneic transplantation. The IFM99-04 trial enrolled 219 patients with two poor prognosis factors and no HLA-identical familial donor. After a similar induction and first high-dose melphalan course, patients received a second melphalan-based intensification (220 mg/m^2), and were randomly assigned to receive or not an anti–interleukin-6 antibody during the conditioning regimen. Finally, IFM99-02* denoted patients treated according to IFM99-02 trial without stratification according to Sβ_2M > 3 mg/l and/or del(13) by FISH, and were treated according to arm A without maintenance therapy. The initial and validation cohorts include patients enrolled in the three IFM99 trials and patients of the IFM99-02* subset. The initial cohort comprised 192 patients with MM. Patients were enrolled from December 21, 1999 to February 15, 2005. The median follow-up time was 58.3 months. The validation cohort comprised 273 patients with newly diagnosed myeloma younger than 66 years

of age. Patients were enrolled from September 3, 1998 to June 23, 2006. The median follow-up time was 42.7 months. The overall survivals were similar between initial and validation cohorts (Appendix Fig A11, only online). Overall survival was calculated from the date of diagnosis. Approval for this study was obtained from medical ethical committees and institutional review boards of the University Hospitals of Nantes, Toulouse, and Grenoble. Informed consent was provided according to the Declaration of Helsinki.

Samples and DNA Extraction

Bone marrow specimens were obtained during standard diagnostic procedures in IFM centers and overnight shipped to the Hematology department at University Hospital in Nantes.

Plasma cell purification was performed as previously described.[17] In all cases, purity of CD138+ plasma cells was higher than 90%, median purity 98% (range, 90% to 100%) as assessed by morphology. Purified plasma cells were cryopreserved in DMSO or frozen at –80°C as dry pellet. DNA was extracted using the DNA AllPrep DNA/RNA MiniKit (Qiagen, Valencia, CA). DNA quality and quantity were assessed using the Nanodrop Spectrophotometer (NanoDrop Technologies Inc, Wilmington, DE).

Hybridization of Myeloma Samples onto 500K Arrays Using WGSA

The GeneChip Human Mapping 500K Array Set (Affymetrix, Santa Clara, CA), contain 534,500 SNPs on two arrays (250K Nsp and 250K Sty) and is used in conjunction with the whole genome sampling assay (WGSA). Each array is designed to interrogate SNPs residing on PCR amplicons of 100 to 1,100 bp in length using

either Nsp1 or Sty1 restriction enzyme digested, adaptor-ligated genomic DNA template. Each SNP is interrogated by 24 features, with six features each for the perfect match (PM) probes and the mismatch (MM) probes. DNA was restriction enzyme digested, PCR-amplified, purified, fragmented, labeled and hybridized to the arrays according to the manufacturer's instructions. Briefly, 250 ng of DNA was digested with Nsp1 or Sty1 (New England Biolabs, Boston, MA). Digested DNA was adaptor-ligated and PCR-amplified using Titanium Taq (Clontech, Mountain View, CA) in three 100 µl PCR reactions for each enzyme-digested sample. PCR products from each set of three reactions were pooled, concentrated, purified and 90 µg of purified products were fragmented using 0.25 units of DNase I. Fragmented PCR products were then labeled, denatured, and hybridized to the arrays. Arrays were then washed and stained using Affymetrix fluidics stations, and scanned using the Gene Chip Scanner 3000 7G. CEL files were generated using Affymetrix GeneChip Operating Software v 1.4. Array SNP calls were generated using Affymetrix GeneChip Genotyping Analysis Software version 4.0.

Interphase Cytogenetics Analysis

Plasma cells were then analyzed using DNA probes specific for the following chromosomal aberrations: del(13q14), t(11;14)(q13;q32), t(4;14)(p16;q32), and del(17p13). The del(13) was analyzed with a probe specific for the D13S319 locus (purchased from Abbott, Rungis, France). Probes specific for the t(4;14) and t(11;14) translocations were kindly provided by Abbott (Chicago, IL). In addition, samples from patients in the validation cohort were analyzed for del(12p13.31) using a BAC probe RP11-433J6 at 6 579 345 – 6 635 864 and for amp(5q31.3) using 5q31 probe (purchased from Abbott, Rungis, France). The chromosomal aberrations: del(13q14),

t(11;14)(q13;q32), t(4;14)(p16;q32), and del(17p13) have already been reported in a previous study.[10]

The concordance between 500K SNPs array and FISH copy number estimates was assessed by comparing the same samples for 13q14 deletion (Appendix Fig A2). The 500K SNP array analysis correctly assigned 94% of del(13) patients determined by FISH analysis and 98% of normal diploid chromosome 13 (Appendix Table A3).

SNP Microarray Analysis dChip Software Analysis

We used the dChip software (www.dchip.org) to normalize all the arrays of 192 MM samples and 10 normal blood samples and compute model-based signal values.[19] For each SNP, the average signal of the normal samples was used to determine the signal for the normal copy number 2 and to scale the signals in the rest of the samples to obtain raw copy numbers. A smoothing window of 11 SNPs was then used to obtain median-smoothed copy numbers.

We then scaled the smoothed copy numbers to recover absolute copy numbers. Specifically, the smoothed copy numbers are binned into copy number intervals of 0.05, and the copy number bin with the largest SNP counts is regarded as corresponding to the absolute copy number 2 in cells. The assumption is that the size of the chromosome regions with altered copy numbers is smaller than un-altered chromosome regions with copy number 2. Then this mode copy number is used to scale the smoothed copy numbers by a factor of (2 / mode copy).

A hidden Markov model considering haplotype was used to infer loss of heterozygosity (LOH) from tumor-only samples http://www.ncbi.nlm.nih.gov/pubmed/16699594),[20] using publicly available genotypes of 60 parental samples from the Centre d'Etude du Polymorphisme Humain collection

in the HapMap dataset as control haplotype genotypes. Inferred LOH events with copy numbers between [1.7, 2.3] were regarded as uniparental disomy (UPD).

The genomic profile and recurrence plot were obtained using dChip software with median smoothing window of 11 SNPs, gains were defined as copy number ≥ 2.5 and losses as copy number ≤ 1.5. We used these thresholds to account for possible normal sample contamination, hyperdiploid genome content, and the signal compression effect of the array platforms.

Molecular Cytogenetic Studies

A molecular karyotype was established with the average value of the SNPs intensities in each cytoband. Gain/amplification was defined as copy number ≥ 2.5 and loss/deletion as copy number ≤ 1.5.

We observed three distinct patterns:

1) Gain or loss of chromosome; if at least cytobands flanking the centromere are gained or lost. In a few cases visual inspection of dChipSNP median-smoothed $\log_2$ ratio plot was used to confirm gain or loss of the centromere region.
2) Gain or loss of whole arm; if all cytobands upstream or downstream centromere are altered.
3) Interstitial gain or loss; if at least one cytoband is altered in a chromosome.

Copy-number abnormalities encompassing these three patterns of the 192 MM cases are shown in the Table A4.

Hyperdiploid status (47 chromosomes or more) was calculated from the number of gained or lost chromosomes (pattern 1). Given that 500K SNP arrays do not have

markers on Y chromosome, the number of chromosomes in male cases is evaluated +/- 1; the 5 male cases with 46 of 47 chromosomes were not considered in this study. Another case was removed because uncertainties on several chromosomes. In seven cases, the copy number of one chromosome was not clearly assessed, but even with the smallest value they remained hyperdiploid (Table A4).

Segmentation Analysis

We applied the circular binary segmentation algorithm to the above $\log_2$ ratios data to identify copy number alterations for each sample.[21] This algorithm recursively splits chromosomes into subsegments based on a maximum t statistic. To decide whether or not to split at each stage, we performed 10,000 permutations to obtain a reference distribution to estimate the significance of the maximum t statistic. The segmentations were performed using a p-value threshold of .01 (significance level $\alpha = .01$). We used version 1.12.0 of the circular binary segmentation algorithm implemented in the DNAcopy package of Bioconductor :

http://www.bioconductor.org/packages/2.2/bioc/html/DNAcopy.html.

Segment Size Controlling

We developed a procedure to assemble short segments (≤ 25 SNPs): (1), a short segment is surrounded by large segments. We proceed to a comparison of both flanking segments: if $P > .05$ (Wilcoxon test) then the segments are fused (creation of a novel segment with a new segment mean) if not, the small segment is removed. (2), a short segment follows a short segment, both segments are fused and the above procedure is repeated.

Finding Minimum Common Genomic Deletion and Amplification in Individual Chromosome

In order to define recurrent altered segment per chromosome we developed an algorithm described as follows:

1) Filtration of segments to obtain candidate genomic lesions (deletion or amplification) by using mean $\log_2$ ratios of the segments obtained as described above with values > 0.2 or < -0.4 per chromosome and per patient; a $\log_2$ ratio is defined as $\log_2$ (SNP's copy number/2). Cutoff values were adjusted for male patients.
2) Identification of the patient with an altered segment (S) most upstream on a chromosome
3) Search among the other patients for the patients with altered segment (N) overlapping partially or totally S segment
4) Search for the most upstream N segment from step 3
5) Selection of intersections between S and N segments, generating minimum common genetic lesions (MCGL)
6) Search for MCGL across all patients
7) Collection of MCGL characteristics (SNP number, occurrence, genomic annotations) then filtration according to SNP number (≥ 25) and to frequency greater than 5%.
8) Filtration of intra-chromosomal MCGL associated with survival by univariate Cox analyses. Cutoff *P* values < .001 with the exception of segments in chromosomes 5, 9, and 12 (*P* < .0001) (Table A7).

Gene Expression Data Generated for the Genes of Interest

Expression profiles of the genes of interest including: the 15 genes of the prognostic signature[26] (Table A17) and the 85 genes residing in three DNA segments at chr1q21-q23, chr5q31, and chr12p12-13 (Table A14, A15, and A16) were obtained for 170 samples of the initial cohort by using Affymetrix, Exon 1.0 St Array. The microarray data analyzed in this study have been deposited in the NIH Gene Expression Omnibus data base at www.ncbi.nlm.nih.gov/geo under the following accession number GSE12896.

Performance of CNA-Based Model When Adjusted for 15-Gene Model Integrating SNP Array-Based Prognostic Segments With Gene Expression Signatures

The expression levels of the 15-gene signature deduced from a data set generated from Exon 1.0 ST Array (Affymetrix) of 170 samples within the initial cohort were used to calculate an expression risk score for each patient (Table A17). Then the cases were ranked according to their score, the quartile 4 delineated high-risk patients. Finally, univariate and multivariate analyses of expression-based and CNA-based risk models were applied to the initial cohort.

Statistical Analysis

SAS System version 9.1 (SAS Institute Inc, Cary, NC), R software (version 2.5.1) and BRB ArrayTools developed by Dr. R. Simon and A. Peng (Bethesda, MD; 2003) were used to perform statistical analyses.

IFM Members:

Dr A. Abdelkefi, Prof J.F. Abgrall, Dr X. Agape, Dr C. Alexandre, Dr M. Alexis, Dr A. Amar, Dr M. Andre, Dr B. Anglaret, Dr Y. Arkam, Dr P. Arnaud, Dr S. Assadourian, Dr

D. Assouline, Prof M. Attal, Dr B. Audhuy, Prof H. Avet-Loiseau, Dr M. Azagury, Dr I. Azais, Prof F.R. Bataille, Dr E. Baumelou, Prof F. Bauters, Dr M. Beaumont, Dr A. Bellange, Dr L. Benboubker, Prof C. Berthou, Dr M.A. Bichoffe, Dr P. Bindi, Dr P. Biron, Dr P. Bironne, Prof D. Blaise, Dr M. Blanc, Dr J.Y. Blay, Dr F. Bonnevie, Prof A. Bosly, Dr K. Bouadballah, Dr R. Bouadballah, Dr J.M. Boulet, Dr J.H. Bourhis, Prof D. Bouscary, Dr G. Bouteiller, Dr A. Brion, Dr S. Cailleres, Dr D. Caillot, Dr J. Camo, Dr L. Campos, Prof P. Casassus, Dr M.T. Caulier, Dr J. Ceccaldi, Dr C. Chaleteix, Dr C. Charbonnel, Dr C. Chatelain, Dr N. Cheron, Dr B. Christian, Dr V. Coiteux, Dr P. Collet, Prof P. Colombat, Dr B. Corront, Prof R. Costello, Dr B. Couret, Dr F. Da Silva, Dr C. Dauriac, Dr De Bock, Dr P. De Jaureguiberry, Dr B. De Prijck, Dr T. De Revel, Dr D. Decaudin , Dr O. Decaux, Dr A. Delannoy, Dr N. Denizon, Dr B. Depernet, Dr I. Dervite, Prof B. Desablens, Dr L. Detourmignies, Dr M. Dib, Dr J.F. Dor, Dr V. Dorvaux, Prof C. Doyen, Prof F. Dreyfus, Prof C. Dumontet, Dr B. Dupriez, Dr J.L. Dutel, Dr M. Eddahbi, Dr H. Eghbali, Dr J.C. Eisenmann, Dr M. Ertault, Dr J.P. Eschard, Dr M. Escofre-Barbe, Prof L. Euller-Ziegler, Prof T. Facon, Dr O. Fain, Prof W. Feremans, Prof A. Ferrant, Dr E. Fleck, Dr M. Flesch, Dr J. Fleury, Dr C. Fohrer, Dr A. Fouilhoux, Dr C. Foussard, Dr J. Frenkiel, Dr C. Fruchart, Prof J.G. Fuzibet, Dr F. Garban, Dr M. Gardembas-Pain, Dr L. Garderet, Dr R. Garidi, Prof J. Gastaut, Dr C. Ghandour, Dr K. Ghomari, Dr B. Gilson, Prof C. Gisselbrecht, Prof N.C. Gorin, Prof B. Grosbois, Prof F. Guilhot, Dr G. Guillerm, Dr H. Guy, Prof D. Guyotat, Dr O. Hamdan, Prof M. Hamidou, Prof J.L. Harousseau, Dr D. Hauteville, Dr O. Herault, Dr F. Huguet, Dr C. Hulin, Dr P. Hutin, Dr A. Huynh, Prof N. Ifrah, Dr D. Jacomy, Dr C. Jacquy, Dr H. Jardel, Dr J. Jaubert, Prof J.P. Jouet, Dr R. Kaphan, Dr A. Kentos, Dr O. Ketelslegers, Dr N. Ketterer, Dr B. Kolb, Dr J.L. Kuntz, Dr L. Lacotte, Prof T. Lamy, Dr S. Lassoued, Dr V. Laurencon, Dr P. Lenain, Dr I. Leduc, Dr L. Legros, Dr

G. Lepeu, Dr J.F. Leseve, Dr X. Levaltier, Dr S. Leyvraz , Dr P. Lionne, Dr B. Lioure, Dr E. Loppinet, Dr P. Lore, Dr F. Luthier, Dr M. Maerevoet, Dr C. Magnier, Dr M. Maigre, Dr H. Maisonneuve, Dr F. Maloisel, Prof J.P. Marie, Prof G. Marit, Dr P. Martiat, Dr C. Martin, Prof J.Y. Mary, Dr C. Mathiot, Dr T. Matthes, Dr M. Merlet, Dr N. Meuleman, Prof M. Michallet, Prof J.L. Michaux, Dr L. Michaux, Dr P. Mineur, Dr S. Minvielle, Prof M. Monconduit, Prof P. Moreau, Dr P. Moreau, Dr P. Morice, Dr G. Morlock, Dr Mulleman, Dr P. Naccache, Prof A. Najman, Dr H. Naman, Dr H. Orfeuvre, Dr M. Pages, Dr B. Pallot-Prades, Dr P.Y. Peaud, Dr B. Pegourie-Bandelier, Dr P. Philippe, Dr I. Plantier, Dr J.P. Pollet, Dr J.F. Ramee, Dr E. Randriamalala, Dr A. Regnier, Dr M. Renaud, Dr C. Richard, Dr B. Rio, Dr P. Rodon, Prof C. Rose, Dr V. Salle , Dr B. Salles, Dr F. Sanderson, Dr D. Schlaifer, Prof G. Sebahoun, Dr C. Sebban, Dr M. Simon, Dr P. Solal-Celigny, Prof E. Solary, Prof J.J. Sotto, Dr A.M. Stoppa, Dr N. Straetmans, Dr F. Teillet-Thiebaud, Prof C. Thieblemont, Prof A. Thyss, Dr M. Tiab, Dr C. Traulle, Dr A. Triffet, Dr J. Troncy, Dr L. Valat, Dr B. Valenza, Dr X. Vallantin, Dr J.M. Vantelon, Dr M.C. Vekemans, Dr A. Vekhoff, Dr J.P. Vilque, Dr L. Voillat, Dr E. Voog, Prof E. Wattel, Dr M. Wetterwald, Dr F. Witz, Dr I. Yakoub-Agha, Prof M. Zandecki, Dr C. Zarnitsky, and Dr D. Zylberait.

Table A1. Clinical Parameters, Genetic Abnormalities, and Outcome in the Initial (SNP subset) and Whole (non SNP subset) Cohorts Across The Separate Trials

	IFM-99-02			IFM-99-03			IFM-99-04		
Parameter	**Initial Cohort (n = 79)**	**Whole Cohort* (n = 615)**	***P***	**Initial Cohort (n = 9)**	**Whole Cohort (n = 48)**	***P***	**Initial Cohort (n = 23)**	**Whole Cohort (n = 175)**	***P***
Age, years **< 50** **≥ 50**	22 57	137 470	0.30	1 8	8 38	1.00	4 19	32 142	1.00
Sβ_2M, mg/L **< 5.5** **≥ 5.5**	69 10	529 78	0.96	6 3	35 13	0.70	13 10	108 67	0.63
Serum albumin, g/L **< 35** **≥ 35**	16 54	139 438	0.82	1 5	11 33	1.00	7 13	47 85	0.96
Del(13) **< 30 %** **≥ 30 %**	56 23	445 165	0.70	0 9	0 48	1.00	1 22	14 161	1.00
Del(17p) **< 60 %** **≥ 60 %**	67 5	295 11	0.20	9 0	23 4	0.55	18 2	95 18	0.74
t(4; 14) **No** **Yes**	66 10	390 44	0.43	6 3	26 12	1.00	21 2	116 29	0.26
t(11; 14) **No** **Yes**	64 14	353 105	0.33	8 1	28 10	0.66	19 3	126 22	1.00
ISS, No. of patients **1** **2** **3**	40 20 10	324 178 78	0.96	0 6 3	7 27 13	0.37	3 10 10	17 87 66	0.91
4-year OS, % **%** **95% CI**	73.2 61.6-81.9	79.4 75.9-82.4	0.59	55.6 20.4-80.5	44.4 29.9-58.0	0.46	63.5 39.8-79.9	47.3 39.5-54.7	0.81

* Denoted whole cohort without initial cohort

Abbreviations: Sβ_2M, serum β2-microglobulin; ISS, International Staging System; OS, overall survival.

Table A2. SNP Array Quality Control Data

	Median SNP call % (Range)
CD138$^+$ 250K NspI	96.77 (91.75-98.84)
CD138$^+$ 250K StyI	97.35 (94.08-98.90)

Table A3. Confusion Matrix of del(13) Status Using 500K SNP Array

Del(13) by FISH	Total	Del(13) by 500K SNP	
		Yes	No
Yes	93	87	6
No	97	2	95

Table A4. Molecular Karyotype of the 192 Myeloma Cases

Table A5. Summary of CNAs Identified by Metaphase-CGH and SNP Arrays

Chromosome	Loss ≥ 10% (%)					Gain ≥ 10% (%)				
	Cigudosa et al[22] Metaphase-CGH (n = 25)	Liebisch et al[23] metaphase-CGH (n = 45)	Gutierrez et al[24] metaphase-CGH (n = 74)	Walker et al[25] SNP array (n = 30)	This study SNP array (n = 192)	Cigudosa et al[22] metaphase-CGH (n = 25)	Liebisch et al[23] metaphase-CGH (n = 45)	Gutierrez et al[24] metaphase-CGH (n = 74)	Walker et al[25] SNP array (n = 30)	This study SNP array (n = 192)

1		1p (13)	1p (10)	1p (23)	1p (24)		1q (26)	1p (10) ; 1q(45)	1q (36)	1q (31)
3								3q (16)	3 (30)	3 (36)
4								4q (10)	4 (10)	
5							5 (11)	5q (24)	5 (23)	5 (39)
6	6q (13)	6q (15)	6q (10)	6q (13)	6q (21)		6p (11)	6p (12)	6p (16)	6p (15)
7								7q (14)	7 (30)	7 (30)
8		8p (15)	8p (10)	8p (30)	8p (22)			8q (10)		
9						9q (10)	9q (15) ; 9p (11)	9q (24)	9 (33)	9 (42)
11						11q (20)	11q (15)	11q (22)	11 (27) ; 11q (13)	11 (35) ; 11q (10)
12					12p (13)	12q (10)		12q (10)		
13	13q (30)	13q (37)	13q (39)	13q (53)	13 (45)					
14			14q (12)		14q (23)					
15						15q (10)	15 (11)	15q (22)	15 (43)	15 (49)
16	16 (17)		16q (18)	16q (20)	16p (15) ; 16q (28)					
17						17q (10)		17q (10)		
18							18q (11)			18 (16)
19						19p (30)		19q (18)	19 (57)	19 (43)
20					20p (10)					
21										21 (21)
22				22q (13)	22 (17)	22q (10)		22q (10)		
X female		X (15)		X (57)	X (50)			Xq (10)	Xq (10)	

Table A6. Univariate Analysis for Overall Analysis in the Initial Cohort

Factor	HR	95% CI	*P*
Age (< 50 years)	-		-
Albumin (< 35 g/L)	2.38	1.33 to 4.17	<.01
Creatinine (> 180 µmol/L)	2.20	1.15 to 4.21	<.05
Hemoglobin (< 10 g/dL)	-		-
Platelets (< 130 10^9/L)	-		-
$S\beta_2M$ (≥ 5.5 mg/L)	2.97	1.91 to 4.60	<.001
ISS	2.04	1.52 to 2.73	<.001
Del(13)	1.01	1.00 to 1.01	<.05
t(4;14)	2.27	1.33 to 3.87	<.01
t(11;14)	-	-	-
Del(17p) (≥ 60%)	2.26	1.61 to 4.39	<.05
Hyperdiploidy (≥ 47 chr)	0.48	0.31 to 0.75	=.001
Loss 1p	1.72	1.1 to 2.7	<.05
Gain 1q	1.85	1.21 to 2.85	<.01
Gain 3	0.56	0.35 to 0.90	<.05
Gain 5	0.43	0.27 to 0.70	<.001
Gain 6p/6 *			-
Loss 6q			-
Gain 7			-
Loss 8p/ 8§			-
Gain 9	0.45	0.28 to 0.72	<.001
Gain 11			-
Gain 11q only			-
Loss 12p	2.56	1.49 to 4.35	<.001
Loss 13			-
Loss 14q (1vs>1)			-
Gain 15	0.61	0.39 to 0.93	<.05
Loss 16p			-
Loss 16q			-
Gain 18			-
Gain 19	0.57	0.37 to 0.90	<.05
Loss 20p			-
Gain 21			-
Loss 22			-

*CNAs count included gains of whole chromosome 6 and sub chromosomal material.
§ CNAs count included losses of whole chromosome 8 and sub chromosomal material.

Table A7. Selected segments that impact overall survival. Focal deletions arising from B- cell antigen receptor gene rearrangements at 2p11.2 (IGK@), 14q32.33 (IGHV@) and 22q11.22 (IGL@) were excluded.

Table A8. Frequencies of All Combinations Among the Three Abnormalities

Combination*	Amp(1q23.3)	Amp(5q31.3)	Del(12p13.31)	No. of Risk Factors	No. of Cases	% of Cases
1	0	1	0	0	54	28.1
2	1	1	0	1	12	6.3
3	0	0	0	1	66	34.4
4	0	1	1	1	2	1.0
5	1	0	0	2	38	19.8
6	1	1	1	2	2	1.0
7	0	0	1	2	12	6.3
8	1	0	1	3	6	3.1

* Overlaps are in gray

Table A9. Associations Between CNAs

CNAs	Amp(5q31.3)	Del(12p13.31)	Hyperdiploidy	Gain or Loss of Subchromosomal Material (≥ 1)
Amp(1q23.3)	.022	.62	.11	< .0001
Amp(5q31.3)		.06	<.0001	.16
Del(12p13.31)			.02	.02
Hyperdiploidy				.33

Table A10. Multivariate Analysis for Event-Free Survival

Prognostic Variable	Multivariate analysis	
	Hazard ratio [95% CI*]	*P*
$S\beta_2M \geq 5.5$ *v* < 5.5mg/L	2.03 [1.41-2.91]	.0001
Amp(5q31.3) yes *v* no	0.53 [0.37-0.76]	.0005
Del(12p13.31) yes *v* no	2.23 [1.38-3.61]	.0011

*CI denotes confidence interval

Table A11. Distribution of the Samples According IFM Trials

Abnormality (N vs Y)	Initial Cohort SNP Array Study (N=192)				
	IFM99-02 (n = 79)	IFM99-02* (n = 81)	IFM99-03 (n = 9)	IFM99-04 (n = 23)	*P*
amp(5q31.3)	45/34	56/25	7/2	14/9	.34
del(12p13.31)	71/8	70/11	9/0	20/3	.80

Table A12. Performance of the CNA-Based Model According to Protocols

	Multivariate Analysis	
	Hazard Ratio [95% CI*]	*P*
IFM 99-02 (n = 79)	2.41 [1.36-4.26]	.003
IFM99-02* (n = 81)	2.87 [1.82-4.52]	<.0001
IFM 99-03 (n = 9)	NA	
IFM99-04 (n = 24)	3.65 [1.29-10.32]	.015

*CI denotes confidence interval

Table A13. Univariate and Multivariate Analyses of Expression-Based and CNA-Based Risk Models Applied to the Initial Cohort

	Univariate Analysis (log-rank test)			Adjusted Cox		
Predictor	HR	95% CI*	*P*	HR	95% CI	*P*
15-gene model	3.16	1.98-5.04	<.0001	2.08	1.29-3.34	.003
CNA-based model	2.94	2.17-3.99	<.0001	2.77	2.03-3.78	<.0001

*CI denotes confidence interval

Table A14. Genes of Interest chr1q21-q23

Gene Symbol	Gene Title	Cytoband	Transcript Cluster ID Exon 1.0 ST	Amp(1q23.3) (Wilcoxon test)	Univariate Cox		
				P	*P*	HR	95% CI
ADAR	adenosine deaminase, RNA-specific	1q21.1-q21.2	2436754	<.0001	.001	1.93	1.30 to 2.86
ALDH9A1	aldehyde dehydrogenase 9 family, member A1	1q23.1	2442103	<.0001	.0003	2.05	1.39 to 3.02
AQP10	aquaporin 10	1q21.3	2360186	-	-		
ATP8B2	ATPase, class I, type 8B, member 2	1q21.3	2360206	<.0001	.03	1.42	1.03 to 1.95
CHRNB2	cholinergic receptor, nicotinic, beta 2 (neuronal)	1q21.3	2360346	-	-		
CKS1B	CDC28 protein kinase regulatory subunit 1B	1q21.2	2360452	<.0001	.001	1.50	1.17 to 1.92
CREB3L4	cAMP responsive element binding protein 3-like 4	1q21.3	2359993	<.0001	.0003	4.19	1.93 to 9.08
GPA33	glycoprotein A33 (transmembrane)	1q24.1	2442493	-	.04		
HAX1	HCLS1 associated protein X-1	1q21.3	2360158	<.0001	-		
IL6R	interleukin 6 receptor	1q21	2360257	<.0001	-		
ILF2	interleukin enhancer binding factor 2, 45kDa	1q21.3	2436132	<.0001	<.0001	2.77	1.70 to 4.49
KCNN3	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3	1q21.3	2436826	<.0001	-		
LMX1A	LIM homeobox transcription factor 1, alpha	1q22-q23	2441940	-	ND		
LRRC52	leucine rich repeat containing 52	1q23.3	2365086	-	ND		
MAEL	maelstrom homolog (Drosophila)	1q24.1	2365597	-	ND		
MGST3	microsomal glutathione S-transferase 3	1q23	2365119	.0002	-		
NPR1	natriuretic peptide receptor A/guanylate cyclase A (atrionatriuretic peptide receptor A)	1q21-q22	2359780	-	ND		
NUP210L	nucleoporin 210kDa-like	1q21.3	2436467	-	ND		
RAB13	RAB13, member RAS oncogene family	1q21.2	3457375	-	ND		
RPS27	ribosomal protein S27	1q21	2360027	-	ND		
RPS27	ribosomal protein S27	1q21	2451242	-	ND		
RXRG	retinoid X receptor, gamma	1q22-q23	2442008	-	ND		
S100A1	S100 calcium binding protein A1	1q21	2359722	-	ND		
S100A1	S100 calcium binding protein A1	1q21	4043381	-	ND		
S100A13	S100 calcium binding protein A13	1q21	4045676	.01	-		
SHC1	SHC (Src homology 2 domain containing) transforming protein 1	1q21	2436985	<.0001	.002	1.78	1.23 to 2.57
SLC27A3	solute carrier family 27 (fatty acid transporter), member 3	1q21.3	2359885	.003	-		
SLC39A1	solute carrier family 39 (zinc transporter), member 1	1q21	2436378	.005	_		
TDRD10	tudor domain containing 10	1q21.3	2360310	-	ND		

TMCO1	transmembrane and coiled-coil domains 1	1q22-q25	2442134	<.0001	.04	1.37	1.01 to 1.85
TPM3	tropomyosin 3	1q21.2	2436526	<.0001	.03	1.92	1.07 to 3.45
UBAP2L	ubiquitin associated protein 2-like	1q21.3	2360083	<.0001	<.0001	3.48	2.11 to 5.76
UBE2Q1	ubiquitin-conjugating enzyme E2Q family member 1	1q21.3	2436716	<.0001	.01	1.92	1.16 to 3.19
UCK2	uridine-cytidine kinase 2	1q23	2365210	<.0001	-		

ND: Not done

Table A15. Genes of Interest chr5q31

Gene Symbol	Gene Title	Cytoband	Transcript Cluster ID Exon 1.0 ST	Amp(5q31.3) (Wilcoxon test)	Univariate Cox		
				P	*P*	HR	95% CI
CENTD3	centaurin, delta 3	5q31.3	2878809	-	ND		
CNOT8	CCR4-NOT transcription complex, subunit 8	5q31-q33	2836856	-	ND		
DIAPH1	diaphanous homolog 1 (Drosophila)	5q31	2878662	.01	_		
FAT2	FAT tumor suppressor homolog 2 (Drosophila)	5q32-q33	2882026	-	ND		
FCHSD1	FCH and double SH3 domains 1	5q31.3	2878778	-	ND		
GALNT10	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 10 (GalNAc-T10)	5q33.2	2836518	-	ND		
GEMIN5	gem (nuclear organelle) associated protein 5	5q33.2	2882897	-	ND		
GNPDA1	glucosamine-6-phosphate deaminase 1	5q21	2879028	.003	-		
HAND1	heart and neural crest derivatives expressed 1	5q33	2882747	-	ND		
HDAC3	histone deacetylase 3	5q31	2878726	<.0001	-		
KIAA0141	KIAA0141	5q31.3	2832963	<.0001	-		
LARP1	La ribonucleoprotein domain family, member 1	5q33.2	2836738	.0002	-		
LOC492311	similar to bovine IgA regulatory protein	5q31	2831591	-	ND		
MRPL22	mitochondrial ribosomal protein L22	5q33.1-q33.3	2836886	.02	-		
NDFIP1	Nedd4 family interacting protein 1	5q31.3	2833078	-	ND		
NRG2	neuregulin 2	5q23-q33	2878074	<.0001	.04	.50	.25 to .97
PCDH1	protocadherin 1	5q32-q33	2878943	-	ND		
PCDH12	protocadherin 12	5q31	2878987	-	ND		
PCDHA13 /// PCDHA8 /// PCDHA1	protocadherin alpha 13 /// protocadherin alpha 8 /// protocadherin alpha 1	5q31	2832115	-	ND		
PCDHB1	protocadherin beta 1	5q31	2832291	-	ND		
PCDHB16 ///	protocadherin beta 16 /// protocadherin beta 13	5q31	2832392	.007	-		

PCDHB13							
PCDHB5	protocadherin beta 5	5q31	2832325	-	ND		
PCDHB2	protocadherin beta 2	5q31	2832297	-	ND		
PCDHB3	protocadherin beta 3	5q31	2832310	-	ND		
PCDHB4	protocadherin beta 4	5q31	2832315	-	ND		
PCDHB6	protocadherin beta 6	5q31	2832355	-	ND		
PCDHB7	protocadherin beta 7	5q31	2832378	-	ND		
PCDHB8 /// PCDHB13	protocadherin beta 8 /// protocadherin beta 13	5q31	2832387	-	ND		
PCDHB9 /// PCDHB10	protocadherin beta 9 /// protocadherin beta 10	5q31	2832403	.016	-		
PSD2	pleckstrin and Sec7 domain containing 2	5q31.3	2831436	-	ND		
PURA	purine-rich element binding protein A	5q31	2831567	.0003	-		
RNF14	ring finger protein 14	5q23.3-q31.1	2833024	.01	-		
SAP30L	SAP30-like	5q33.2	2836665	.007	-		
SLC36A1	solute carrier family 36 (proton/amino acid symporter), member 1	5q33.1	2835848	.009	-		
SLC36A2	solute carrier family 36 (proton/amino acid symporter), member 2	5q33.1	2881950	-	ND		
SLC36A3	solute carrier family 36 (proton/amino acid symporter), member 3	5q33.1	2881923	-	ND		
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	5q31.3-q32	2882098	-	ND		

ND : Not done

Table A16. Genes of Interest chr12p12-13

Gene Symbol	Gene Title	Cytoband	Transcript Cluster ID Exon 1.0 ST	Del(12p13.31) (Wilcoxon test)	Univariate Cox		
				P	***P***	**HR**	**95% CI**
ACRBP /// LPAR5	acrosin binding protein /// lysophosphatidic acid receptor 5	12p13.31	3442150	-	ND		
CD27	CD27 molecule	12p13	3402506	.002	.0002	1.33	1.14 to 1.53
CHD4 /// SCARNA11	chromodomain helicase DNA binding protein 4 /// small Cajal body-specific RNA 11	12p13	3442054	-	ND		
COPS7A	COP9 constitutive photomorphogenic homolog subunit 7A (Arabidopsis)	12p13.31	3402697	.0005	-		
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	12p13	3402625	-	ND		
ING4	inhibitor of growth family, member 4	12p13.31	3442176	.0002	-		

LAG3	lymphocyte-activation gene 3	12p13.32	3402757	.001	.004	1.94	1.23 to 3.06
MLF2	myeloid leukemia factor 2	12p13	3442282	<.0001	-		
MRPL51	mitochondrial ribosomal protein L51	12p13.3-p13.1	3441955	.002	-		
NOL1	nucleolar protein 1, 120kDa	12p13	3442024	<.0001	-		
PTMS	parathymosin	12p13	3402736	-	ND		
TAPBPL	TAP binding protein-like	12p13.31	3402522	.0002	.003	1.64	1.19 to 2.26
VAMP1	vesicle-associated membrane protein 1 (synaptobrevin 1)	12p	3441941	.001	-		
ZNF384	zinc finger protein 384	12p12	3442205	<.0001	-		

ND : Not done

Table A17. Table of Correspondences Between UMGC Probes and Affymetrix Probe Sets

UMGC Probe[27]	Gene Symbol	Affymetrix Probe Set Exon 1.0 ST
UMGC_17217	*FAM49A*	2541699
UMGC_11702	*ATF4*	3945914
UMGC_09916	*CTSF*	3378344
UMGC_02996	*ALDH2*	3432090
UMGC_06566	*CNDP2*	3793760
UMGC_01066	*STMN1*	2402459
UMGC_06118	*AFG3L2*	3799415
UMGC_05764	*STK38*	2951916
UMGC_01969	*PARP1*	2458773
UMGC_08943	*CPSF6*	3421446
UMGC_00460	*MGAT5/ LOC151162*	2506903
UMGC_11582	*TOX2*	3886294
UMGC_11580	*FRY*	3484497
UMGC_10992	*FLJ21438*	3853453
UMGC_07324	*MGST1*	3406589

Fig A1. Kaplan-Meier analysis of overall survival in patients of the non SNP subset and of the SNP subset.

Fig A2. Validation of 500K copy number method for estimating individual chromosome loss compared to FISH. 500K platform for chromosomal 13q14 are plotted on the y-axis as a function of FISH-derived % cells with anomalous 13q14 copy number. Red and blue frames mark del(13) status of the patients obtained with 500K platform and FISH respectively.

Fig.A3. Hierarchical clustering of MM according to group 1 and group 2 CNAs. Patients without CNAs were omitted; * indicated dead patients, hyperdiploid MM are in blue, red is for gain and green for loss.

Fig A4. Prognostic impact of amp(1q23.3), amp(5q31.3), del(12p13.31) on event-free survival.

Fig A5. Prognostic impact of amp(5q31.3), del(12p13.31) and $S\beta_2M \geq 5.5$ mg/L on event-free survival.

Fig A6. Independent prognostic impact of 15-gene predictor and the CNA-based model in initial cohort

Fig. A7. Kaplan-Meier plot of overall survival considering the expression of the genes of interest residing in 1q21-23. Genes are ordered by increasing P-value (log-rank)

ILF2 (NF45)
ADAR (ADAR1, IFI4, G1P1, DSRAD, DRADA)
ALDH9A1 (ALDH4, ALDH7, ALDH9, E3, TMABADH)
UBAP2L (NICE-4)
UBE2Q1 (GTAP, NICE-5, UBE2Q)
SHC1 (SHC, SHCA)
CKS1B (CKS1)
TMCO1 (PCIA3, TMCC4)
CREB3L4 (AIBZIP, ATCE1, CREB3, CREB4, JAL)

Fig. A8. Kaplan-Meier plot of overall survival considering the expression of the gene of interest residing in 5q31: NRG2 (DON1, HRG2, NTAK).

Fig. A9. Kaplan-Meier plot of overall survival considering the expression of the genes of interest residing in 12p31. Genes are ordered by increasing P-value (log-rank).

CD27 (TNFRSF7)
LAG3 (CD223)
TAPBPL

Fig A10. Impact of del(16q12.1) (A) and del(16q23.2) (B) on OS.

Fig A11. Kaplan-Meier analysis of overall survival in patients of initial and validation cohorts.