EUROPEAN HEMATOLOGY ASSOCIATION
haematologica
Journal of the European Hematology Association

Circulating tumor DNA as a liquid biopsy in plasma cell dyscrasias

by Bernhard Gerber, Martina Manzoni, Valeria Spina, Alessio Bruscaggin, Marta Lionetti, Sonia Fabris, Marzia Barbieri, Gabriella Ciceri, Alessandra Pompa, Gabriela Forestieri, Erika Lerch, Paolo Servida, Francesco Bertoni, Emanuele Zucca, Michele Ghielmini, Agostino Cortelezzi, Franco Cavalli, Georg Stussi, Luca Baldini, Davide Rossi, and Antonino Neri

Haematologica 2018 [Epub ahead of print]

Citation: Bernhard Gerber, Martina Manzoni, Valeria Spina, Alessio Bruscaggin, Marta Lionetti, Sonia Fabris, Marzia Barbieri, Gabriella Ciceri, Alessandra Pompa, Gabriela Forestieri, Erika Lerch, Paolo Servida, Francesco Bertoni, Emanuele Zucca, Michele Ghielmini, Agostino Cortelezzi, Franco Cavalli, Georg Stussi, Luca Baldini, Davide Rossi, and Antonino Neri. Circulating tumor DNA as a liquid biopsy in plasma cell dyscrasias.
Haematologica. 2018; 103:xxx
doi:10.3324/haematol.2017.184358

Circulating tumor DNA as a liquid biopsy in plasma cell dyscrasias

Bernhard Gerber,[1] Martina Manzoni,[2] Valeria Spina,[3] Alessio Bruscaggin,[3] Marta Lionetti,[2] Sonia Fabris,[4] Marzia Barbieri,[4] Gabriella Ciceri,[2] Alessandra Pompa,[4] Gabriela Forestieri,[3] Erika Lerch,[5] Paolo Servida,[5] Francesco Bertoni,[3] Emanuele Zucca,[5] Michele Ghielmini,[5] Agostino Cortelezzi,[2,4] Franco Cavalli,[3,5] Georg Stussi,[1] Luca Baldini,[2,4] Davide Rossi[1,3] and Antonino Neri[2,4]

[1]Division of Hematology, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland;

[2]Department of Oncology and Hemato-oncology, University of Milan, Italy;

[3]Institute of Oncology Research, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland;

[4]Hematology Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy;

[5]Division of Oncology, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland.

BG and MM contributed equally to this work

Running head: liquid biopsy genotyping in plasma cell dyscrasias

CORRESPONDING AUTHORS

Antonino Neri, MD, PhD, Department of Oncology and Hemato-Oncology, University of Milan, Hematology-CTMO, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, via F. Sforza, 35, Milan 20122, Italy. E-mail: antonino.neri@unimi.it

Davide Rossi, MD, PhD, Hematology, Oncology Institute of Southern Switzerland and Institute of Oncology Research, 6500 Bellinzona, Switzerland; Ph +41 91 820 03 62; Fax +41 91 820 03 97; E-mail davide.rossi@eoc.ch

Word count: 1499

Number of tables: 1

Number of figures: 1

Number of Supplementary files: 1

FUNDING

This work was supported by grants from the AIRC (Associazione Italiana per la Ricerca sul Cancro) to AN (IG10136 and IG16722) and by a grant from ABREOC 2016 to BG

Multiple myeloma (MM) is a clinically and genetically heterogeneous malignant proliferation of plasma cells (PCs) with a typical multifocal distribution in the bone marrow (BM) and occasional extra-medullary dissemination.[1] Advances in the genetic knowledge of MM are increasingly translated into biomarkers to refine diagnosis, prognostication and treatment of patients.[2]

MM genotyping has so far relied on the analysis of purified PCs from the bone marrow (BM) aspirate, which may fail in capturing the postulated spatial heterogeneity of the disease and imposes technical hurdles limiting its transfer in the routine and clinical grade diagnostic laboratory. In addition, longitudinal monitoring of disease molecular markers may be limited by patient discomfort caused by repeated BM samplings during disease course. Circulating tumor DNA is shed into the peripheral blood (PB) by tumor cells and can be used as source of tumor DNA for the identification of cancer-gene somatic mutations, with obvious advantages in terms of accessibility. In addition, the systemic origin of cell-free DNA (cfDNA) allows catching the entire tumor heterogeneity.[3] Tumor cfDNA was identified in MM patients by preliminary studies tracking the clonotypic V(D)J rearrangement as disease fingerprint,[4] or genotyping a highly restricted set of cancer genes that were not specifically addressed to resolve the typical MM mutational landscape.[5-7] We developed a CAPP-seq ultra-deep targeted next-generation sequencing (NGS) approach to genotype a gene panel specifically designed to maximize the mutation recovery in plasma cell tumors, and compared the mutational profiling of cfDNA and tumor genomic DNA (gDNA) of purified PCs from BM aspirates in a consecutive series of patients representative of different clinical stages of PC tumors ranging from monoclonal gammopathy of undetermined significance (MGUS), to smoldering MM, and symptomatic MM.

The study was based on a series of 28 patients with PC disorders, whose clinical and molecular characteristics were consistent with an unselected cohort of PC dyscrasia patients (Supplementary Table S1) [two had MGUS, five smoldering MM (SMM), and 21 symptomatic MM]. The study was conducted according to good clinical practice and the ethical principles outlined in the Declaration of Helsinki. All patients provided written informed consent. The following material was collected: cfDNA isolated from plasma; tumor gDNA from CD138+ purified BM PCs for comparative purposes, and germline gDNA extracted from PB granulocytes after Ficoll gradient separation, to filter out polymorphisms. The sampling was done in 25 newly diagnosed and three relapsed/refractory treated patients. A targeted resequencing gene panel, including coding exons and splice sites of 14 genes (target region: 31 kb: *BRAF*, *CCND1*, *CYLD*, *DIS3*, *EGR1*, *FAM46C*, *IRF4*, *KRAS*, *NRAS*, *PRDM1*, *SP140*, *TP53*, *TRAF3*, *ZNF462*; Supplementary Table S2) was specifically designed and optimized to allow *a priori* the recovery of at least one mutation in 68% (95% confidence interval: 58-76%) of patients, based on literature data.[8-10] Ultra-deep NGS was performed

on MiSeq (Illumina) using the CAPP-seq library preparation strategy (NimbleGen).[11] The somatic function of VarScan2 was used to call non-synonymous somatic mutations, and a stringent bioinformatic pipeline was developed and applied to filter out sequencing errors (detection limit $3x10^{-3}$). The sensitivity and specificity of plasma cfDNA genotyping were calculated in comparison with tumor gDNA genotyping as the gold standard. Details of the experimental procedures are given in the Supplementary Methods.

cfDNA was detectable in plasma samples with an average of ~11 000 haploid genome-equivalents per mL of plasma (range: 19-52562 hGE/mL; median: 6617 hGE/mL). The amount of cfDNA correlated with clinic-pathological parameters reflecting tumor load/extension, including BM PC infiltration (Spearman's rho coefficient=0.42, *P*=0.02; Supplementary Figure S1A), and clinical stage. Indeed patients presenting with ISS stage 3 had significantly higher amounts of cfDNA compared with MGUS/SMM samples and MM cases at ISS stages 1-2 (*P*=0.01; Supplementary Figure S1B, Mann-Whitney test). Conversely, we did not observe differences in cfDNA concentration between newly diagnosed and relapsed/refractory MM patients (data not shown). More than 90% of the target region was covered ≥1000X in all plasma samples, and ≥2000X in 23/28 (Supplementary Figure S2 and Supplementary Table S3). Overall, within the interrogated genes, 18/28 (64%) patients had at least one non-synonymous somatic mutation detectable in cfDNA (Figure 1A and Table 1A); 28 total variants were identified, with a range of 1-4 mutations per patient. Quite consistent with the typical spectrum of mutated genes in MM, plasma cfDNA genotyping revealed somatic variants of *NRAS* in 25%; *KRAS* in 14%; *TP53*, *TRAF3* and *FAM46C* in 11%, respectively, *CYLD* and *DIS3* in 7%, respectively, and *BRAF* and *IRF4* in 4% of cases, respectively. Variants in *NRAS*, *KRAS* and *BRAF* genes occurred in a mutually exclusive manner, and they overall involved 43% of patients. *TP53* mutations were positively associated with the deletion of the remaining allele as revealed by fluorescence *in situ* hybridization on purified PCs (*P*=0.02, Fisher-exact test). Overall, the molecular spectrum of mutations discovered in tumor cfDNA reflected previous observations in genomic studies based on PC genotyping (see representative example for the two most frequently mutated genes in Supplementary Figure S3), thus supporting the tumor origin of the mutations identified in cfDNA.

To validate the tumor origin of mutations discovered in cfDNA and to derive the accuracy of our approach in resolving tumor genetics, the genotype of cfDNA was matched with that of gDNA from purified BM PCs in all the patients. Sequencing of tumor gDNA identified 39 somatic mutations in 20/28 (71.4%) patients (Figure 1A). cfDNA genotyping correctly identified 72% of mutations (n=28/39) that were discovered in tumor PCs (Supplementary Figure S4A); overall the variant allele frequencies in plasma samples correlated with those in tumor biopsies (Pearson correlation coefficient=0.58, *P*=9.6e-05; Supplementary

Figure S4B) and with the degree of bone marrow involvement (Pearson correlation coefficient=0.5, *P*=0.006). Specifically, of the 28 mutations correctly identified in tumor cfDNA, four were detected in two SMM patients out of a total of 7 biopsy-confirmed mutations (4/7, 57%) in three SMM patients, and 24 were detected in 16 MM cases out of a total of 32 biopsy-confirmed mutations (24/32, 75%) in 17 MM cases. Notably, BM PC confirmed mutations not discovered in cfDNA (n=11) had a low representation in the tumor (median allelic frequency: 2.5%; range: 1.1-4.96%) (Table 1B, Figure 1B). Since circulating tumor DNA is diluted in cfDNA from normal cells,[12, 13] variants that are already rare in tumor gDNA are much less represented in plasma and may fall below the sensitivity threshold of the CAPP-seq under the experimental conditions adopted in this work. Consistently, based on ROC analysis, cfDNA genotyping has the best performance in detecting tumor PC confirmed mutations when they are represented in at least 5% of the alleles of tumor plasma cells (Supplementary Figure S4C). Above this threshold, cfDNA genotyping detected 100% of biopsy-confirmed mutations. Noteworthy, cfDNA genotyping was still able to detect almost half (10/21) of low-abundance mutations in tumor PC (i.e. allelic frequency <20%), indicating a good capacity of tumor cfDNA to mirror also the subclonal composition of the tumor. Of course, these data concerning the sensitivity of cfDNA genotyping refer to the depth of coverage used in the paper, and higher depth may allow a better overlap of gDNA and cfDNA. In none of the cases cfDNA genotyping identified additional somatic mutations not detected in the purified BM PCs, thus suggesting that, as far as our limited patient cohort is concerned, the genotype of PC collected from a single tumor site is already representative of the entire tumor genetics. Alternatively, spatial genomic heterogeneity, supported by very recent findings in MM,[14] may exist but involving minor subclones not sufficiently represented to be detectable in plasma.

Our results provide the proof of principle that circulating tumor cfDNA genotyping is a feasible, non-invasive, real-time approach that reliably detects clonal and subclonal somatic mutations represented in at least 5% of alleles in tumor PCs. Despite the genetic heterogeneity characterizing MM, and the inclusion in the study cohort of seven patients at pre-malignant/asymptomatic disease stages, the designed gene-panel employed in our study proved to be very effective, in that it allowed the recovery of at least one mutation in tumor gDNA of 20/28 (71%) cases. To the best of our knowledge, this is the first gene panel specifically created to maximize mutational recovery in MM patients by using an affordable number of genes, and by virtue of this potentially effective and manageable even in clinical practice in a hopefully near future.

One of the original findings of the study is that cfDNA genotyping can resolve tumor genetics also in cases at early disease stages as SMM patients, who may benefit the most from this non-invasive approach. Indeed, among asymptomatic patients cfDNA genotyping could allow a non-invasive longitudinal molecular

monitoring of clonal evolution and the identification of the switch point on which the disease acquires high-risk genetic features. This has been prevented so far by the unfeasibility of serial BM sampling in the clinical routine.

An immediate clinical application of cfDNA genotyping in MM could be the incorporation of this minimally-invasive method in clinical trials for the identification of patients carrying actionable mutations and their longitudinal genetic monitoring during targeted therapy administration or for the estimation of minimal residual disease.

REFERENCES

1. Kumar SK, Rajkumar V, Kyle RA, et al. Multiple myeloma. Nat Rev Dis Primers. 2017;3:17046.
2. Lionetti M, Neri A. Utilizing next-generation sequencing in the management of multiple myeloma. Expert Rev Mol Diagn. 2017;17(7):653-663.
3. Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol. 2014;32(6):579-586.
4. Oberle A, Brandt A, Voigtlaender M, et al. Monitoring multiple myeloma by next-generation sequencing of V(D)J rearrangements from circulating myeloma cells and cell-free myeloma DNA. Haematologica. 2017;102(6):1105-1111.
5. Kis O, Kaedbey R, Chow S, et al. Circulating tumour DNA sequence analysis as an alternative to multiple myeloma bone marrow aspirates. Nat Commun. 2017;8:15086.
6. Mithraprabhu S, Khong T, Ramachandran M, et al. Circulating tumour DNA analysis demonstrates spatial mutational heterogeneity that coincides with disease relapse in myeloma. Leukemia. 2017;31(8):1695-1705.
7. Rustad EH, Coward E, Skytoen ER, et al. Monitoring multiple myeloma by quantification of recurrent mutations in serum. Haematologica. 2017;102(7):1266-1272.
8. Bolli N, Avet-Loiseau H, Wedge DC, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun. 2014;5:2997.
9. Kortuem KM, Braggio E, Bruins L, et al. Panel sequencing for clinically oriented variant screening and copy number detection in 142 untreated multiple myeloma patients. Blood Cancer J. 2016;6:e397.
10. Lohr JG, Stojanov P, Carter SL, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell. 2014;25(1):91-101.
11. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med. 2014;20(5):548-554.
12. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. Nat Rev Clin Oncol. 2013;10(8):472-484.
13. Abbosh C, Birkbak NJ, Wilson GA, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature. 2017;545(7655):446-451.
14. Rasche L, Chavan SS, Stephens OW, et al. Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. Nat Commun. 2017;8(1):268.

Table 1A. Somatic non-synonymous mutations discovered by cfDNA genotyping and their validation in tumor gDNA

ID Sample	*Gene*	*CHR*	*Absolute position**	*REF*	*VAR*	*cDNA change*[§]	*Protein change*	*cfDNA allele fraction*	*gDNA allele fraction*
ID1	*CYLD*	chr16	50820803	A	T	c.1987A>T	p.R663W	0.95%	26.75%
ID2	*KRAS*	chr12	25380276	T	A	c.182A>T	p.Q61L	25.01%	44.72%
ID3	*NRAS*	chr1	115258747	C	A	c.35G>T	p.G12V	3.08%	63.07%
ID5	*KRAS*	chr12	25380279	C	T	c.179G>A	p.G60D	1.05%	15.42%
ID7	*FAM46C*	chr1	118166229	T	C	c.739T>C	p.Y247H	3.82%	53.38%
ID7	*NRAS*	chr1	115256529	T	C	c.182A>G	p.Q61R	6.72%	54.57%
ID7	*TRAF3*	chr14	103363617	A	-	c.839_839delA	p.E280fs*3	9.66%	76.97%
ID8	*CYLD*	chr16	50813911	G	A	c.1474G>A	p.G492S	0.87%	3.93%
ID11	*KRAS*	chr12	25398281	C	T	c.38G>A	p.G13D	4.39%	16.82%
ID12	*NRAS*	chr1	115256529	T	C	c.182A>G	p.Q61R	3.33%	35.14%
ID13	*NRAS*	chr1	115256530	G	T	c.181C>A	p.Q61K	32.52%	19.11%
ID15	*DIS3*	chr13	73337723	C	T	c.1993G>A	p.E665K	37.86%	86.29%
ID15	*TP53*	chr17	7578269	G	A	c.580C>T	P.L194F	36.29%	81.79%
ID17	*TP53*	chr17	7577610	T	A	c.673-2A>T	p.224?	8.84%	79.53%
ID18	*IRF4*	chr6	394920	G	T	c.316G>T	p.D106Y	1.48%	39.08%
ID18	*TRAF3*	chr14	103336686	A	G	c.148A>G	p.K50E	0.29%	4.86%
ID19	*FAM46C*	chr1	118165764	G	C	c.274G>C	p.D92H	0.68%	6.98%
ID19	*NRAS*	chr1	115256521	A	C	c.190T>G	p.Y64D	0.65%	9.97%
ID21	*NRAS*	chr1	115256529	T	G	c.182A>C	p.Q61P	0.54%	26.06%
ID21	*TP53*	chr17	7578406	C	T	c.524G>A	p.R175H	0.73%	38.91%
ID26	*FAM46C*	chr1	118165699	G	C	c.209G>C	p.R70P	1.22%	5.16%
ID26	*FAM46C*	chr1	118166036	C	G	c.546C>G	p.D182E	5.35%	18.83%
ID26	*NRAS*	chr1	115256529	T	C	c.182A>G	p.Q61R	16.08%	32.59%
ID26	*NRAS*	chr1	115256530	G	T	c.181C>A	p.Q61K	11.55%	15.04%
ID27	*DIS3*	chr13	73337723	C	T	c.1993G>A	p.E665K	0.64%	51.36%
ID27	*TRAF3*	chr14	103363719	C	T	c.941C>T	p.S314F	0.42%	33.81%
ID28	*BRAF*	chr7	140453136	A	T	c.1799T>A	p.V600E	1.43%	32.88%
ID29	*KRAS*	chr12	25398281	C	T	c.38G>A	p.G13D	11.36%	43.4%

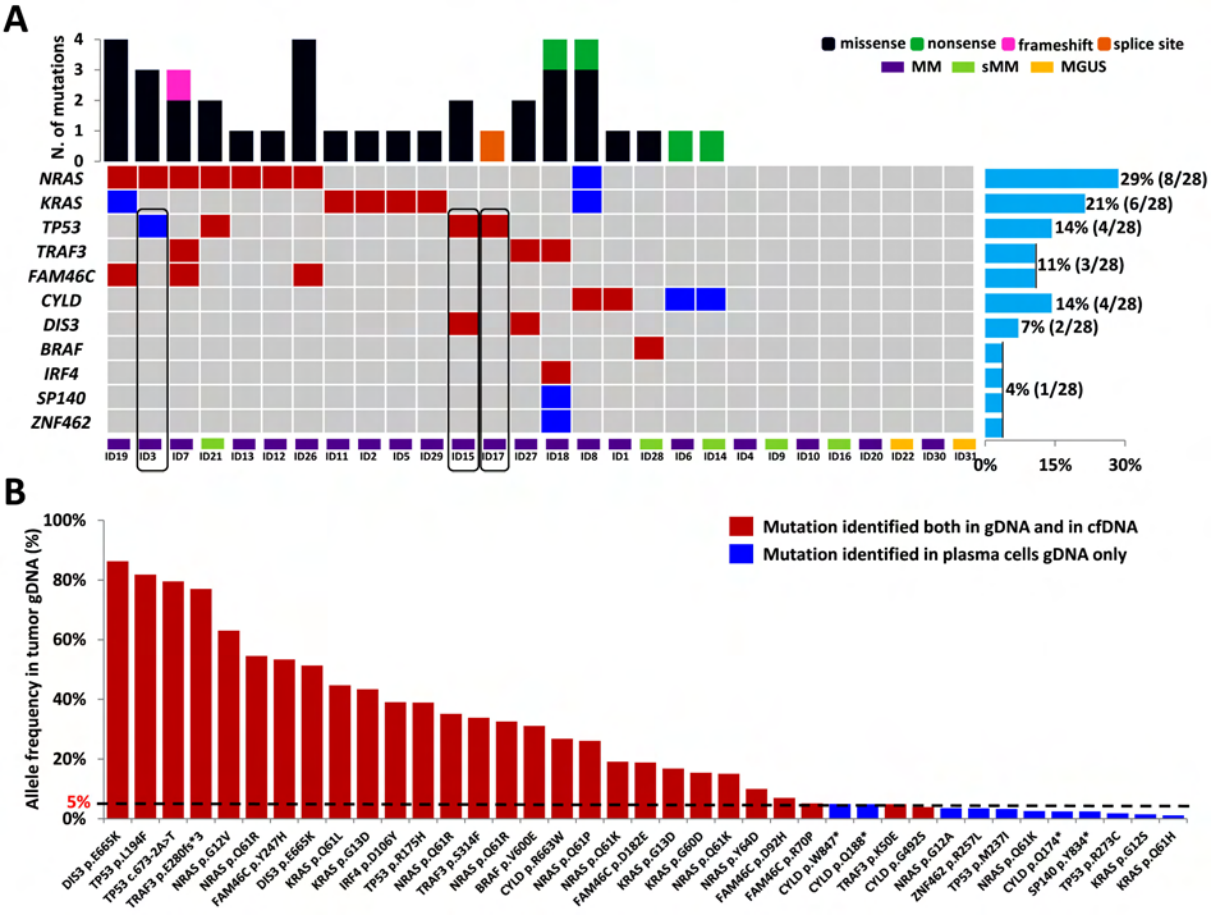
Table 1B. Somatic non-synonymous mutations discovered in tumor gDNA genotyping and missed in plasma cfDNA

ID Sample	*Gene*	*CHR*	*Absolute position**	*REF*	*VAR*	*cDNA change*[§]	*Protein change*	*cfDNA allele fraction*	*gDNA allele fraction*
ID3	*TP53*	chr17	7577570	C	T	c.711G>A	p.M237I	-	3.31%
ID3	*TP53*	chr17	7577121	G	A	c.817C>T	p.R273C	-	1.83%
ID6	*CYLD*	chr16	50785530	C	T	c.520C>T	p.174Q*	-	2.44%
ID8	*CYLD*	chr16	50785572	C	T	c.562C>T	p.188Q*	-	4.88%
ID8	*KRAS*	chr12	25380275	T	A	c.183A>T	p.Q61H	-	1.14%
ID8	*NRAS*	chr1	115256530	G	T	c.181C>A	p.Q61K	-	2.55%
ID14	*CYLD*	chr16	50828193	G	A	c.2540G>A	p.W847*	-	4.96%
ID18	*SP140*	chr2	231176307	C	A	c.2502C>A	p.Y834*	-	2.43%
ID18	*ZNF462*	chr9	109686963	G	T	c.770G>T	p.R257L	-	3.5%
ID19	*KRAS*	chr12	25398285	C	T	c.34G>A	p.G12S	-	1.46%
ID19	*NRAS*	chr1	115258747	C	G	c.35G>C	p.G12A	-	3.58%

Abbreviations: CHR, chromosome; REF, reference allele; VAR, variant allele.
*Absolute chromosome coordinates of each variant based on the hg19 version of the human genome assembly.
[§]cDNA change determined on the following RefSeq: NM_015247.2 for *CYLD*, NM_033360.3 for *KRAS*, NM_002524.4 for *NRAS*, NM_017709.3 for *FAM46C*, NM_003300.3 for *TRAF3*, NM_014953.3 for *DIS3*, NM_000546.5 for *TP53*, NM_002460.3 for *IRF4*, NM_004333.4 for *BRAF*, NM_007237.4 for *SP140*, NM_021224.4 for *ZNF462*.

FIGURE LEGEND

Figure 1. Overview of the mutations identified in the PC dyscrasia series. (**A**) Mutations detected in plasma cfDNA and confirmed in tumor gDNA are filled in red; mutations detected in tumor gDNA only are filled in blue. Each column represents one tumor sample and each row represents one gene. The fraction of tumors with mutations in each gene is plotted (right). The number and the type of mutations in a given tumor are plotted above the heat map. Patients positive for del(17p) are framed in black. (**B**) Bar graph of the allele frequencies in tumor gDNA of the variants that were discovered in plasma cfDNA (red bars) or missed in plasma cfDNA (blue bars). The dashed line tracks the 5% allelic frequency threshold.


A
N. of mutations
missense
nonsense
frameshift
splice site
MM
sMM
MGUS
NRAS
KRAS
TP53
TRAF3
FAM46C
CYLD
DIS3
BRAF
IRF4
SP140
ZNF462
29% (8/28)
21% (6/28)
14% (4/28)
11% (3/28)
14% (4/28)
7% (2/28)
4% (1/28)
ID19 ID3 ID7 ID21 ID13 ID12 ID26 ID11 ID2 ID5 ID29 ID15 ID17 ID27 ID18 ID8 ID1 ID28 ID6 ID14 ID4 ID9 ID10 ID16 ID20 ID22 ID30 ID31
0% 15% 30%
B
Allele frequency in tumor gDNA (%)
Mutation identified both in gDNA and in cfDNA
Mutation identified in plasma cells gDNA only
5%
DIS3 p.E665K
TP53 p.I194F
TP53 c.673-2A>T
TRAF3 p.E280fs*3
NRAS p.G12V
NRAS p.Q61R
FAM46C p.Y247H
DIS3 p.E665K
KRAS p.Q61L
KRAS p.G13D
IRF4 p.D106Y
TP53 p.R175H
NRAS p.Q61R
TRAF3 p.S314F
NRAS p.Q61R
BRAF p.V600E
CYLD p.R663W
NRAS p.Q61P
NRAS p.Q61K
FAM46C p.D182E
KRAS p.G13D
KRAS p.G60D
NRAS p.Q61K
NRAS p.Y64D
FAM46C p.D92H
FAM46C p.R70P
CYLD p.W847*
CYLD p.Q188*
TRAF3 p.K50E
CYLD p.G492S
NRAS p.G12A
ZNF462 p.R257L
TP53 p.M237I
NRAS p.Q61K
CYLD p.Q174*
SP140 p.Y834*
TP53 p.R273C
KRAS p.G12S
KRAS p.Q61H

SUPPLEMENTARY MATERIAL

Circulating tumor DNA as a liquid biopsy in plasma cell dyscrasias

Bernhard Gerber,[1] Martina Manzoni,[2] Valeria Spina,[3] Alessio Bruscaggin,[3] Marta Lionetti,[2] Sonia Fabris,[4] Marzia Barbieri,[4] Gabriella Ciceri,[2] Alessandra Pompa,[4] Gabriela Forestieri,[3] Erika Lerch,[5] Paolo Servida,[5] Francesco Bertoni,[3] Emanuele Zucca,[5] Michele Ghielmini,[5] Agostino Cortelezzi,[2,4] Franco Cavalli,[3,5] Georg Stussi,[1] Luca Baldini,[2,4] Davide Rossi[1,3] and Antonino Neri[2,4]

[1]Division of Hematology, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland;

[2]Department of Oncology and Hemato-oncology, University of Milano, Italy;

[3]Institute of Oncology Research, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland;

[4]Hematology Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy;

[5]Division of Oncology, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland.

BG and MM contributed equally to this work

Supplementary Methods

Patients

The study had a prospective, observational, nonintervention design and consisted in the collection of peripheral blood (PB) samples and clinical data from plasma cell (PC) dyscrasia patients. Inclusion criteria were: (1) male or female adults ≥18 years old; (2) diagnosis of multiple myeloma (MM) or monoclonal gammopathy of undetermined significance (MGUS) after pathological revision; (3) evidence of signed informed consent. A total of 28 patients fulfilled the inclusion criteria and were recruited for the study from September 2016 to May 2017 (Supplementary Table S1). The following biological material was collected: (1) cfDNA isolated from plasma, (2) tumor genomic DNA (gDNA) from the CD138+ purified PCs from BM aspiration, for comparative purposes, and (3) normal germline gDNA extracted from peripheral blood (PB) granulocytes after Ficoll separation. Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki.

Isolation and analysis of plasma cfDNA

PB (20 ml maximum) was collected in Cell-Free DNA BCT tubes that allow obtaining stable cfDNA samples while preventing gDNA contamination that may occur due to nucleated cell disruption during sample storage, thus avoiding pre-analytical issues affecting cfDNA genotyping. PB was centrifuged at 820 g for 10 min to separate plasma from cells. Plasma was then further centrifuged at 20000 g for 10 min to pellet and remove any remaining cells and stored at -80°C until DNA extraction. cfDNA was extracted from 1-3 ml aliquots of plasma (to allow the recovery of enough genomic equivalents of DNA to reach a genotyping sensitivity of 10^{-3}) using the QIAamp circulating nucleic acid kit (Qiagen) and quantized using Quant-iT™ PicoGreen dsDNA Assay kit (ThermoFisher Scientific). Contamination of plasma cfDNA from gDNA released by blood nucleated cell disruption was ruled out by checking, through the Bioanalyzer (Agilent Technologies) instrument, the size of the DNA extracted from plasma.

gDNA extraction

PB granulocytes were separated by Ficoll gradient density centrifugation as a source of normal germline gDNA. Tumor gDNA was isolated from PCs purified using CD138 immunomagnetic microbeads as previously described [1, 2] (CD138+ cell percentage was ≥90% in all cases). gDNA was extracted according to standard procedures.

Library design for hybrid selection

A targeted resequencing gene panel, including coding exons and splice sites of 14 genes that are recurrently mutated in MM patients, was specifically designed for this project (target region: 30989bp: *BRAF*, *CCND1*, *CYLD*, *DIS3*, *EGR1*, *FAM46C*, *IRF4*, *KRAS*, *NRAS*, *PRDM1*, *SP140*, *TP53*, *TRAF3*, *ZNF462*; Supplementary Table S2). Inclusion *criteria* of gene panel design were based on publicly available sequencing data from three distinct datasets [3-5] and were as follows: (i) genes that were recurrently mutated in ≥ 3% of MM tumors; (ii) genes that were cross-validated in at least two of the considered genomic datasets. An *in silico* validation of the designed gene panel in the three aforementioned patients cohorts resulted in the recovery of at least one clonal mutation in 68% (95% confidence interval [CI]: 58 to 76) of MM cases.

CAPP-seq library preparation and ultra-deep NGS

The gene panel was analyzed in plasma cfDNA, and for comparative purposes to filter out polymorphisms, in normal gDNA from the paired granulocytes as source of germline material. The gDNA from the paired CD138+ purified plasma cells from BM aspiration was also investigated in the same cases to assess the accuracy of plasma cfDNA genotyping. Tumor and germline gDNA (median 400 ng) were sheared through sonication before library construction to obtain 200-bp fragments. Plasma cfDNA, which is naturally fragmented, was used (average: 59 ng; median: 48 ng; range: 0.05-400 ng) for library construction without additional fragmentation. Targeted ultra-deep-next generation sequencing was performed on plasma cfDNA, tumor and germline gDNA by using the CAPP-seq approach, which has been validated for plasma cfDNA genotyping [6]. The NGS libraries were constructed using the KAPA Library Preparation Kit (Kapa Biosystems) and hybrid selection was performed with the custom SeqCap EZ Choice Library (Roche NimbleGen). The manufacturer's protocols were modified as previously reported [6]. Multiplexed libraries were sequenced using 300-bp paired-end runs on an Illumina MiSeq sequencer. Each run included 24 multiplexed samples in order to allow >2000x coverage in >80% of the target region.

Bioinformatic pipeline for variant calling

Mutation calling in plasma cfDNA was performed separately and in blind from mutation calling in tumor gDNA from purified PCs. We deduped FASTQ sequencing reads by utilizing FastUniq v1.1. The deduped FASTQ sequencing reads were locally aligned to the hg19 version of the human genome using BWA v.0.6.2, and sorted, indexed and assembled into a mpileup file using SAMtools v.1. The aligned reads were processed with mpileup. Single nucleotide variations and indels were called in plasma cfDNA vs normal gDNA, and tumor gDNA vs normal gDNA, respectively, by using the somatic function of VarScan2 (a minimum Phred quality score of 30 was imposed). The variant called by VarScan 2 were annotated by using SeattleSeq Annotation 138. Variants annotated as SNPs, intronic variants mapping >2 bp before the start or after the end of coding exons, and synonymous variants were filtered out. To filter out variants below the base-pair resolution background frequencies, the Fisher's exact test was used to test whether the variant frequency called by VarScan 2 in cfDNA or tumor gDNA, respectively, was significantly higher from that called in the corresponding paired germline gDNA, after adjusting for multiple comparison by Bonferroni test (Bonferroni-adjusted P=4.03252e-7). To further filter out systemic sequencing errors, a database containing all background allele frequencies in all the specimens analyzed was assembled. Based on the assumption that all background allele fractions follow a normal distribution, a Z-test was employed to test whether a given variant differs significantly in its frequency from typical DNA background at the same position in all the other DNA samples, after adjusting for multiple comparison by Bonferroni. Variants that did not pass this filter were not further considered. Variant allele frequencies for the resulting candidate mutations and the background error rate were visualized using IGV (see Supplementary Figure S5 for a representative example).

Statistical analysis

The sensitivity and specificity of plasma cfDNA genotyping were calculated in comparison with tumor gDNA genotyping as the gold standard. The analysis were performed with the Statistical Package for the Social Sciences (SPSS) software (Chicago, IL) and with R statistical package (http://www.r-project.org).

Supplementary Table S1. Patients' characteristics

ID	*Age*	*Gender*	*Diagnosis*	*Phase*	*% of PCs in BM biopsy*	*Monoclonal component*	*FLC ratio*	*ISS stage*	*del(13q)*	*del(17p)*	*t(4;14)*	*t(14;16)*	*t(11;14)*	*HD*	*1p loss*	*1q gain*
1	46	F	MM	ND	50	Micromolecular λ	λ/κ FLC = 753	3	neg	neg	neg	neg	pos	neg	neg	neg
2	52	M	MM	ND	90	IgGκ	κ/λ FLC = 708	3	pos	neg	neg	neg	neg	pos	neg	pos
3	70	M	MM	ND	30	IgGκ	κ/λ FLC = 104	1	neg	pos	neg	neg	neg	pos	neg	neg
4	53	M	MM	RR	25	IgAκ	κ/λ FLC = 3	2	n.d.	n.d.	n.d.	n.d.	n.d.	pos	neg	pos
5	56	M	MM	ND	80	IgGκ	κ/λ FLC = 36	3	neg	neg	neg	neg	neg	neg	pos	pos
6	66	M	MM	ND	28	IgAλ	λ/κ FLC = 182	1	neg	neg	neg	neg	neg	pos	neg	pos
7	46	F	MM	RR	30	Micromolecular λ	λ/κ FLC > 27000	1	pos	neg	neg	neg	neg	pos	neg	pos
8	52	F	MM	ND	45	Micromolecular λ	λ/κ FLC = 446	1	pos	neg	neg	neg	neg	pos	neg	pos
9	76	M	sMM	ND	55	IgGκ	κ/λ FLC = 4	n.a.	neg	neg	neg	neg	neg	pos	neg	pos
10	76	F	MM	ND	30	IgAλ	λ/κ FLC = 6	1	n.d.	neg	neg	neg	neg	neg	neg	pos
11	54	M	MM	ND	40	IgAκ	κ/λ FLC = 19	2	pos	neg	neg	neg	neg	pos	neg	neg
12	77	M	MM	ND	30	IgGκ	κ/λ FLC = 17	2	neg	neg	neg	neg	neg	pos	neg	neg
13	64	M	MM	ND	60	IgAλ	λ/κ FLC = 210	3	n.d.	neg	neg	neg	neg	neg	neg	pos
14	61	F	sMM	ND	55	IgGκ	κ/λ FLC = 29	n.a.	pos	neg	neg	neg	neg	pos	neg	neg
15	68	M	MM	ND	70	IgAλ	λ/κ FLC = 65	3	pos	pos	pos	neg	neg	neg	neg	pos
16	76	F	sMM	ND	18	IgGκ	κ/λ FLC = 44	n.a.	n.d.	neg	neg	neg	neg	pos	neg	neg
17	59	M	MM	ND	90	Micromolecular λ	λ/κ FLC = 129	3	pos	pos	neg	neg	pos	pos	pos	neg
18	68	F	MM	ND	40	IgGλ	λ/κ FLC = 338	1	n.d.	neg	neg	neg	neg	neg	neg	pos
19	64	F	MM	RR	65	IgGκ	κ/λ FLC = 15	2	neg	neg	neg	neg	pos	neg	neg	neg
20	82	F	MM	ND	11	IgGκ	κ/λ FLC = 21	1	n.d.	neg	neg	neg	pos	neg	neg	neg
21	59	F	sMM	ND	10	IgGλ	λ/κ FLC = 5	n.a.	n.d.	neg	neg	neg	pos	neg	neg	neg
22	78	M	MGUS	ND	8	IgMκ	κ/λ FLC = 3	n.a.	neg	neg	neg	neg	pos	neg	neg	neg
26	47	F	MM	ND	50	IgAλ	λ/κ FLC = 6	1	n.d.	neg	neg	neg	neg	neg	neg	neg
27	71	M	MM	ND	28	IgGκ	κ/λ FLC = 109	1	n.d.	neg	neg	neg	neg	neg	neg	pos
28	68	M	sMM	ND	38	IgAλ	λ/κ FLC = 23	n.a.	n.d.	neg	neg	neg	neg	pos	neg	neg
29	69	M	MM	ND	70	IgGκ	κ/λ FLC = 200	2	n.d.	neg	neg	neg	neg	neg	neg	pos
30	50	F	MM	ND	60	IgAκ	κ/λ FLC = 108	1	n.d.	neg	neg	neg	neg	neg	neg	neg
31	61	F	MGUS	ND	7	IgGκ	κ/λ FLC = 4	n.a.	neg	neg	neg	neg	neg	pos	neg	neg

Abbreviations: F, female; M, male; MM, multiple myeloma; sMM, smoldering multiple myeloma; MGUS, monoclonal gammopathy of undetermined significance; ND, newly diagnosed; RR, relapsed/refractory; BM, bone marrow; FLC, free light-chain; ISS, International Staging System; n.a., not applicable; n.d., not determined; HD, hyperdiploidy.

Supplementary Table S2. Target region

Gene	*chromosome*	*coding exon start plus splice site (2bp)*	*coding exon stop plus splice site (2bp)*
NRAS	chr1	115251156	115251277
		115252188	115252351
		115256419	115256601
		115258669	115258783
FAM46C	chr1	118165491	118166666
CCND1	chr11	69456082	69456281
		69457797	69458016
		69458598	69458761
		69462760	69462912
		69465884	69466050
KRAS	chr12	25368375	25368496
		25378546	25378709
		25380166	25380348
		25398206	25398318
DIS3	chr13	73333933	73334018
		73334665	73334791
		73335499	73335661
		73335782	73335954
		73336059	73336277
		73337587	73337747
		73340108	73340198
		73342921	73343052
		73345040	73345128
		73345217	73345285
		73345931	73346036
		73346295	73346415
		73346829	73346979
		73347820	73347961
		73348082	73348199
		73349347	73349515
		73350061	73350232
		73351556	73351633
		73352323	73352520
		73354982	73355143
		73355741	73355970
TRAF3	chr14	103336537	103336785
		103338252	103338307
		103341959	103342067
		103342693	103342864
		103352524	103352608
		103355895	103355973
		103357660	103357756
		103363596	103363740
		103369590	103369768
		103371548	103372121
CYLD	chr16	50783610	50784115
		50785513	50785819
		50788228	50788337
		50810088	50810190
		50811734	50811854
		50813574	50813957
		50815155	50815324
		50816234	50816379
		50818238	50818364
		50820764	50820859
		50821695	50821765
		50825467	50825603
		50826506	50826618
		50827455	50827577
		50828121	50828341
		50830233	50830419

Supplementary Table S2. *(continued)*

Gene	*chromosome*	*coding exon start plus splice site (2bp)*	*coding exon stop plus splice site (2bp)*
TP53	chr17	7572927 7573925 7576851 7577017 7577497 7578175 7578369 7579310 7579698 7579837	7573010 7574035 7576928 7577157 7577610 7578291 7578556 7579592 7579723 7579912
SP140	chr2	231090560 231101796 231102926 231106117 231108444 231109701 231110576 231112629 231113598 231115694 231118029 231120165 231134245 231134550 231135299 231149059 231150465 231152605 231155173 231157359 231158984 231162134 231174637 231175456 231175867 231176165 231177299	231090620 231101977 231103098 231106204 231108528 231109797 231110657 231112782 231113685 231115778 231118134 231120249 231134335 231134670 231135356 231149128 231150549 231152683 231155281 231157505 231159035 231162179 231174756 231175568 231175948 231176312 231177399
EGR1	chr5	137801451 137802444	137801759 137803770
IRF4	chr6	393153 394819 395845 397106 398826 401422 405016 407453	393370 395009 395937 397254 398937 401779 405132 407598
PRDM1	chr6	106534429 106536074 106543488 106547173 106552698 106554244 106554784	106534472 106536326 106543611 106547429 106553810 106554376 106555361

Supplementary Table S2. *(continued)*

Gene	*chromosome*	*coding exon start plus splice site (2bp)*	*coding exon stop plus splice site (2bp)*
BRAF	chr7	140434397	140434572
		140439610	140439748
		140449085	140449220
		140453073	140453195
		140453985	140454035
		140476710	140476890
		140477789	140477877
		140481374	140481495
		140482819	140482959
		140487346	140487386
		140494106	140494269
		140500160	140500283
		140501210	140501362
		140507758	140507864
		140508690	140508797
		140534407	140534674
		140549909	140550014
		140624364	140624503
ZNF462	chr9	109685665	109685886
		109686412	109692042
		109692804	109692972
		109694725	109694832
		109697782	109697904
		109701195	109701390
		109734284	109734555
		109736416	109736556
		109746465	109746692
		109765573	109765709
		109771824	109771951
		109773102	109773311

Absolute chromosome coordinates are based on the hg19 version of the human genome assembly.

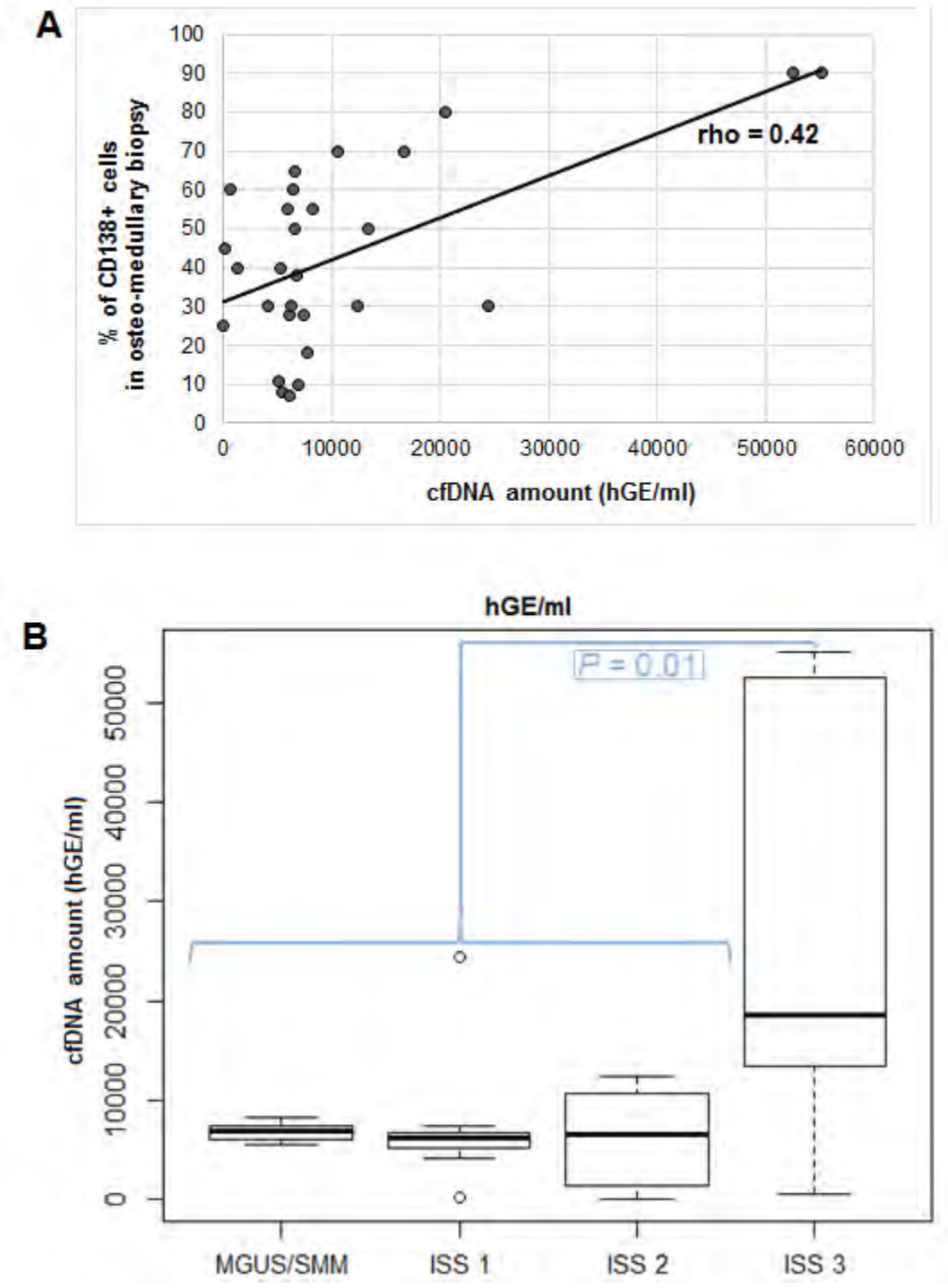
Supplementary Table S3. Percentage of target region covered ≥1000X and ≥2000X in distinct patient samples.

ID	*Sample*	*Target Region Coverage (%)*	
		≥ 1000X	*≥ 2000X*
	GL	100.0	99.5
ID1	PCS	99.6	98.3
	PL	99.4	97.9
	GL	99.5	99.5
ID2	PCS	99.4	96.9
	PL	97.8	83.5
	GL	98.7	85.9
ID3	PCS	99.9	99.5
	PL	98.9	95.8
	GL	98.3	43.0
ID4	PCS	100.0	100.0
	PL	97.6	58.5
	GL	97.7	63.4
ID5	PCS	100.0	99.6
	PL	98.2	93.3
	GL	100.0	99.5
ID6	PCS	99.5	99.5
	PL	99.4	97.4
	GL	99.6	99.5
ID7	PCS	99.6	98.8
	PL	99.5	97.9
	GL	99.5	98.6
ID8	PCS	99.5	99.5
	PL	98.6	91.9
	GL	99.5	98.8
ID9	PCS	99.5	98.9
	PL	99.4	97.8
	GL	99.5	99.0
ID10	PCS	99.3	96.9
	PL	99.5	98.5
	GL	99.5	98.7
ID11	PCS	99.5	99.3
	PL	99.3	95.5
	GL	99.5	99.0
ID12	PCS	99.5	99.1
	PL	94.5	21.9
	GL	99.5	98.9
ID13	PCS	99.5	98.7
	PL	98.7	86.0
	GL	99.5	98.8
ID14	PCS	99.6	98.8
	PL	99.4	96.5

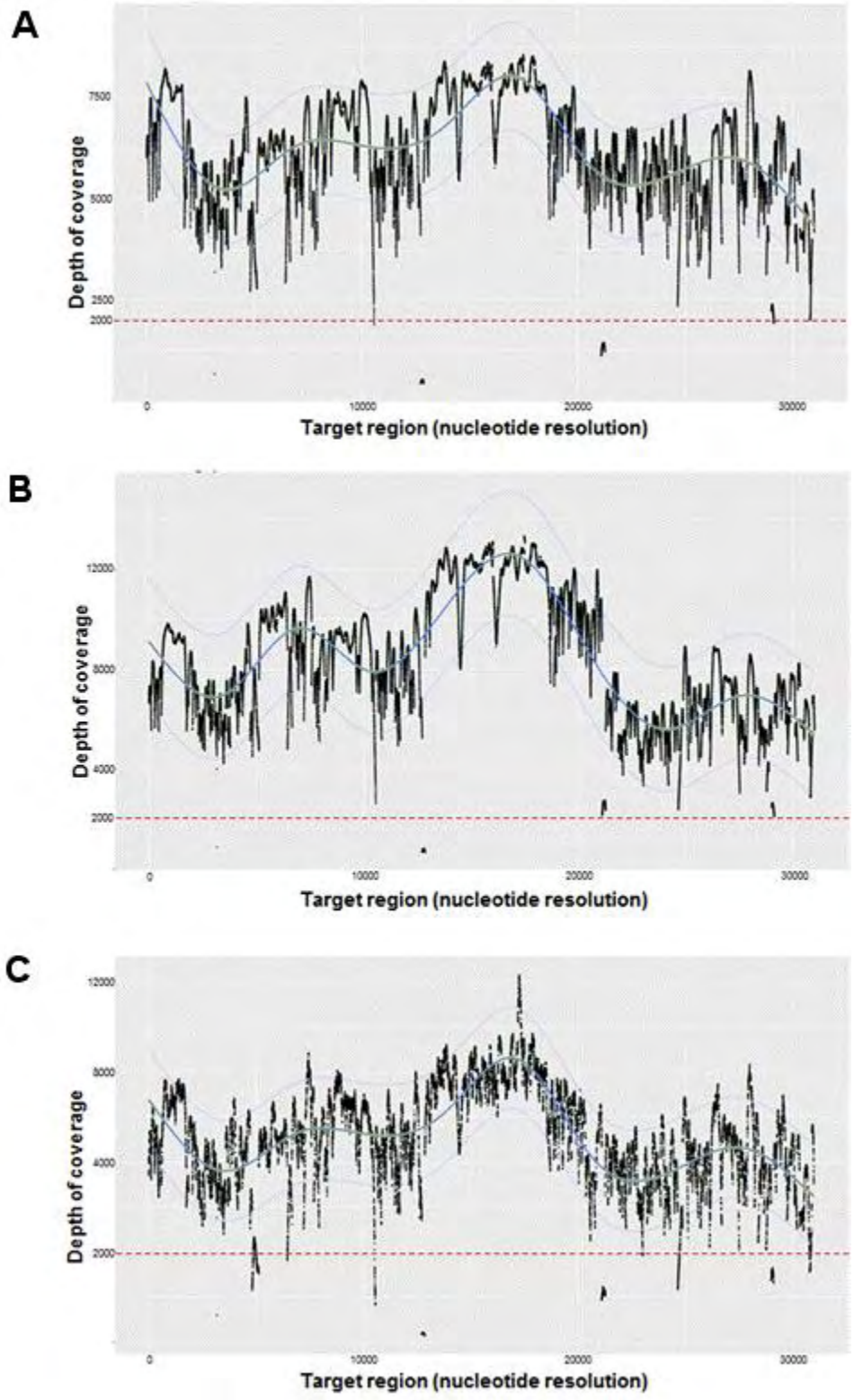
Supplementary Table S3. *(continued)*

ID	*Sample*	*Target Region Coverage (%)*	
		≥ 1000X	*≥ 2000X*
	GL	99.5	98.7
ID15	PCS	99.5	98.6
	PL	99.4	98.0
	GL	99.5	98.9
ID16	PCS	99.5	98.3
	PL	97.3	62.8
	GL	98.6	97.5
ID17	PCS	99.5	95.2
	PL	98.5	97.0
	GL	98.7	96.3
ID18	PCS	94.7	84.0
	PL	98.7	97.8
	GL	98.4	92.0
ID19	PCS	96.4	74.9
	PL	98.7	97.5
	GL	98.6	96.5
ID20	PCS	97.7	95.2
	PL	98.7	96.8
	GL	98.5	94.9
ID21	PCS	98.0	94.9
	PL	98.6	96.7
	GL	98.7	97.5
ID22	PCS	98.7	97.8
	PL	98.7	96.5
	GL	98.9	98.2
ID26	PCS	99.5	99.5
	PL	99.4	97.1
	GL	99.5	99.0
ID27	PCS	99.5	98.8
	PL	99.5	98.8
	GL	99.5	99.0
ID28	PCS	99.5	99.5
	PL	98.7	90.3
	GL	98.2	94.5
ID29	PCS	99.6	99.5
	PL	99.6	99.4
	GL	99.5	98.9
ID30	PCS	100.0	99.5
	PL	99.5	98.1
	GL	99.6	98.9
ID31	PCS	99.6	98.9
	PL	99.4	98.3

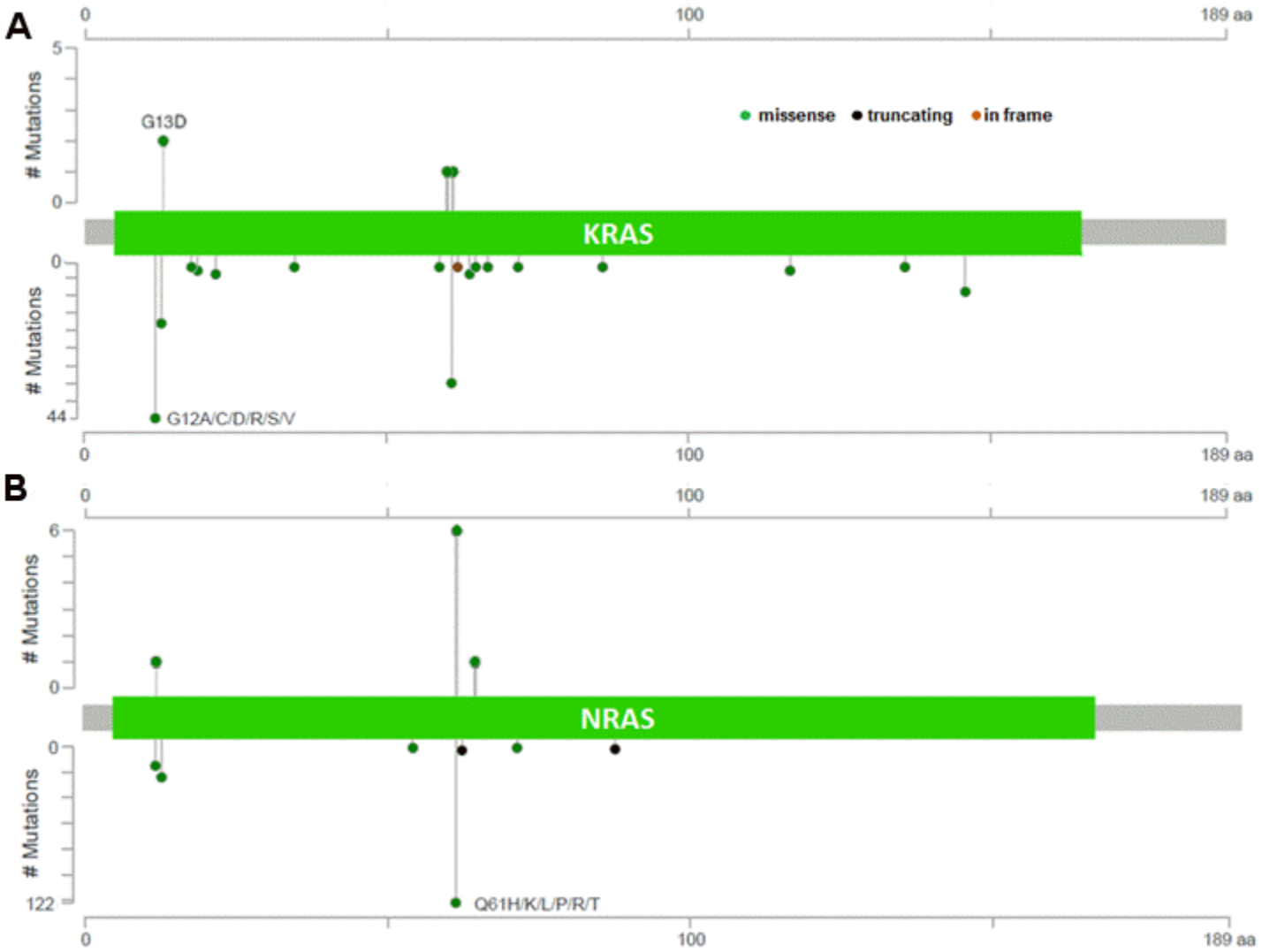
Abbreviations: GL, normal germline DNA from granulocytes; PCS, tumor genomic DNA from plasma cells; PL, cfDNA from plasma.



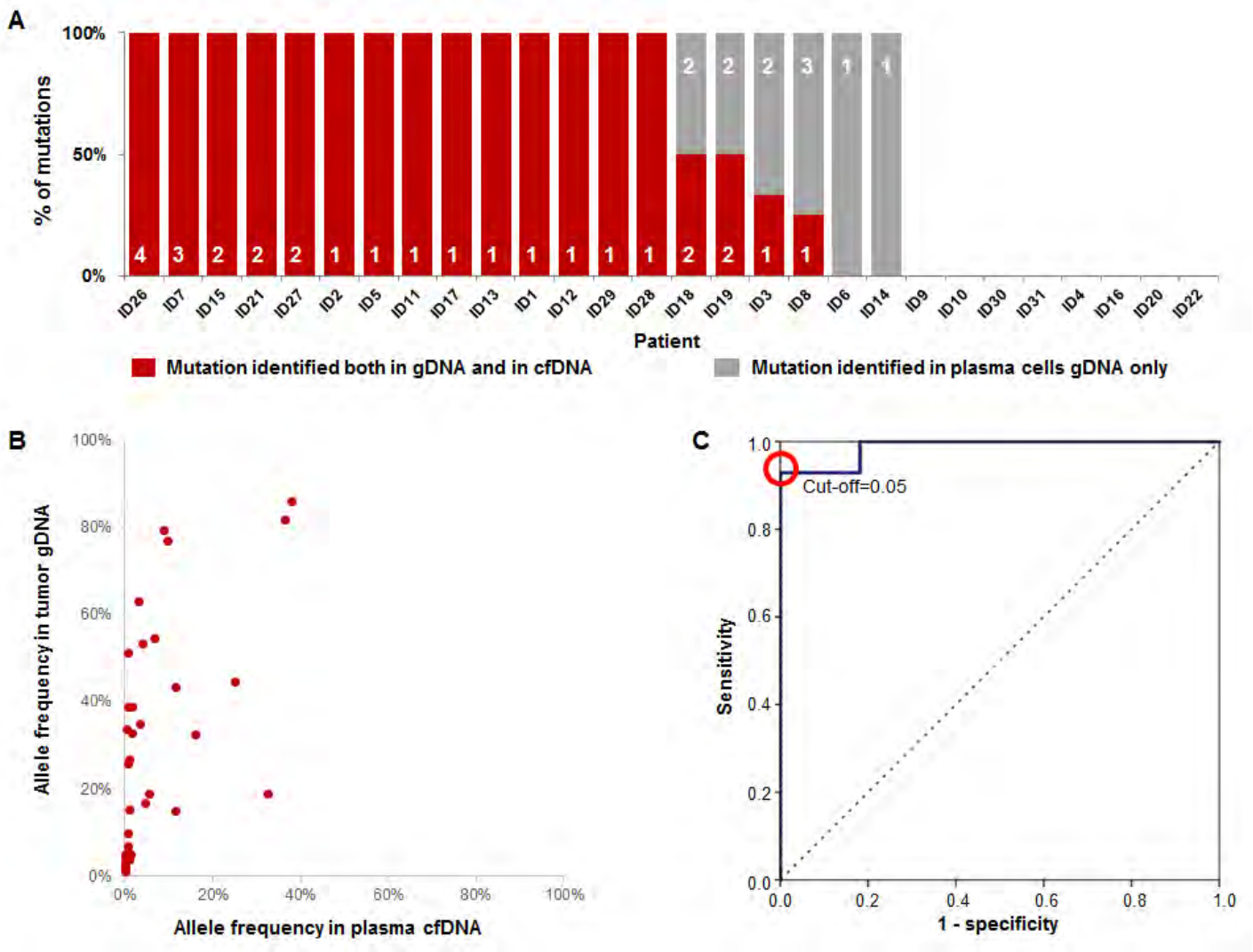
Supplementary Figure S1. (A) Correlation between cfDNA amount and bone marrow plasma cell infiltration. **(B)** cfDNA amount according to diagnosis/risk stratification: the levels of cfDNA are significantly higher in MM patients at ISS stage 3 compared with MGUS/SMM samples and MM cases at ISS stages 1-2 (*P*=0.01; Mann-Whitney test).



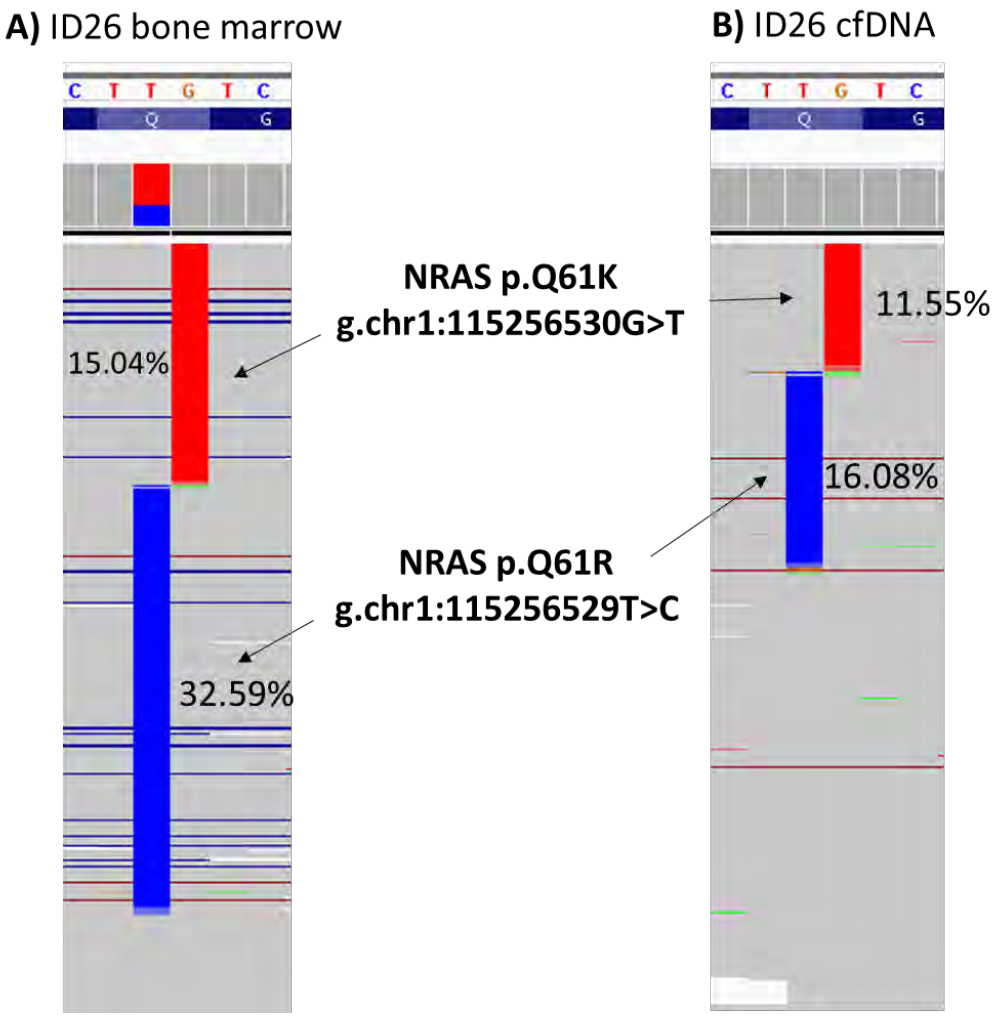
Supplementary Figure S2. Coverage across the target region. Depth of coverage (y axis) across the target region (x axis) by CAPP-seq of **(A)** gDNA from the germline (granulocytes) samples, **(B)** tumor gDNA from bone marrow plasma cells, and **(C)** plasma cfDNA. Each dot represents the sequencing depth on that specific position of the target region of one single individual sample. The solid blue line shows the median depth of coverage, while the dash blue lines show the interquartile range. The dashed red line shows the 2000X coverage.



Supplementary Figure S3. Prevalence and molecular spectrum of nonsynonymous somatic mutations discovered in plasma cfDNA. The most mutated genes are reported: (**A**) *KRAS* gene and (**B**) *NRAS* gene. The molecular spectrum of nonsynonymous somatic mutations identified in plasma cfDNA (in the upper part of the figure) compared with the molecular spectrum of nonsynonymous somatic mutations that have been detected in the tumor gDNA in published MM series and reported in the COSMIC database (version 81) [7] (in the lower part of the figure). Mutation maps were obtained through Mutation Mapper version 1.0.1. Color codes indicate the type of the mutations: truncating mutations include nonsense, frameshift deletion, frameshift insertion, splice site.



Supplementary Figure S4. Concordance between plasma cfDNA and tumor gDNA genotyping. **(A)** The fraction of tumor biopsy–confirmed mutations that were detected in plasma is shown. Patients are ordered by decreasing detection rates. The red portion of the bars indicates the prevalence of tumor biopsy–confirmed mutations that were detected in plasma cfDNA. The gray portion of the bars indicates the prevalence of tumor biopsy-confirmed mutations that were not detected in plasma cfDNA. **(B)** The mutation abundance in plasma cfDNA vs the mutation abundance in tumor gDNA is comparatively represented in the scatter plot for each identified variant. **(C)** ROC analysis illustrating the performance of gDNA genotyping in discriminating the ability of cfDNA genotyping to detect biopsy-confirmed mutations according to the variant allele frequency of mutations in tumor gDNA.



Supplementary Figure S5. Visualization of deep sequencing data in BM gDNA (A) and cfDNA (B) of patient ID26 by Integrated Genome Viewer software. Two adjacent base substitutions affecting the same codon and originating distinct *NRAS* p.Q61R and p.Q61K mutations are shown. The lack of sequencing reads carrying both mutations suggested that these two substitutions likely involved different tumor subclones. Reads were sorted by base at chr1:115,256,529 locus and then again sorted by base at chr1:115,256,530 locus. Red bars show G>T substitutions at the chr1:115,256,530 locus. Blue bars show T>C substitution at the chr1:115,256,529 locus.

References

1. Fabris S, Agnelli L, Mattioli M, et al. Characterization of oncogene dysregulation in multiple myeloma by combined FISH and DNA microarray analyses. Genes, chromosomes & cancer. 2005;42(2):117-27.
2. Mattioli M, Agnelli L, Fabris S, et al. Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma. Oncogene. 2005;24(15):2461-73.
3. Bolli N, Avet-Loiseau H, Wedge DC, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nature communications. 2014;5:2997.
4. Kortuem KM, Braggio E, Bruins L, et al. Panel sequencing for clinically oriented variant screening and copy number detection in 142 untreated multiple myeloma patients. Blood Cancer J. 2016;6:e397.
5. Lohr JG, Stojanov P, Carter SL, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell. 2014;25(1):91-101.
6. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nature medicine. 2014;20(5):548-54.
7. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic acids research. 2017;45(D1):D777-D83.