MCL1 gene co-expression module stratifies multiple myeloma and predicts response to proteasome inhibitor-based therapy

Ayaz Ali Samo[1], Jiuyi Li[1], Min Zhou[2], Yingyu Sun[1], Yuan Yang[1], Yunqiu Zhang[1], Jing Li[4], Mark van Duin[5], Xuzhang Lu[3, *], and Xiaolong Fan[1, *]

[1]Beijing Key Laboratory of Gene Resource and Molecular Development, Laboratory of Neuroscience and Brain Development, Beijing Normal University, Beijing, China.

[2]Department of Hematology, Changzhou No. 3 People's Hospital, Changzhou, China.

[3]Department of Hematology, Changzhou No. 2 People's Hospital, The Affiliated Hospital of Nanjing Medical University, Changzhou, China.

[4]Yichuang Biotechnology Industry Research Institute, Beijing Yizhuang Biomedical Park, Kechuang Street No. 6, Daxing district, Beijing, China

[5]Department of Hematology, Erasmus MC Cancer Center, Rotterdam, The Netherlands.

Running title: MCL1 module-based molecular classification of multiple myeloma

[*]Corresponding Authors: Xiaolong Fan, Beijing Key Laboratory of Gene Resource and Molecular Development, Laboratory of Neuroscience and Brain Development, Beijing Normal University, Beijing, 100875 China; e-mail: xfan@bnu.edu.cn;

Xuzhang Lu, Department of Hematology, Changzhou No. 2 People's Hospital, The Affiliated Hospital of Nanjing Medical University, Changzhou, 213003 China; e-mail: luxuzhang2008@163.com.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/gcc.2

Abstract

Multiple myeloma (MM) is the second most common hematologic cancer, characterized by abnormal accumulation of plasma cells in the bone marrow. The extensive biological and clinical heterogeneity of MM hinders effective treatment and etiology research. Several molecular classification systems of prognostic impact have been proposed, but they do not predict the response to treatment nor do they correlate to plasma cell development pathways. Here we describe the classification of MM into two distinct subtypes based on the expression levels of a gene module co-expressed with *MCL1* (MCL1-M), a regulator of plasma cell survival. The classification system enabled prediction of the prognosis and the response to bortezomib-based therapy. Moreover, the two MM subtypes were associated with two different plasma cell differentiation pathways (enrichment of a preplasmablast signature versus aberrant expression of B cell genes). 1q gain, harboring 56 of the 87 MCL1-M members including *MCL1*, was found in about 80% of the MM with upregulated MCL1-M expression. Clonal analysis showed that 1q gain tended to occur as an early clonal event. Members of MCL1-M captured both MM cell-intrinsically acting signals and the signals regulating the interaction between MM cells with bone marrow microenvironment. MCL1-M members were co-expressed in mouse germinal center B cells. Together, these findings indicate that MCL1-M may play previously inadequately recognized, initiating role in the pathogenesis of MM. Our findings suggest that MCL1-M signature-based molecular clustering of MM constitutes a solid

framework toward understanding the etiology of this disease and establishing personalized care.

Article Summary: A pathogenic mechanism-guided molecular classification would facilitate treatment decision and etiology research of multiple myeloma. Based on the expression levels of a gene module co-expressed with MCL1, we have established a classification scheme assigning multiple myeloma into two subtypes with distinct prognosis, treatment responses and pathogenic backgrounds.

Introduction

Multiple myeloma (MM) represents the second most common hematologic malignancy, characterized by an abnormal accumulation of clonal antibody-producing plasma cells in bone marrow (BM) [1]. With the introduction of proteasome-inhibitors such as bortezomib and immunomodulatory drugs, outcome in MM has been significantly improved. However, MM remains incurable [1,2]. MM is characterized by a high degree of variabilities in overall survival (OS) and responses to treatment with individual drugs or their combinations. Moreover, the biological mechanisms underlying these clinical differences hindering the development of personalized care are inadequately understood.

To facilitate treatment decisions and improve our understanding of tumor biology, a simple and reliable molecular classification system is needed. Several such systems have been proposed on the basis of gene expression profiling (GEP). Bergsagel et al. [3] identified eight MM subtypes with distinct cyclin D expression and translocation patterns. Using unbiased hypothesis-free transcriptome analysis, Zhan et al. and Broyl et al. identified 7–10 molecular subtypes of MM [4,5]. This enabled the division of MM into high-risk or low-risk groups in studies with patients receiving total therapy 2 (TT2) [5] or TT3 [6]. Moreover, GEP signatures correlated with the prognosis of MM have been described (for example UAMS-70 and the related UAMS-17, UAMS-80, IFM-15, Millennium-100 and EMC-92) [6-10]. Further, gene proliferation index (GPI-50), MRC-IX-6, and centrosome amplification index were found to be

prognostic for MM patients [11-13]. Recently, Palumbo et al. proposed a revised international staging system (R-ISS) based on the serum β2-microglobulin level, the serum lactate dehydrogenase (LDH) level, and high-risk chromosome abnormalities [including del(17p), t(4;14), and t(14;16) translocations], which also provides prognostic information [14]. These studies were important steps forward in establishing an objective molecular classification for MM. However, three issues remain unaddressed: (1) the capacity to predict the treatment response [15,16], (2) the correlation between subtypes and specific pathways of plasma cell development [17], and (3) the relatedness of classifiers to MM pathogenesis.

To address these questions, we explored whether gene co-expression modules around key signaling pathways conserved between germinal centre (GC) formation and MM pathogenesis might enable the molecular classification of MM. We were specifically interested in the dysregulation of gene networks controlling B cell differentiation to plasma cells in GC because these may play a crucial role in MM initiation [18]. Here we describe the identification of genes consistently co-expressed with *Myeloid cell leukemia 1* (*MCL1*-module, MCL1-M) in MM, and their application in classifying MM into two major subtypes that are characterized by differences in prognosis, response to bortezomib-based treatment, and correlation to plasma cell development. These findings pave the way for establishing personalized care and improving our understanding of the etiology of MM.

Materials and Methods

The data of normalized global transcriptome, array CGH (aCGH) and whole exome sequencing for bone marrow $CD138^{+}$ plasma cells of newly diagnosed and untreated MM, and the corresponding clinical characteristics and treatment responses from GSE2658 [5], GSE19784 [4,19], GSE26863 [20], GSE29023 [21], and Lohr et al. [22] were retrieved from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/), or the MMRF website (https://research.themmrf.org/), or dbGap (accession no. phs000348.v2. p1). Pearson correlation coefficient analysis in GSE2658 [5] was used to identify gene co-expression modules with the regulators of GC formation [23,24] as the seed genes. MCL1-M contained the top 106 probe sets most correlated with *MCL1* expression (200797_s_at). MCL1-M, complemented with the MCL1-M low MM identifier, was used in unsupervised consensus clustering [25] to define stable and robust clusters of MCL1-M low or MCL1-M high MM.

Details regarding survival analysis, correlation to plasma cell development, the frequency and clonality of SCNAs and driver mutations in the MCL1-M signature-defined MM subtypes, and Cox regression analysis are provided in *Supplementary Materials*.

Results

Identification of MCL1-M and molecular clustering of MM based on MCL1-M signature

To identify gene network(s) conserved in plasma cell development, we used transcriptome data from the MM data set GSE2658 [5] to construct gene co-expression modules around key regulators of GC formation, including BCL6, BLIMP1, IRF4, IRF8, MCL1, MTA3, NF-kB, PAX5, and XBP1 [23,24]. Unsupervised hierarchical clustering in five data sets (GSE2658 [5], GSE19784 [4,19], MMRF, GSE26863 [20] and GSE29023 [21], **Table S1**) showed that MM samples were clustered into two subtypes with high or low expression of MCL1-M and XBP1 module (XBP1-M); the expression of the other co-expression modules did not enable clustering of MM samples. However, frequent alteration of *XBP1* in MM genome has not been reported, and the signature of XBP1-M, with members predominantly residing in protein synthesis pathway, did not enable a clustering of MM samples into subtypes with prognostic and predictive features. Therefore, the remaining investigations focused on the MCL1-M signature.

MCL1 is essential for germinal center formation and memory B cell generation [26], and is indispensable for the maintenance of long-lived plasma cells in the BM [27]. Frequent gain of the chromosomal region 1q21, where *MCL1* is located, was reported to be associated with MM progression and short survival of the patients [28]. Of the 87 known genes in MCL1-M, 63 are located on 1q (**Figure S1** and **Table S2)**; a gene dosage-dependent expression was found for 39 of these 63 in MM with "MCL1-M

high" signature (**Table S2**). With the exceptions of *MCL1*, *IL6R*, *LDH* (also known as *LDHA*), *PSMB4*, *PSMD4*, *UBE2Q1*, *ANP32E*, *ATF3*, *CCT3*, and *COPA*, which are known to be involved in myeloma pathogenesis [7,29-31], the remaining MCL1-M members have not been extensively studied in MM biology. MCL1-M genes are enriched in Reactome pathways of antigen processing and presentation functions, cytokine signaling pathways and interferon response signaling [32] **(Table S3)**. To facilitate stable clustering, we also identified a set of 46 known genes enriched in the MCL1-M low MM (named MCL1-M low MM identifier, **Figure S2** and **Table S4**). The MCL1-M low MM identifier included important regulators of B cell identity (PAX5 [33]), B cell receptor signalling (PAK1, PLCG2, STAP1 [34]), B cell survival (TNFRSF13C (BAFFR) [35]), Wnt signalling (CCND1, FRZB, TRABD2A [36]) and NF-kB signalling (IRAK2 [37])**.** Based on the signatures of the MCL1-M and MCL1-M low MM identifier, we performed unsupervised consensus clustering [25] to robustly assign 1650 newly diagnosed MM in the above-mentioned data sets into MCL1-M high or MCL1-M low MM subtypes, with 44.7% of all MM in the MCL1-M high subtype and the remaining in the MCL1-M low subtype (**Figure S3** and **Figure S4**).

Prognostic and predictive features of MCL1-M signature-based MM clustering

Since prognosis is an integral outcome determined by disease biology, treatment regimen, and length of follow-up, we applied reiterative survival analysis to assess the prognostic distinctions between patients with MCL1-M high or MCL1-M low MM. Patients in GSE19784 were randomly assigned to VAD (vincristine, doxorubicin, and dexamethasone) or PAD (bortezomib, doxorubicin, and dexamethasone) treatment

arms in the phase III HOVON-65/GMMG-HD4 trial [19]. In the follow-up years five till eight, patients with MCL1-M high MM showed significant poorer OS compared with patients with MCL1-M low MM (**Figure 1A** and **1B**). Among the patients treated with TT2 regimen in GSE2658, the same prognostic trend was found in follow-up years three and four (**Figure 1C** and **1D**). MMRF data set included patients treated with diverse regimens, the poor prognosis in patients with MCL1-M high MM was found between the follow-up years two to four (**Figure 1D** and **1F**). These findings show that due to the difference in treatment regimens, the prognostic impact between MCL1-M high and MCL1-M low MM groups was manifested at different follow-up periods.

We further integrated MCL1-M clustering-based survival analysis with treatment arms in GSE19784 [19]. Compared with the VAD treatment, PAD treatment improved progression-free survival (PFS) in patients of MCL1-M high MM in follow-up years five to eight (**Figure 2A** and **2B**) (median PFS 19 *v* 27 months; HR = 1.58; 95% CI = 1.07 to 2.42; $P = 0.02$), the improvement in OS was however only observed in follow-up years three to four (**Figure 2C** and **2D**). In contrast, in MCL1-M low MM patients, no difference in PFS or OS was observed between the VAD and PAD treatment arms (**Figure 2E** and **2F**). The superior outcome with bortezomib-based treatment was reported in patients with del(13q14) and del(17p13) [19,38]. Though more MCL1-M high MM harbored del(13q14) ($P < 0.0001$, x^2 test, **Table S5**), del(17p13) occurred at comparably low frequencies in both MCL1-M high and MCL1-M low MM.

Collectively, these findings show that MM with high MCL1-M signature is characterized by poorer prognosis and might benefit from bortezomib-based treatment during the induction and maintenance phases.

Comparison of MCL1-M signature-based classification with other classification schemes or MCL1 expression alone

We compared MCL1-M signature-based MM classification with the previously reported prognostic classification schemes. Analyses of the β2-microglobulin, albumin, and LDH levels in the GSE19784 and MMRF MM samples in MCL1-M high and low groups demonstrated that MCL1-M clustering provides a different result from the R-ISS classification [14] **(Table S6)**. Comparison with the prognostic MM clusters defined by Zhan et al. [5] showed that the good prognostic CD-1, CD-2, and HY clusters predominantly contained MCL1-M low MM, and most of the MM in the poor prognostic PR, MS, and MF clusters showed the MCL1-M high signature **(Table S7)**. However, most good prognostic LB MM showed the MCL1-M high signature. In GSE19784, the LB cluster was not confirmed [4]. The assignments of the same MM samples into the 10 clusters reported in GSE19784 [4] did not correlate with the MCL1-M clustering **(Table S8)**. Unlike the distinct response to VAD/PAD treatment arms between patients with MCL1-M high or MCL1-M low MM, nine of the ten previously defined clusters (except the MF cluster) [4] did not show distinct response to treatment regimens with or without bortezomib (**Figure S5**).

The prognostic impact of *MCL1* expression has been controversial [39,40]. Based on the results shown in **Figure 1**, we used univariate Cox regression analysis to assess the prognostic effect of all MCL1-M members. In both GSE2658 and GSE19784 data sets, MCL1 expression alone was not prognostic, but the other 15 MCL1-M members, including key regulators of osteoclast formation (ANXA2 [41]), chromatin structure (ANP32E [42] and H3F3A [43]) and oncogenic transformation (TPM3 [44]) were strongly prognostic (**Table S9**). Thus, though the prognostic impact of individual MCL1-M members may vary, the expression level of the whole MCL1-M is a robust prognostic marker.

Plasma cell development-related transcriptomic features of MCL1-M signature-defined MM subtypes

To study the correlation of the MCL1-M signature-defined MM molecular subtypes with the B cell to plasma cell differentiation stages, we performed GSEA analysis to compare the enrichment of gene signatures characteristic of the different stages of human plasma cell development between MCL1-M high and MCL1-M low MM [45]. Human B cell differentiation into plasma cells can be divided into eight stages with distinct gene expression signatures: naïve B cells, centroblasts, centrocytes, memory B cells, preplasmablasts, plasmablasts, early plasma cells, and bone marrow plasma cells [46-49]. In all five data sets analyzed, gene signatures characteristic of pre-plasmablasts (as defined by the molecular atlas of human B cell differentiation [46]) was consistently enriched in MCL1-M high MM **(Figure 3A, Figure S6** and **Table**

S10). Further, large subsets of both MCL1-M and MCL1-M low MM identifier showed enriched expression in the GC B cells of the murine transcriptome data set for B cell to plasma cell development, GSE60927 [50] (**Figure 3B** and **3C**). MCL1-M low MM identifier genes, including PAX5, showed high expression in B cell populations (**Figure 3C**). MCL1-M low MM showed relatively enriched expression of the PAX5 targets CD19, FCER2(CD23), CD40, CD79A and STAP1 (**Figure S6**). Together, the enrichment of a pre-plasmablast signature in MCL1-M high MM and aberrant expression of B cell genes in MCL1-M low MM suggest distinct involvement of plasma cell differentiation pathways in these two MM subtypes.

Characteristic somatic copy number alterations (SCNA) in MCL1-M high and low MM subtypes

To characterize the genomic abnormalities in the MCL1-M high and low MM subtypes, we analyzed fluorescent *in situ* hybridization (FISH) data of cytogenetic abnormalities in GSE19784 samples [4,19]. 1q21 gain was identified in 65% and 16% of the MCL1-M high and MCL1-M low MM, respectively (P = 0.0001, Chi-square test, **Table S5)**. Conversely, 68% of the MCL1-M low MM and 38% of the MCL1-M high MM harbored trisomies of chromosome 9, 11 and 15 (P = 0.0002, Chi-square test, **Table S5)**.

We further analyzed myeloma-characteristic SCNAs in MCL1-M signature-defined MM subtypes using the GSE29023 (n = 115) and GSE26863 (n = 180) array CGH data. In both data sets, 18% of the MCL1-M low subtype harbored 1q

gain; but up to 77% of the MCL1-M high MM harbored 1q gain. Further, high copy gain (defined by more than three copies of 1q) was found nearly exclusively in MCL1-M high MM **(Table S5** and **Figure S7)**. In both GSE29023 and GSE26863, the ratio between the MCL1-M expression average and the expression average of MCL1-M low MM identifier in the MCL1-M high MM without 1q gain was significantly higher than that in all MCL1-M low MM **(Figure S8)**, indicated that mechanisms other than 1q gain, such as bone marrow microenvironment [31], might contribute to up-regulated expression of MCL1-M genes in those MCL1-M high MM samples without 1q gain. In all three data sets analyzed, 13q14 loss occurred more frequently in MCL1-M high MM, whereas trisomies of chromosome 9, 11 and 15 were more frequent in MCL1-M low MM (**Table S5**).

Five other key genomic alterations (IgH split and deletions of 1p, 8p, 16q, and 17p) occurred with comparable frequencies in MCL1-M high and MCL1-M low MM subtypes (**Table S5** and **Table S11**). Further, occurrence of the recently reported MM driver gene mutations *KRAS*, *NRAS*, *BRAF*, *FAM46C*, and *TP53* [22,51] was not clearly associated with either subtype (**Table S12**). These findings together demonstrate 1q gain as the most recurrent somatic genomic alterations in MCL1-M high MM.

1q gain as an early and clonal event during MM development

Previous reports of FISH analysis in unrelated patients have suggested 1q gain as a marker of MM progression [28,52]. To infer the temporal order of 1q gain and other characteristic MM genomic abnormalities, we employed BubbleTree [53] to perform

clonality analysis with the whole exome sequencing data for 112 MM samples reported by Lohr et al [22]. Among them, we identified thirty-nine samples harboring 1q gain, with 12 of them harboring high gain of 1q, for the clonality analysis (**Table S13**).

The number of clones with the SCNAs identified above ranged from 1 to 4; 18 MM samples harbored only one clone. Most of the major SCNAs appeared to be highly clonal. Thirteen, seven, and one MM samples harbored two, three and four clones, respectively; most SCNAs in samples with two multi-clones were clonal. 1q gain was clonal in 36 of the 39 MM samples analyzed. In three samples with subclonal 1q gain, the ploidy for 1q was 4 or 5 (**Table S13**). These samples might have initially gained one (clonal) 1q segment and subsequently gained additional (subclonal) 1q segments (**Figure S9** and **Table S14**). A representative sample (MMRC0571) with clonal 1q gain is shown in **Figure 4**. Three clones were identified. The triploid heterozygous 1q gain belonged to the major clone harboring most copy number variations (gain of 1q, trisomy of odd-numbered chromosomes, and loss of 13q). This major clone made up ca. 90% of all cells, which also harbored the *NRAS* mutation.

We further assessed the cancer cell fraction (CCF) containing the driver mutations in MM harboring the 1q gain. Only a fraction of MM samples with 1q gain harbored these mutations, and they were either clonal or subclonal (**Figure 4** and **Figure S10**). These findings suggest that 1q gain and other large SCNAs occurred earlier than the driver mutations, and nearly all neoplastic cells of patients with a 1q

gain harbored the 1q gain. Hence, MCL1-M members located in 1q may exhibit upregulated expression already in the early phases of MM development.

Discussion

The biological and clinical heterogeneity hinders elucidation of MM biology and treatment of MM patients. Our findings show that MCL1-M signature-based classification scheme assigns MM into two biologically and clinically separate subtypes. The MCL1-M contains a set of previously inadequately recognized, early acting drivers of MM pathogenesis that are consistently co-expressed with the key regulator of plasma cell survival, MCL1. The two molecular subtypes showed distinct prognoses. Patients with MCL1-M high MM responded to bortezomib-based treatment, whereas patients with MCL1-M low MM did not. The two MM subtypes may involve distinct pathways of plasma cell development and thus require different treatment strategies. These findings constitute a step toward establishing personalized care for MM patients.

The introduction of proteasome inhibitors and immunomodulatory drugs has significantly improved the median survival of MM patients. However, treatment responses in individual MM patients have been highly variable and unpredictable [19]. Previously reported GEP signatures have limited power in predicting treatment response [15]. Our analysis of MCL1-M signature in the phase III HOVON-65/GMMG-HD4 trial showed that only the patients with MCL1-M high MM benefited from bortezomib-based treatment, whereas those patients with MCL1-M low MM did not respond to bortezomib-based treatment. To the best of our knowledge, this is the first MM GEP signature enabling potential treatment decisions in newly diagnosed patients. Furthermore, whereas correlation to plasma cell development was unclear for

the previously defined MM molecular subtypes [4-6,8-13,54], MCL1-M signature-based MM molecular subtypes showed distinct involvement of plasma cell differentiation pathways. Our findings appear to be in contrary with a report by Leung-Hagesteijn et al. on the worse repose to bortezomib in MM with pre-plasmablast features [55]. The crucial differences between these two studies could be the selection of patients. The study by Leung-Hagesteijn et al. utilized the gene expression profiling data of bone marrow $CD138^+$ plasma cells and the treatment response data from the relapsed patients with 1 to 3 prior therapies [9]. In contrast, data from newly diagnosed and untreated MM were used in our study.

The unique treatment responses between the two MM subtypes may arise from the features of MCL1-M. MCL1-M captures both MM cell-intrinsically acting signals and the signals regulating the interaction between MM cells with bone marrow microenvironment. MCL1 itself is a key regulator of plasma cell survival [27]. Other MCL1-M members regulate chromatin structure (ANP32E [42] and H3F3A [43]) and oncogenic transformation (TPM3 [44]), as well as cytokine signaling (IL6R [56]) and osteoclast formation (ANXA2 [41]). Further, MCL1-M members are co-expressed in murine GC B cells, and frequently amplified in MM genome with a gene dosage-dependent expression. These findings suggest that an integral activity of MCL1-M may drive the development of MM.

Of the 87 known genes in MCL1-M, 63 genes (including *MCL1* itself) are located on 1q. Gain of this chromosomal region is a common finding in MM and is associated with MM progression and short patient survival [28,52]. To examine the temporal order

of 1q gain and the other key genomic abnormalities in MM, we employed new genomic analysis approaches to assess the clonal feature of the characteristic SCNAs in MM harboring 1q gain, using the exome sequencing data from a large number of MM samples. In the majority of the MM with high MCL1-M signature, 1q gain was clonal, indicating that the 1q gain might have occurred together with the other SCNAs during a short period of catastrophic genomic crisis at an early stage of MM development. The enrichment of a pre-plasmablast signature in MCL1-M high MM suggests that this genomic crisis occurred in a B cell clone during GC development. Unlike the clonal nature of 1q gain in nearly all MCL1-M high MM, driver mutations in *NRAS*, *KRAS*, *BRAF*, *FAM46C*, and *TP53* genes occurred in a small fraction of MM samples, and they were clonal or subclonal in individual MM samples, as was also reported previously [22,51]. Thus, 1q gain and the upregulation of MCL1-M may play an initiating role in MM pathogenesis whereas the relatively less frequent clonal or subclonal driver mutations [22,51] may have been acquired during MM progression.

Over 20% of the MCL1-M high MM did not harbor 1q gain. In such MM samples, signals derived from BM microenvironment may have contributed to high MCL1-M signature [31]. This coincides with the facts that 31 MCL1-M members are not located on 1q, and that MCL1 expression alone was not prognostic, but the expression levels of 15 other MCL1-M members, which are not all located on 1q, were prognostic.

Though the distinct treatment responses between the two MM subtypes require further validation in independent clinical trials, and mechanisms of better response to

bortezomib-based therapy in MCL1-M high MM are currently unknown, diagnosis of MM according to MCL1-M signature-based classification scheme could support the clinical trials aiming to define the efficacy of new proteasome inhibitors [57]. MCL1-M signature-based classification scheme may greatly facilitate the etiology studies in the context of plasma cell development and the search of subtype-specific therapies. In summary, our findings suggest that MCL1-M signature-based MM classification constitutes a solid framework toward establishing personalized care of MM patients and understanding of MM etiology.

Acknowledgments

We thank Robert Smail Jack, Claudia Berek, Qiong Yang, Ge Gao, and Dan Csontos for insightful comments on the project and manuscript. This work was supported by the National Natural Science Foundation of China (Grants 81472348, 81773015, and 31711530156), The National Key Research and Development Plan (Grant 2016YFC0902500), Beijing Cooperative Construction Project, the Erik Philip-Sörensen Foundation, and High-Level Medical Talents Training Project (NO. 2016CZLJ027) of Changzhou.

Disclosure: The authors have declared no conflicts of interest.

References

1. Raab MS, Podar K, Breitkreutz I, Richardson PG, Anderson KC. Multiple myeloma. Lancet. 2009;374(9686):324-39.
2. Corre J, Munshi N, Avet-Loiseau H. Genetics of multiple myeloma: another heterogeneity level? Blood. 2015;125(12):1870-6.
3. Bergsagel PL, Kuehl WM, Zhan F, Sawyer J, Barlogie B, Shaughnessy J, Jr. Cyclin D dysregulation: an early and unifying pathogenic event in multiple myeloma. Blood. 2005;106(1):296-303.
4. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. Blood. 2010;116(14):2543-53.
5. Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S, et al. The molecular classification of multiple myeloma. Blood. 2006;108(6):2020-8.
6. Shaughnessy JD, Jr., Zhan F, Burington BE, Huang Y, Colla S, Hanamura I, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. Blood. 2007;109(6):2276-84.
7. Shaughnessy JD, Jr., Qu P, Usmani S, Heuck CJ, Zhang Q, Zhou Y, et al. Pharmacogenomics of bortezomib test-dosing identifies hyperexpression of proteasome genes, especially PSMD4, as novel high-risk feature in myeloma treated with Total Therapy 3. Blood. 2011;118(13):3512-24.
8. Kuiper R, Broyl A, de Knegt Y, van Vliet MH, van Beers EH, van der Holt B, et al. A gene expression signature for high-risk multiple myeloma. Leukemia. 2012;26(11):2406-13.
9. Mulligan G, Mitsiades C, Bryant B, Zhan F, Chng WJ, Roels S, et al. Gene expression profiling and

correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood. 2007;109(8):3177-88.

10. Decaux O, Lode L, Magrangeas F, Charbonnel C, Gouraud W, Jezequel P, et al. Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myelome. Journal of Clinical Oncology. 2008;26(29):4798-805.

11. Chng WJ, Ahmann GJ, Henderson K, Santana-Davila R, Greipp PR, Gertz MA, et al. Clinical implication of centrosome amplification in plasma cell neoplasm. Blood. 2006;107(9):3669-75.

12. Dickens NJ, Walker BA, Leone PE, Johnson DC, Brito JL, Zeisig A, et al. Homozygous deletion mapping in myeloma samples identifies genes and an expression signature relevant to pathogenesis and outcome. Clin Cancer Res. 2010;16(6):1856-64.

13. Hose D, Reme T, Hielscher T, Moreaux J, Messner T, Seckinger A, et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. Haematologica. 2011;96(1):87-95.

14. Palumbo A, Avet-Loiseau H, Oliva S, Lokhorst HM, Goldschmidt H, Rosinol L, et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. Journal of Clinical Oncology. 2015;33(26):2863-9.

15. Amin SB, Yip WK, Minvielle S, Broyl A, Li Y, Hanlon B, et al. Gene expression profile alone is inadequate in predicting complete response in multiple myeloma. Leukemia. 2014;28(11):2229-34.

16. Chng WJ, Chung TH, Kumar S, Usmani S, Munshi N, Avet-Loiseau H, et al. Gene signature combinations improve prognostic stratification of multiple myeloma patients. Leukemia. 2016;30(5):1071-8.

17. Robiou du Pont S, Cleynen A, Fontan C, Attal M, Munshi N, Corre J, et al. Genomics of Multiple Myeloma. Journal of Clinical Oncology. 2017 Mar 20;35(9):963-7.

18. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. Nature Reviews Cancer. 2012;12(5):335-48.

19. Sonneveld P, Schmidt-Wolf IG, van der Holt B, El Jarari L, Bertsch U, Salwender H, et al. Bortezomib induction and maintenance treatment in patients with newly diagnosed multiple myeloma: results of the randomized phase III HOVON-65/ GMMG-HD4 trial. Journal of Clinical Oncology. 2012;30(24):2946-55.

20. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, et al. Initial genome sequencing and analysis of multiple myeloma. Nature. 2011;471(7339):467-72.

21. Zhou Y, Zhang Q, Stephens O, Heuck CJ, Tian E, Sawyer JR, et al. Prediction of cytogenetic abnormalities with gene expression profiles. Blood. 2012;119(21):e148-50.

22. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell. 2014;25(1):91-101.

23. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. Nat Rev Immunol. 2008;8(1):22-33.

24. De Silva NS, Klein U. Dynamics of B cells in germinal centres. Nat Rev Immunol. 2015;15(3):137-48.

25. Monti S, Tamayo P, Mesirov J, Golub T. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. Machine Learning. 2003;52:91-118.

26. Vikstrom I, Carotta S, Luthje K, Peperzak V, Jost PJ, Glaser S, et al. Mcl-1 is essential for germinal

center formation and B cell memory. Science. 2010;330(6007):1095-9.

27. Peperzak V, Vikstrom I, Walker J, Glaser SP, LePage M, Coquery CM, et al. Mcl-1 is essential for the survival of plasma cells. Nat Immunol. 2013;14(3):290-7.

28. Hanamura I, Stewart JP, Huang Y, Zhan F, Santra M, Sawyer JR, 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(5):1724-32.

29. Narita T, Ri M, Masaki A, Mori F, Ito A, Kusumoto S, et al. Lower expression of activating transcription factors 3 and 4 correlates with shorter progression-free survival in multiple myeloma patients receiving bortezomib plus dexamethasone therapy. Blood Cancer J. 2015;5:e373.

30. Zheng P, Guo H, Li G, Han S, Luo F, Liu Y. PSMB4 promotes multiple myeloma cell growth by activating NF-kappaB-miR-21 signaling. Biochem Biophys Res Commun. 2015;458(2):328-33.

31. Gupta VA, Matulis SM, Conage-Pough JE, Nooka AK, Kaufman JL, Lonial S, et al. Bone marrow microenvironment-derived signals induce Mcl-1 dependence in multiple myeloma. Blood. 2017;129(14):1969-79.

32. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic acids research. 2014;42(Database issue):D472-7.

33. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. Nat Immunol. 2007;8(5):463-70.

34. Milner JD, Holland SM. The cup runneth over: lessons from the ever-expanding pool of primary immunodeficiency diseases. Nat Rev Immunol. 2013;13(9):635-48.

35. Mackay F, Browning JL. BAFF: a fundamental survival factor for B cells. Nat Rev Immunol. 2002;2(7):465-75.

36. Reya T, O'Riordan M, Okamura R, Devaney E, Willert K, Nusse R, et al. Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. Immunity. 2000;13(1):15-24.

37. Wan Y, Kim TW, Yu M, Zhou H, Yamashita M, Kang Z, et al. The dual functions of IL-1 receptor-associated kinase 2 in TLR9-mediated IFN and proinflammatory cytokine production. J Immunol. 2011;186(5):3006-14.

38. Neben K, Lokhorst HM, Jauch A, Bertsch U, Hielscher T, van der Holt B, et al. Administration of bortezomib before and after autologous stem cell transplantation improves outcome in multiple myeloma patients with deletion 17p. Blood. 2012;119(4):940-8.

39. Wuilleme-Toumi S, Robillard N, Gomez P, Moreau P, Le Gouill S, Avet-Loiseau H, et al. Mcl-1 is overexpressed in multiple myeloma and associated with relapse and shorter survival. Leukemia. 2005;19(7):1248-52.

40. Mylin AK, Rasmussen T, Lodahl M, Dahl IM, Knudsen LM. Upregulated MCL1 mRNA expression in multiple myeloma lacks association with survival. Br J Haematol. 2009;144(6):961-3.

41. Takahashi S, Reddy SV, Chirgwin JM, Devlin R, Haipek C, Anderson J, et al. Cloning and identification of annexin II as an autocrine/paracrine factor that increases osteoclast formation and bone resorption. J Biol Chem. 1994;269(46):28696-701.

42. Obri A, Ouararhni K, Papin C, Diebold ML, Padmanabhan K, Marek M, et al. ANP32E is a histone chaperone that removes H2A.Z from chromatin. Nature. 2014;505(7485):648-53.

43. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature. 2012;482(7384):226-31.

44. Butti MG, Bongarzone I, Ferraresi G, Mondellini P, Borrello MG, Pierotti MA. A sequence analysis of the genomic regions involved in the rearrangements between TPM3 and NTRK1 genes producing TRK oncogenes in papillary thyroid carcinomas. Genomics. 1995;28(1):15-24.

45. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. PNAS. 2005;102(43):15545-50.

46. Kassambara A, Reme T, Jourdan M, Fest T, Hose D, Tarte K, et al. GenomicScape: an easy-to-use web tool for gene expression data analysis. Application to investigate the molecular events in the differentiation of B cells into plasma cells. PLoS Comput Biol. 2015;11(1):e1004077.

47. Jourdan M, Caraux A, Caron G, Robert N, Fiol G, Reme T, et al. Characterization of a transitional preplasmablast population in the process of human B cell to plasma cell differentiation. J Immunol. 2011;187(8):3931-41.

48. Jourdan M, Caraux A, De Vos J, Fiol G, Larroque M, Cognot C, et al. An in vitro model of differentiation of memory B cells into plasmablasts and plasma cells including detailed phenotypic and molecular characterization. Blood. 2009;114(25):5173-81.

49. Caron G, Le Gallou S, Lamy T, Tarte K, Fest T. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. J Immunol. 2009;182(12):7595-602.

50. Shi W, Liao Y, Willis SN, Taubenheim N, Inouye M, Tarlinton DM, et al. Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. Nat Immunol. 2015;16(6):663-73.

51. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun. 2014;5:2997.

52. Bochtler T, Hegenbart U, Cremer FW, Heiss C, Benner A, Hose D, et al. Evaluation of the cytogenetic aberration pattern in amyloid light chain amyloidosis as compared with monoclonal gammopathy of undetermined significance reveals common pathways of karyotypic instability. Blood. 2008;111(9):4700-5.

53. Zhu W, Kuziora M, Creasy T, Lai Z, Morehouse C, Guo X, et al. BubbleTree: an intuitive visualization to elucidate tumoral aneuploidy and clonality using next generation sequencing data. Nucleic acids research. 2016;44(4):e38.

54. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. Journal of Clinical Oncology. 2005;23(26):6333-8.

55. Leung-Hagesteijn C, Erdmann N, Cheung G, Keats JJ, Stewart AK, Reece DE, et al. Xbp1s-negative tumor B cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. Cancer Cell. 2013;24(3):289-304.

56. Hideshima T, Mitsiades C, Tonon G, Richardson PG, Anderson KC. Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. Nature Reviews Cancer. 2007;7(8):585-98.

57. Moreau P, Masszi T, Grzasko N, Bahlis NJ, Hansson M, Pour L, et al. Oral Ixazomib, Lenalidomide, and Dexamethasone for Multiple Myeloma. N Engl J Med. 2016;374(17):1621-34.

Figure legends

Figure 1. Prognostic impact of the MCL1-M signature-based MM clustering. Kaplan-Meier plots show significant differences in OS between patients with MCL1-M high and those with MCL1-M low MM in data sets GSE19784 (A and B), GSE2658 (C and D), and MMRF (E and F). The two graphs for each data set represent the first and last follow-up years with significant differences in OS. The median OS for patients of MCL1-M high MM in GSE19784 was 55 months; the median OS for patients in GSE2658, MMRF and the patients of MCL1-M low MM in GSE19784 was not reached at the time of analysis.

Figure 2. Distinct treatment responses in patients with MCL1-M high or MCL1-M low MM in the HOVON-65/GMMG-HD4 trial (data set GSE19784). Kaplan-Meier plots from (A) to (D) indicating the first and the last follow-up years with significant difference in PFS or OS in patients with MCL1-M high MM following their random assignment into the VAD (without bortezomib) or PAD (with bortezomib) treatment arm are shown. No such difference in PFS (E) or OS (F) was observed in patients with MCL1-M low MM.

Figure 3. Transcriptomic features in MCL1-M signature-defined MM subtypes in the context of B cell to plasma cell differentiation. GSEA plots (A) demonstrating enrichment of the pre-plasmablast gene set in the MCL1-M high MM in data sets GSE2658 and GSE19784 are shown. Large subsets of both MCL1-M (B) and MCL1-M low identifiers (C) were enriched in the GC B cells in a murine transcriptome data base for B cell to plasma cell differentiation, GSE60927.

Compared to plasma cell populations, MCL1-M low identifiers showed higher expression in the B cell populations. One-way anova test was used to compare the expression average of MCL1-M or MCL1-M low identifiers between the cell populations (***: $p < 0.0001$). FoB, MZB, B1, GCB, SplPB, SplPC, and BMPC stand for follicular B cells, marginal-zone B cells, B1 cells, GC B cells, spleen plasmablasts, spleen plasma cells, and bone marrow plasma cells, respectively.

Figure 4. Clonality analysis of characteristic SCNAs and driver mutations in MM samples with a clonal 1q gain. (A) R-HDS plot generated by BubbleTree of a representative MM sample (MMRC0571) with a clonal 1q gain. The R score represents the copy number ratio between the tumor and a matched normal sample; 1, <1, and >1 indicate normal copy number, loss, or gain of chromosome segments, respectively. HDS is a measure for the deviation of heterozygous state; 0, normal circumstance; >0, gain, loss, or LOH. The higher the HDS, the more cells carry the genomic changes. The branches of the tree represent pre-build copy number states. B, AB, BB, ABB, and BBB represent one copy loss, normal disomy, LOH, heterozygous gain, and homozygous gain, respectively. The numbers labelled on the branch represent the frequency of SCNA occurrence. Each circle on the plot corresponds to a chromosome segment. SCNAs with a similar frequency tend to belong to the same clone. (B) Clonal structure of this sample inferred by the BubbleTree analysis. In the sample, triploid heterozygous 1q gain belongs to the major clone harboring most SCNAs (gain of 1q, 5, 7, 9, 15, and 19p; and loss of 13). This major clone comprises ca. 90% of cells. The other two subclones, marked by the losses of 4q and 8p, or 8q

gain, make up 77% and 46% of cells, respectively. (C) Inferred temporal order of the major SCNAs and *NRAS* mutation in this sample.

John Wiley & Sons

Figure 1.

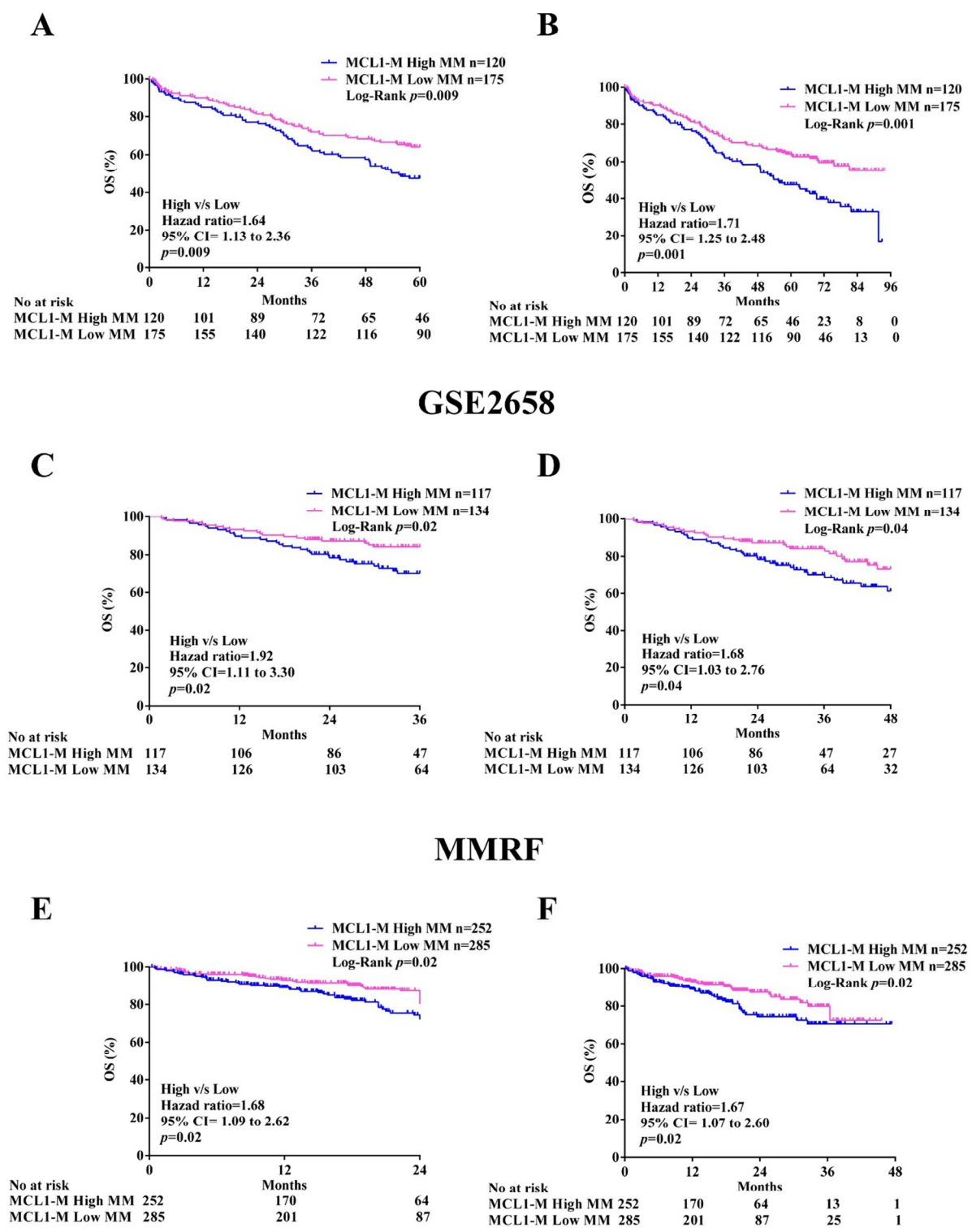

Figure 2.

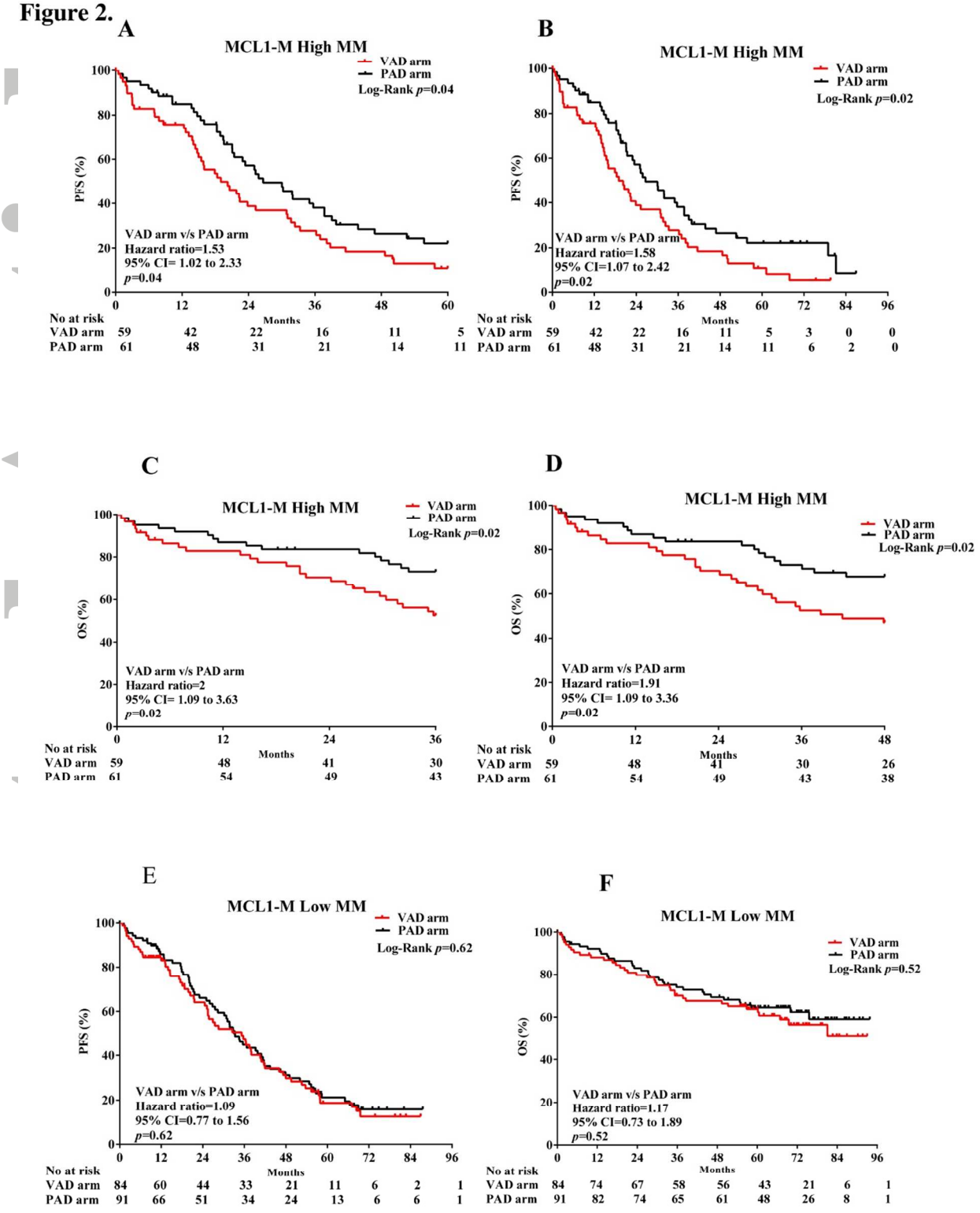

Figure 3.

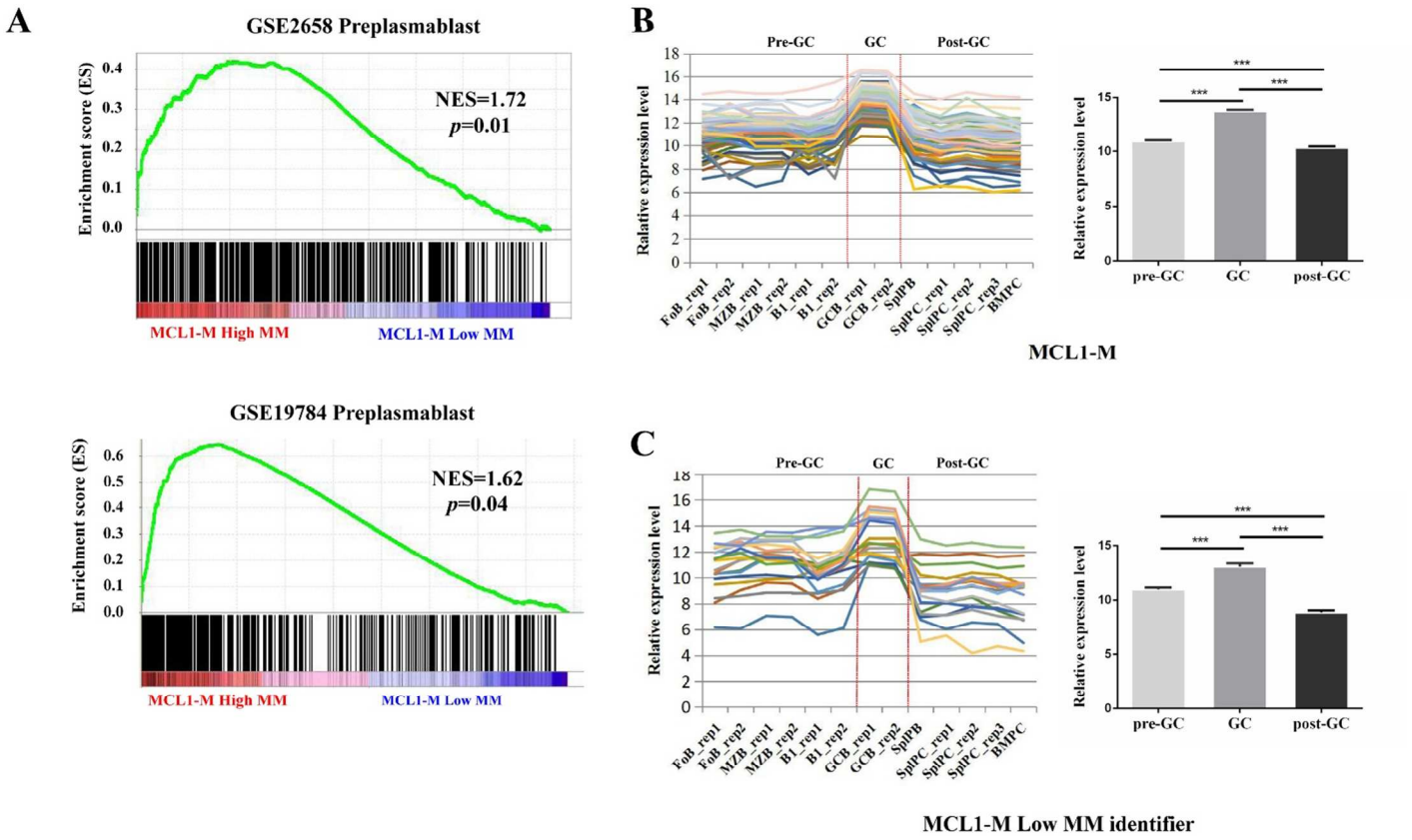

Figure 4.