FAM46C-mediated tumor heterogeneity predicts extramedullary metastasis and poorer survival in multiple myeloma

Weilong Zhang[1,*], Chaoling Wu[1,*], Shuang Geng[1,2,*], Jing Wang[1,*], Changjian Yan[1], Xiannian Zhang[1,2], Jia-jia Zhang[3], Fan Wu[1,2], Yuhong Pang[1,2], Yuping Zhong[3], Jianbin Wang[4], Wei Fu[1], Xin Huang[1], Wenming Wang[1], Xiaoqing Lyu[1,2], Yanyi Huang[1,2], Hongmei Jing[1]

[1]Department of Hematology, Biodynamic Optical Imaging Center (BIOPIC) and Lymphoma Research Center, Third Hospital, Peking University, Beijing 100084, China
[2]Beijing Advanced Innovation Center for Genomics (ICG), School of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100084, China
[3]Department of Hematology, Beijing Chaoyang Hospital West, Capital Medical University, Beijing 100054, China
[4]School of Life Sciences, Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100190, China
*Equal contribution

Correspondence to: Yanyi Huang, Hongmei Jing; email: yanyi@pku.edu.cn, hongmeijing@bjmu.edu.cn
Keywords: FAM46C, tumor heterogeneity, multiple myeloma, extramedullary metastasis, single cell sequencing
Received: March 22, 2019 Accepted: April 22, 2023 Published: May 6, 2023

ABSTRACT

Cancers originate from a single cell according to Nowell's theory of clonal evolution. The enrichment of the most aggressive clones has been developed and the heterogeneity arises for genomic instability and environmental selection. Multiple myeloma (MM) is a multiple relapse plasma cell cancer generated from bone marrow. Although there were accumulating researches in multiple myeloma pathogenesis, the heterogeneity remains poorly understood. The participants enrolled in this study were 4 EMP+ (EMP, Extramedullary plasmacytoma) and 2 EMP- primarily untreated MM patients. Single cell RNA sequencing and analysis were conducted for the single cell suspension, which was sorted by flow cytometry from peripheral blood mononuclear cells or bone marrow cells. In our research, the results of single cell RNA sequencing show that FAM46C determines MM tumor heterogeneity predicting extramedullary metastasis by influencing RNA stability. Further, we integrated and analyzed 2280 multiple myeloma samples from 7 independent datasets, which uncover that FAM46C mediated tumor heterogeneity predicts poorer survival in multiple myeloma.

INTRODUCTION

Multiple myeloma (MM) is a multiple relapse plasma cell malignancy in the bone marrow, which is characterized by clonal proliferation of plasma cells in the bone marrow and vast secretion of monoclonal immunoglobulin [1, 2]. In recent years, the incidence of MM is clearly increased, which accounts for 2% of all hematological malignancies and severely affects the life quality of older individuals [3]. The median survival time of MM patients ranges from several months to more than 10 years. Therefore, the exact predicting of prognosis is crucial to patients with MM. As a method predicting prognosis of patients with MM, the study of MM staging system has developed in recent years. The Durie-Salmon staging system, proposed in 1975, was first used to stage MM patients [4]. The international staging system (ISS) was proposed by the International Myeloma Working Group (IMWG) in 2006, which stages MM by serum

albumin and beta-2-microglobulin (β2-MG) levels [5]. Combined with molecular diagnostic techniques, a revised International Classification System (R-ISS) was proposed in 2015, which made a modification to the ISS system: Patients with high level of serum lactate dehydrogenase (LDH) and high risk of chromosomal abnormalities, including chromosomal 17p deletion and/or t (14; 16) and/or t (4; 14), are considered to be the ISS III [6, 7]. Nowadays, chemotherapy and autologous stem cell transplantation (ASCT) are main treatments for MM patients [8, 9]. With the development of genetic researches, more accurate medical treatments in MM prognosis prediction and therapy will be realized.

FAM46C, a member of nucleotidyltransferase super-family [10–16], encodes an active non-canonical poly (A) polymerase which enhances mRNA stability [17]. Moreover, FAM46C is a high-frequency mutation gene with many frameshift and nonsense mutations which suggests that it may act as a tumor suppressor gene [18, 19]. The deletion of FAM46C was reported in approximately 20% of MM patients, whose progression-free survival (FPS) and overall survival (OS) were lower [20, 21]. Overexpression of FAM46C also enhances replication of some viruses, which is identified as a type I interferon-stimulated gene [22, 23]. In addition, the mutation of FAM46C gene in mouse causes anemia [24] and loss-of-function FAM46C is associated with bone abnormalities in mouse [10, 22, 25, 26].

Cancers originate from a single cell according to Nowell's theory of clonal evolution [27]. The enrichment of the most aggressive clones has been developed and the heterogeneity arises for genomic instability and environmental selection [27]. Studies of next-generation sequencing in MM tumor cells have observed the presence of significant clonal heterogeneity [19, 28]. However, the study regarding the heterogeneity of MM remains limited. In our current study, we used single-cell sequencing to analyze the MM tumor cells and identified that FAM46C, by influencing RNA stability, meditated tumor heterogeneity to predict extramedullary metastasis and poorer survival.

RESULTS

Immunoglobulin genes specifically expressed in each multiple myeloma patient

In order to determine the gene expression levels of immunoglobulin in each single cell isolated from patients with MM, we performed gene expression analysis of each single cell. A total of 9 immuno-

globulin genes were detected at the single cell transcriptome level, of which 6 genes were specifically expressed in different MM patients (Figure 1 and Supplementary Figure 1). Immunoglobulin genes are specifically expressed in each multiple myeloma patients at the single cell transcriptome level through one-way analysis of variance. Immunoglobulin gene IGKV1-33 is specifically expressed in P23 multiple myeloma patients (Figure 1, P = 5.3E-30, one-way analysis of variance), while IGHV4-28 is specifically expressed in P21 multiple myeloma patients (Figure 1, P = 1.8E-12, one-way analysis of variance). IGKV3D-7 and IGHA1 were specially expressed in P20 and P19 MM patients, respectively (Figure 1). Moreover, IGHV1-24 and IGHV3-43 were specially expressed in P17 and P14 MM patients, respectively (Figure 1).

Intriguingly, 3 immunoglobulin genes also were expressed in various MM patients. IGKV3D-20 and IGKV4-1 were specifically expressed in both P20 and P21 multiple myeloma patients (Supplementary Figure 1, P = 1.0E-24 and 5.1E-15, one-way analysis of variance). IGKC was expressed in several MM patients, including P19, P20, and P23 (Supplementary Figure 1, P = 2.2E-14, one-way analysis of variance). The specifically expressed immunoglobulin genes in each multiple myeloma patient with an extremely significant P-value suggest that we have obtained the right sample and our method of single-cell RNA sequencing is robust.

The reduced expression of FAM46C in clonal plasma cells (CPCs) is related to RNA stability

To determine the differentially expressed genes in CPCs compared with bone marrow mononuclear cells (BMMCs), we performed differential expression analysis in CPCs and BMMCs isolated from same EMP-positive patient (P17 and P20). A total of 95 genes were differentially expressed (Supplementary Table 1), of these, FAM46C was significantly lower expressed in CPCs compared with BMMCs (Figure 2A, P17: P = 9.4E-05, FC = −1.85; P20: P = 9.4E-05, FC = −3.53; Unpaired t test, two sided).

FAM46C acts as an active non-canonical poly (A) polymerase which enhances mRNA stability. To explore whether the decreased FAM46C level influences mRNA stability in MM, we analyzed the expression level of the whole transcriptome in both CPCs and BMMCs. As compared to BMMCs, the significantly lower level of the whole transcriptome was identified in CPCs isolated from P17 and P20 patient (Figure 2B, P17: P = 0.0003; P20: P < 0.0001; Kolmogorov-Smirnov test), indicating that decreased level of FAM46C reduces mRNA stability in MM.

FAM46C mediated tumor heterogeneity in multiple myeloma

To further identify the heterogeneity of FAM46C in MM, we conducted an unsupervised clustering analysis for BMMCs transcriptome data of all patients. At the single-cell transcriptome level of P14 patient, two clusters (C1 and C2) of BMMCs were divided through the silhouette method according to the biggest average silhouette width (Figure 3A, 3B). The level of whole transcriptome was lower in C1 compared to C2 (Figure 3C). FAM46C was downregulated in C1 compared with C2 (Figure 3D), thereby C1 was identified as a low-FAM46C group while C2 was identified as a high-FAM46C group. Due to the decreased level of FAM46C influencing the expression of immunoglobulin genes, we analyzed the expression level of immunoglobulin genes in both low-FAM46C
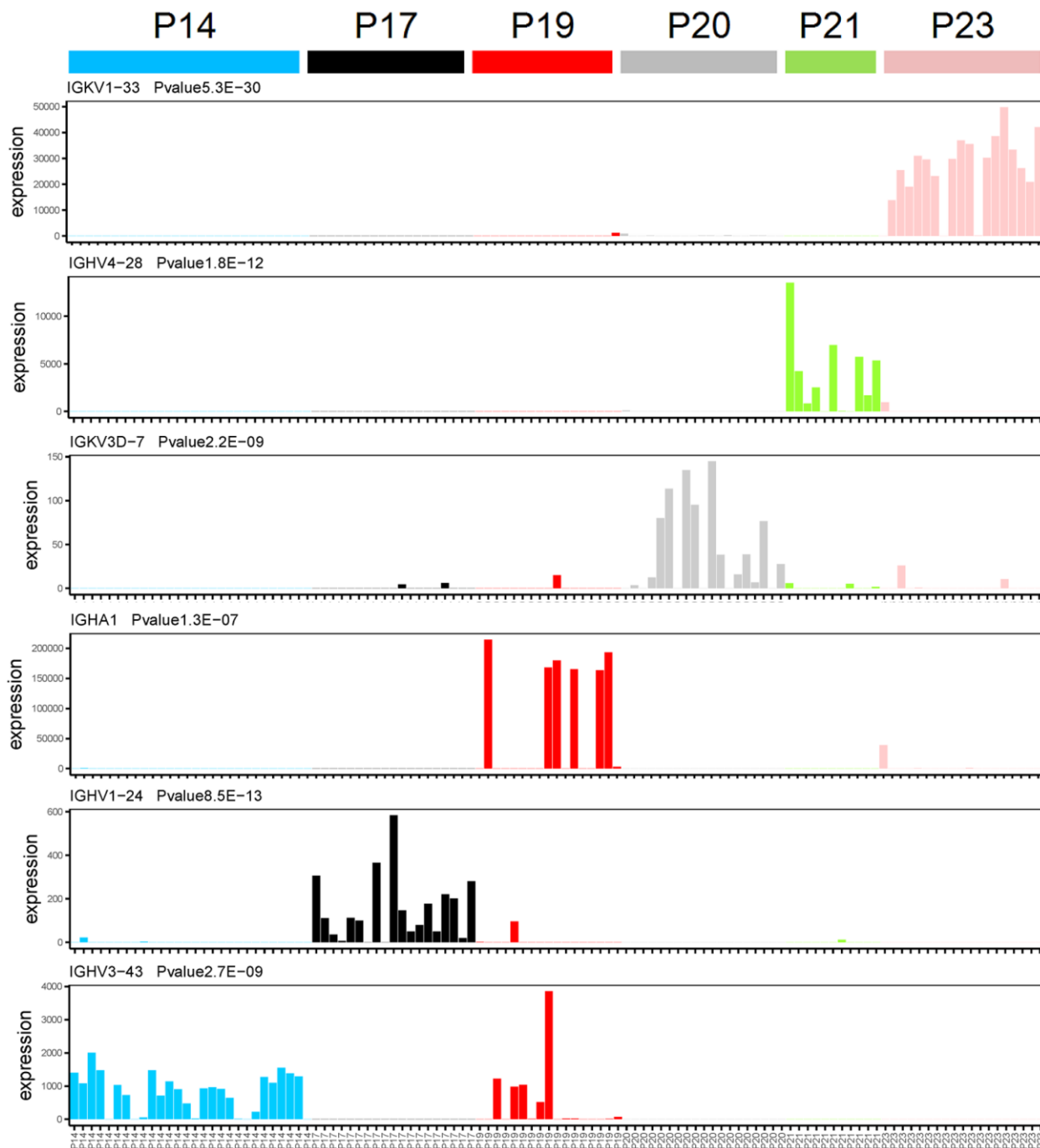

Figure 1. Immunoglobulin genes specifically expressed in each patient at the single cell transcriptome level. The X-axis and Y-axis represent the single cell and expression level (FPKM), respectively. One-way analysis of variance, the gene name and *P*-value were showed in top left corner of each figure.

and high-FAM46C group. Similar to FAM46C, immunoglobulin genes were also differentially expressed between low-FAM46C and high-FAM46C group. The majority of immunoglobulin genes, such as IGHG1, IGHG3, IGHG4, IGHV3-43, IGKC, IGKV1-39, and IGKV1D-39, were downregulated in low-FAM46C group (Figure 3D). In contrast, IGHV4-31 was downregulated in high-FAM46C group.

In addition, FAM46C mediated tumor heterogeneity was also observed in BMMCs isolated from P17, P19, P20, P21 and P23 multiple myeloma patients. There were 3 clusters of BMMCs were divided in both P17 (Supplementary Figure 2A, 2B) and P19 (Supplementary Figure 3A). In P19 BMMCs, the C1 and C3 were FAM46C low group and C2 was FAM46C high group (Supplementary Figure 3B, $P < 0.05$,
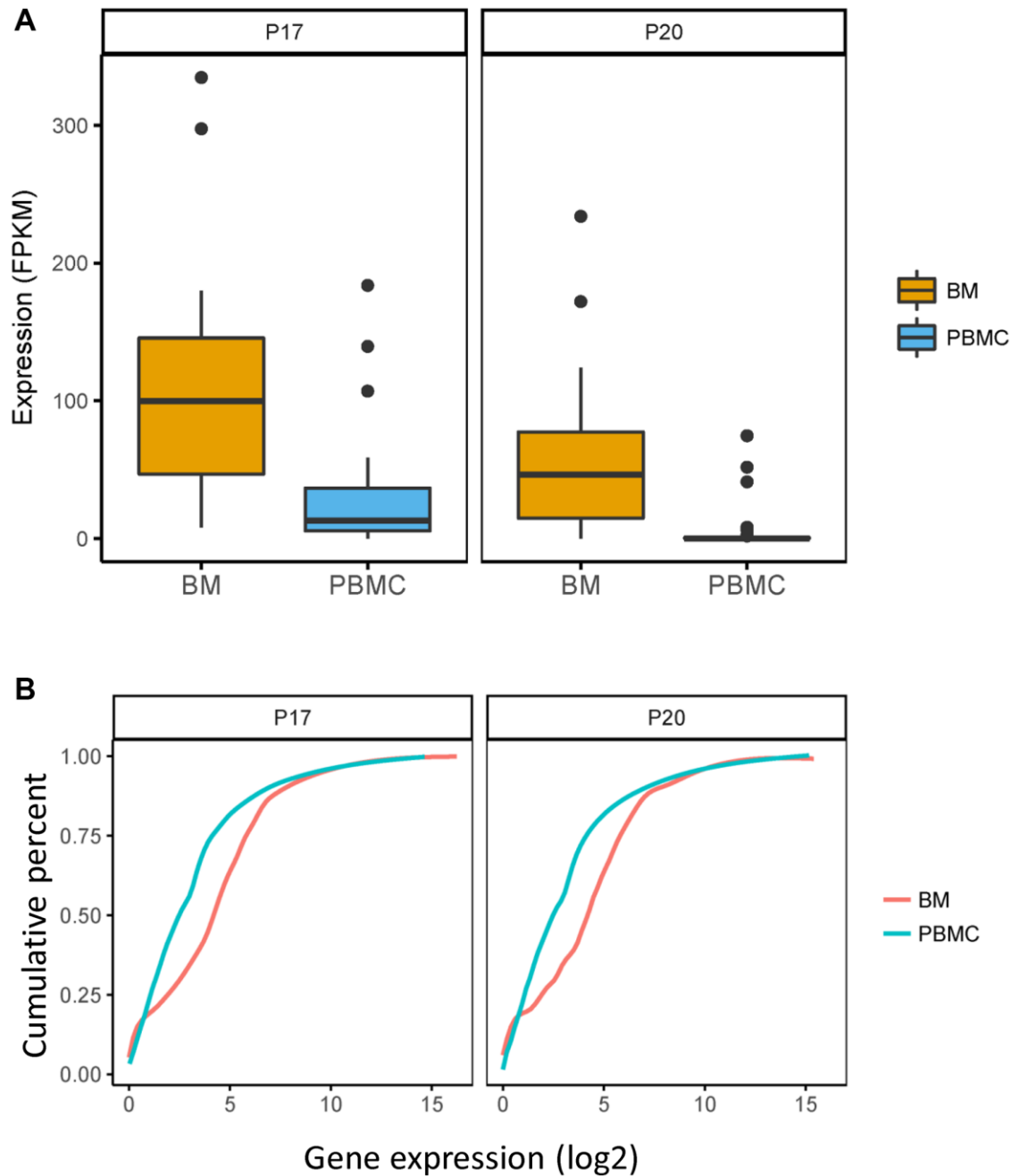

Figure 2. FAM46C expressed differently in CPCs versus BMMCs in P17 and P20 patient. Comparison of expression level of FAM46C and the whole transcriptome in CPCs and BMMCs of P17 and P20 patient. (**A**) FAM46C decreases in CPCs compared with BMMCs in P17 and P20 patients. P17: $P = 9.4\text{E-}05$, FC $= -1.85$; P20: $P = 9.4\text{E-}05$, FC $= -3.53$; Unpaired t test, two sided. (**B**) Lower transcriptome expression level in CPCs compared with BMMCs of single cell transcriptome in P17 and P20 patients. P17: $P = 0.0003$; P20: $P < 0.0001$; Kolmogorov-Smirnov test.

unpaired *t* test, two sided). Immunoglobulin gene IGLV3-1 decreased in the C1 compared with C2 in P19 BMMCs (Supplementary Figure 3B, *P* < 0.05, unpaired *t* test, two sided). In P17 BMMCs, C1 + C3 was low-FAM46C group and C2 was high-FAM46C group (Supplementary Figure 2B). In contrast to P14 BMMCs, IGKC, IGKV1-39, IGKV1D-39 showed the upregulated trend in low-FAM46C group compared with high-FAM46C group, but the difference was not significant (Supplementary Figure 2B). IGLV3-1 was
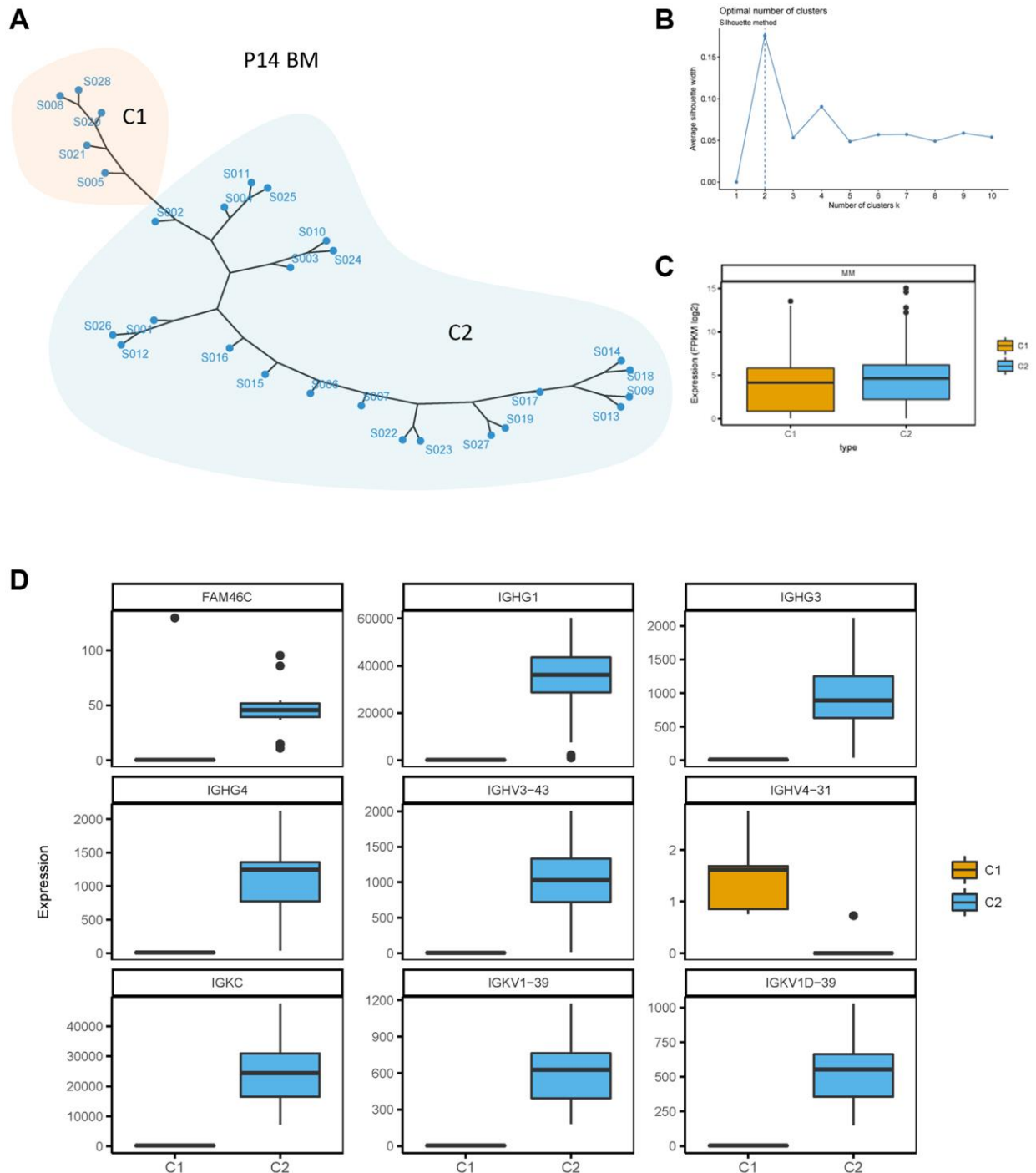

Figure 3. Unsupervised clustering of BMMCs of patient P14. (**A**) Unsupervised clustering of BMMCs of patient P14 at the single cell transcriptome level. One point means one single cell of P14 BMMCs. C1 and C2 means cluster 1 and cluster 2, showed by two different colors. (**B**) Silhouette method for cluster selecting. Average silhouette width was calculated for one to ten cluster. The cluster number was chosen with the biggest average silhouette width (dotted line). (**C**) Lower transcriptome expression level in C1 compared with C2 of single cell transcriptome in P14 patients. *P* = 0.0003, Mann Whitney test, two sided. (**D**) FAM46C and immunoglobulin genes were differently expressed between C1 and C2. *P* < 0.05, Mann Whitney test, two sided.

significant downregulated in C1 compared with C2 (Supplementary Figure 3B).

According to silhouette method, the BMMCs of P20, P21, and P23 were also divided into 2 clusters (Supplementary Figures 4A, 4B, 5A, 5B, 6A and 6B). In P20 BMMCs, FAM46C showed elevated expression in C1 compared with C2, but the difference was not significant (Supplementary Figure 4B). However, IGLV3-19 was significantly upregulated in C1 compared with C2 (Supplementary Figure 4B). Similar to P20 BMMCs, FAM46C also has not significant difference in C1 and C2 of P23 BMMCs (Supplementary Figure 6B) while IGKV1-33 and IGKV1D-33 were significantly downregulated in C1 compared with C2 (Supplementary Figure 6B). In P21 BMMCs, the C1 was low-FAM46C group while C2 was high-FAM46C group (Supplementary Figure 5A, $P < 0.05$, unpaired t test, two sided). A total of 11 immunoglobulin genes were downregulated in the low-FAM46C group compared with high-FAM46C group (Supplementary Figure 5B). Immunoglobulin genes such as IGHG1, IGHG3, IGHG4, IGKC, IGKV3-11, IGKV3-20, IGKV3D-11, IGKV3D-20, IGKV4-1, IGLL5 decreased in the C1 compared with C2 in P21 BMMCs (Supplementary Figure 5B, $P < 0.05$, unpaired t test, two sided). In P23 BMMCs, the C1 trended to be FAM46C high group and C2 trended to be FAM46C low group (Supplementary Figure 6, $P > 0.05$ but show the trend, unpaired t test, two sided). Immunoglobulin genes such as IGKV1-33, IGKV1D-33 decreased in the C1 compared with C2 in P23 BMMCs (Supplementary Figure 6, $P < 0.05$, unpaired t test, two sided). Collectively, these results further demonstrate that FAM46C mediated tumor heterogeneity in MM.

Single cell model of extramedullary metastasis in EMP[+] patient

Due to having identified that FAM46C-mediated tumor heterogeneity of MM and lower expression of FAM46C in CPCs compared with BMMCs, we constructed a single sell model to show the contribution of FAM46C-meditated tumor heterogeneity for extramedullary metastasis. In P17 patient, the low-FAM46C group (C1+C3) provided 83.3% of the contribution for extramedullary metastasis, whereas the high-FAM46C group (C2) provided only 16.7% (Figure 4A). Then, we further assessed the metastasis rate (means the contribution for extramedullary metastasis) of each BMMC, the metastasis rate of the low-FAM46C BMMCs was significantly higher than high-FAM46C BMMCs (Figure 4B). Similarly, the low-FAM46C group (C2), in P20 MM patient, provided more contribution than the high-FAM46C group (C1) for extramedullary metastasis (Figure 4C), and the

metastasis rate of the low-FAM46C BMMCs was also significantly higher than high-FAM46C BMMCs (Figure 4D).

FAM46C expression is an independent prognostic factor

FAM46C mediated tumor heterogeneity contributes to extramedullary metastasis of multiple myeloma. To further explore whether FAM46C could predict poor survival of MM patients at the population level, we integrated 7 independent datasets, a total of 2280 MM samples (2072 patients), and subsequently analyzed the relationship between FAM46C gene expression and clinical characteristics of MM.

We analyzed the FAM46C expression in different ISS stages and molecular subtypes. FAM46C was lower expressed in ISS stage II compared with ISS stage I, and was no difference between ISS stage III and ISS stage II/I (Supplementary Figure 7A, $P = 0.0078$; Kruskal-Wallis test). After combining with molecular subtype, similar results were also identified in immunoglobulin G subtype of MM (Supplementary Figure 7B, $P = 0.032$; Kruskal-Wallis test). In another independent dataset with 293 multiple myeloma samples, FAM46C was lower expressed in ISS stage III compared with ISS stage I and II (Supplementary Figure 8, $P = 0.011$, $P = 0.028$; Kruskal-Wallis test). FAM46C was lower expressed in FISH 1q21 amplification samples (≥ 4 copies) compared with 2 and 3 copies (Supplementary Figure 9A, $P = 0.00094$, $P = 0.009$; Kruskal-Wallis test). FAM46C was lower expressed in proliferation and nuclear factor kB (NFKB) molecular subtypes (Supplementary Figures 9B and 10, $P < 0.01$, $P < 0.001$; ANOVA test).

After the further integration of 7 independent datasets, we got a total of 559 MM samples with survival time. We analyzed the relationship between FAM46C gene expression and survival time. Multivariate Cox regression analysis showed that FAM46C was an independent prognostic factor of OS and EFS. FAM46C is a prognostic factor independent of β2-MG, HGB, ALB, MRI and BMPC (Figure 5A). For ISS stage I MM patients, FAM46C can also distinguish some patients based on poor survival (Supplementary Figure 11A, EFS, $P = 0.0029$; OS $P = 0.016$; Log-rank test). For stage II and III MM patients with lower FAM46C expression, an extremely poorer survival can be observed in these patients (Supplementary Figure 11B, EFS, $P < 0.0001$; OS $P = 0.00025$; Log-rank test). The 559 multiple myeloma patients can be divided into FAM46C-low group (57 patients) and FAM46C-high group (502 patients) (Supplementary Figures 11, 12).
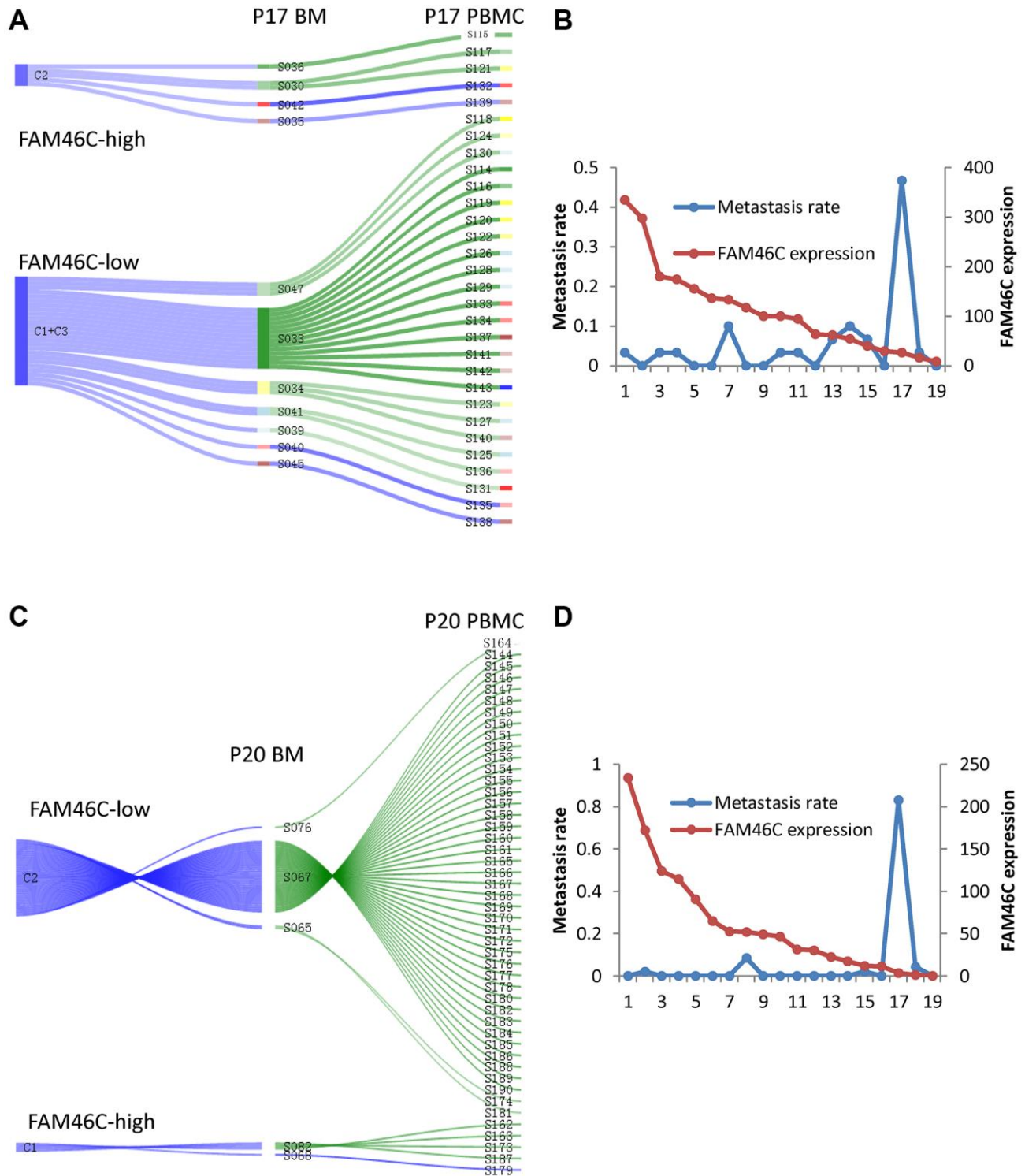
Figure 4. The model of extramedullary metastasis in EMP⁺ patient P17 and P20 at single cell transcriptome level. (**A**) Sankey diagram shows the model of extramedullary metastasis in EMP⁺ patient P17 at single cell transcriptome level. 3 clones were shown the right is all the CPCs in EMP⁺ patient P17. The middle is the related BMMCs in EMP⁺ patient P17. The left is the cluster of the BMMCs in EMP⁺ patient P17. The line between CPCs and BMMCs means that BMMC with the best Pearson correlation was chosen for a certain CPC. (**B**) The relationship between FAM46C expression level and metastasis rate in EMP⁺ patient P17. (**C**) Sankey diagram shows the model of extramedullary metastasis in EMP⁺ patient P20 at single cell transcriptome level. 2 clones were shown. The right is all the CPCs in EMP⁺ patient P20. The middle is the related BMMCs in EMP⁺ patient P20. The left is the cluster of the BMMCs in EMP⁺ patient P20. The line between CPCs and BMMCs means that BMMC with the best Pearson correlation was chosen for a certain CPC. (**D**) The relationship between FAM46C expression level and metastasis rate in EMP⁺ patient P20.

We compared the whole transcriptome expression level between the low-FAM46C and the high-FAM46C group. Lower transcriptome expression level was identified in the low-FAM46C group compared with the high-FAM46C group in 559 MM patients (Figure 5B).

Moreover, we further investigated whether clinical outcomes were different between the low-FAM46C and the high-FAM46C group. For 295 ISS stage I MM patients, Kaplan-Meier survival curves indicated that patients with lower FAM46C expression had poorer

A

| | HR | 95% CI for HR | | P-value |
		Lower	Upper	
EFS				
B2M (>= 3.5 mg/l)	1.46	1.08	1.99	1.54E-02
ALB (>= 35 g/l)	0.87	0.61	1.24	4.46E-01
HGB (>= 100 g/l)	0.78	0.58	1.06	1.13E-01
MRI (>= 3 focal lesions)	1.37	1.05	1.81	2.19E-02
BMPC (>= 35%)	1.45	1.06	1.99	2.18E-02
FAM46C (>=13.6)	0.54	0.36	0.80	2.11E-03
OS				
B2M (>= 3.5 mg/l)	1.68	1.15	2.45	7.40E-03
ALB (>= 35 g/l)	0.73	0.49	1.09	1.21E-01
HGB (>= 100 g/l)	0.86	0.60	1.23	4.13E-01
MRI (>= 3 focal lesions)	1.83	1.30	2.58	5.03E-04
BMPC (>= 35%)	1.38	0.93	2.05	1.06E-01
FAM46C (>=13.6)	0.59	0.38	0.93	2.17E-02

B
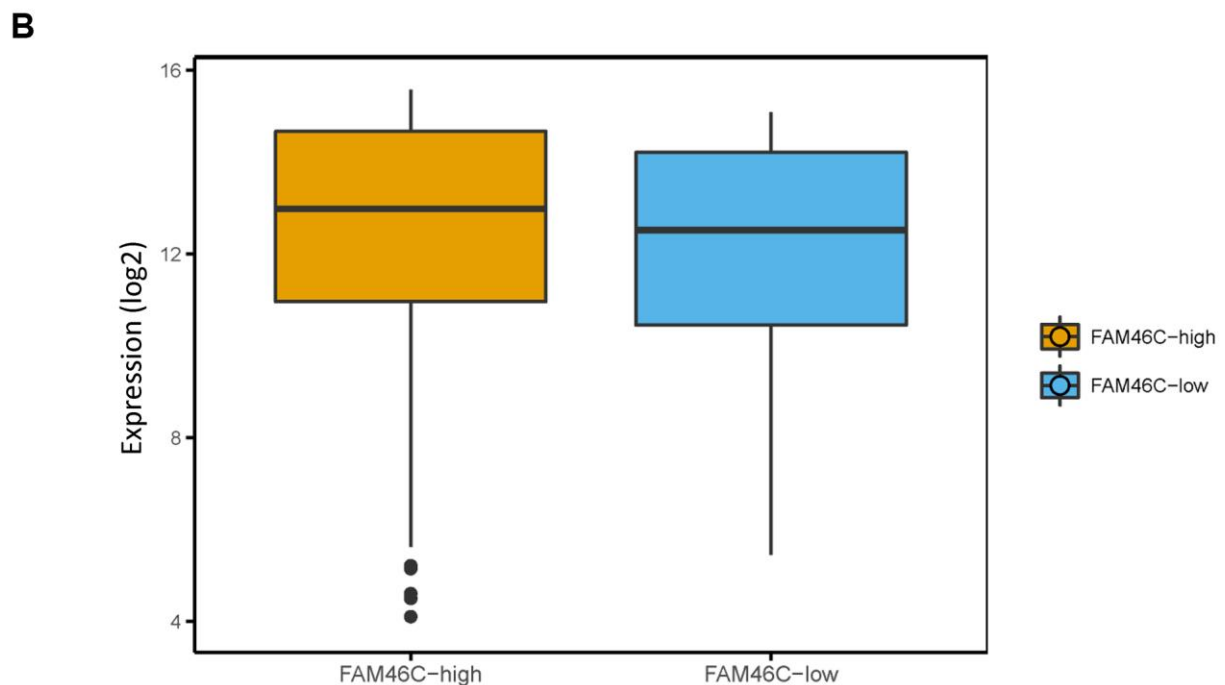
Figure 5. The expression of FAM46C is an independent prognostic factor in multiple myeloma. (**A**) Cox regression analysis identified FAM46C expression (cut off FAM46C ≥ 13.6) as an independent prognostic factor in 559 multiple myeloma patients. OS, overall survival; EPS, progress free survival. Abbreviations: β2-MG: β2 microglobulin; ALB: serum albumin; HR: hazard ratio; CI: confidence interval. (**B**) Lower transcriptome expression level in FAM46C-low group compared with FAM46C-high group in 559 multiple myeloma patients. $P = 0.0043$, unpaired t test, two sided (cut off FAM46C ≥ 13.6).

clinical outcomes (shorter OS and EFS time) (Supplementary Figure 12A). Similar results were also observed in 263 ISS stage II and III MM patients (Supplementary Figure 12B).

FAM46C were related to LDH level and clinical therapy in multiple myeloma patients

Further analysis of clinical characteristics between low-FAM46C and high-FAM46C group revealed that they were similar in age, sex, race, etc. However, several blood indexes (Figure 6A, $P = 0.016$, $P < 0.001$, $P = 0.045$, unpaired t test, two sided), especially LDH, were lower in high-FAM46C group. Among them, LDH is the most significantly different clinical characteristic (Figure 6A, $P < 0.001$, unpaired t test, two sided). LDH is active in the anaerobic glycolysis, which seems closely related to FAM46C. So, we compared the expression of hypoxia-related genes, including HIF1A, HIF1AN, and HIF3A, between low-FAM46C and high-FAM46C group, but none of them were significantly expressed (Supplementary Figure 13,

A

|		FAM46C-low	FAM46C-high	p.value
n		57	502	
AGE (mean (sd))		58.15 (8.62)	57.07 (9.55)	0.413
SEX (%)	female	18 (31.6)	204 (40.6)	0.237
	male	39 (68.4)	298 (59.4)	
RACE (%)	other	3 (5.3)	59 (11.8)	0.209
	white	54 (94.7)	443 (88.2)	
ISOTYPE (%)	FLC	11 (22.0)	73 (14.9)	0.324
	IgA	10 (20.0)	123 (25.2)	
	IgD	1 (2.0)	2 (0.4)	
	IgG	28 (56.0)	285 (58.3)	
	Nonsecretory	0 (0.0)	6 (1.2)	
B2M (mean (sd))		6.36 (9.08)	4.55 (4.75)	0.016
CRP (mean (sd))		12.90 (19.50)	11.49 (23.37)	0.663
CREAT (mean (sd))		1.58 (1.92)	1.29 (1.18)	0.105
LDH (mean (sd))		206.58 (113.67)	168.05 (56.97)	<0.001
ALB (mean (sd))		4.01 (0.62)	4.05 (0.58)	0.601
HGB (mean (sd))		11.05 (1.70)	11.28 (1.82)	0.383
ASPC (mean (sd))		43.43 (26.06)	42.59 (24.16)	0.812
BMPC (mean (sd))		51.59 (28.05)	45.79 (26.03)	0.118
MRI (mean (sd))		14.96 (17.01)	10.62 (14.20)	0.045

B
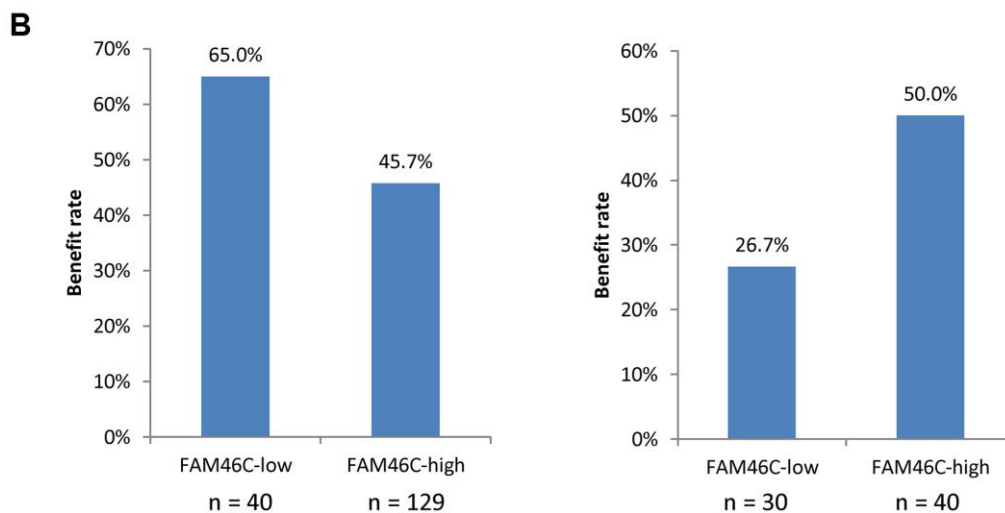
Figure 6. FAM46C was related to LDH level and clinical therapy in in multiple myeloma patients. (**A**) The comparison of clinical information between FAM46C-low and FAM46C-high group. (**B**) FAM46C was related to bortezomib and dexamethasone therapy in multiple myeloma patients. Left, bortezomib, $P = 0.03$, Chi-square test, two sided. Right, dexamethasone, $P = 0.048$, Chi-square test, two sided.

P > 0.05, unpaired *t* test, two sided). Interestingly, 2 other hypoxia related genes (CITED2 and TAZ) were significantly expressed in FAM46C-low vs. FAM46C-high group (Supplementary Figure 13, *P* < 0.0001, unpaired *t* test, two sided). CITED2 was the top 1 down regulated gene in FAM46C-low vs. FAM46C-high group (Supplementary Figure 13, *P* < 0.0001, unpaired *t* test, two sided).

Moreover, we analyzed gene expression profiles and the clinical therapy response in another 239 multiple myeloma patients. FAM46C was related to bortezomib and dexamethasone therapy in multiple myeloma patients (Supplementary Figure 14). FAM46C was more highly expressed in MR patients with dexamethasone therapy (*P* < 0·01; *P* < 0·05; ANOVA test; Supplementary Figure 14) and in PR patients with VAD (Vincristine, Adriamycin, and Dexamethasone) therapy (*P* < 0·05; ANOVA test; Supplementary Figure 15). The low-FAM46C group harbored a much higher benefiting rate to the bortezomib treatment than the high-FAM46C group (Figure 6B). Conversely, low-FAM46C group had only a 26.7% benefiting rate to the dexamethasone treatment, which was much lower compared to the high-FAM46C group (50%) (Figure 6B).

Furthermore, we assessed the gene expression of FAM46C before and after treatment. As compared to baseline, the gene expression level of FAM46C was significantly lower in relapsed MM with TT2 treatment (autologous hematopoietic stem-cell transplantation and thalidomide therapy) or TT3 treatment (incorporating bortezomib up-front into a tandem transplant regimen) (Supplementary Figure 16). This suggested that FAM46C-mediated tumor sub clone (lower FAM46C sub clone) faced a positive selection during the tumor clonal evolution. Consistently, this clonal evolution was also observed in two other independent datasets. FAM46C is lowly expressed in relapsed MM with at least one autologous hematopoietic stem-cell transplantation (ASCT) compared with presentation MM (Supplementary Figure 17A). Moreover, a similar result was observed in relapsed MM after chemotherapy and pre-1st bone marrow transplant compared with baseline MM (Supplementary Figure 17B).

The cell proliferation capacity of H929 cell was potentiated with the knockdown of FAM46C

The best FAM46C-siRNA was selected based on the results of western blot (Figure 7A). To identify whether FAM46C could affect the proliferation capacity,
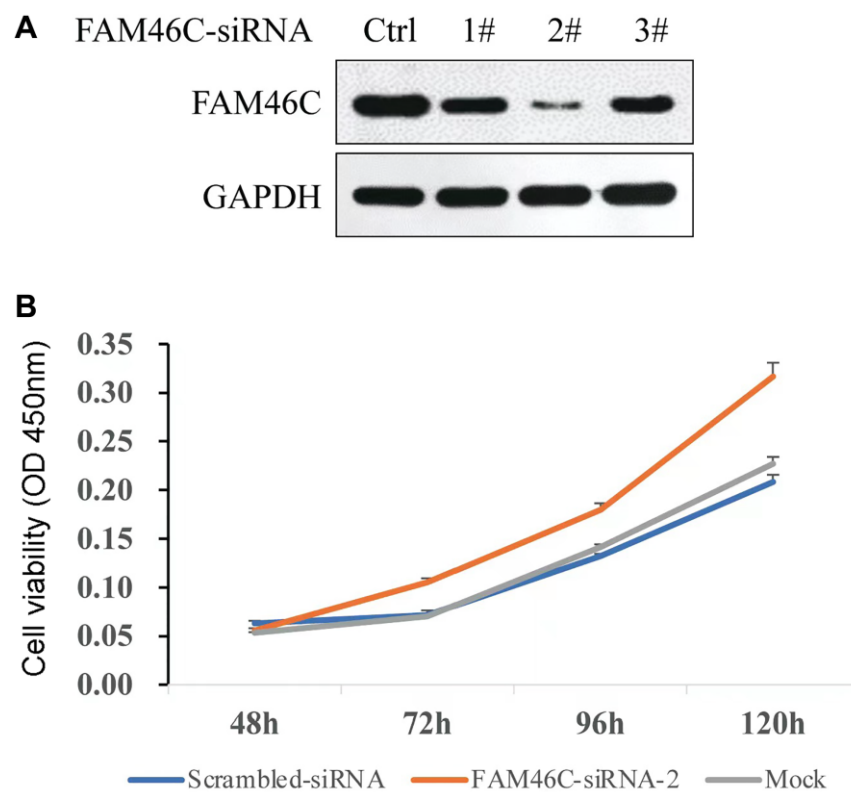

Figure 7. The capacity of cell proliferation was potentiated by the reduction of FAM46C expression in H929 cell. (**A**) The selection of FAM46C-siRNA by western-blot, Ctrl: control; 1#, 2#, and 3# represent the number of each FAM46C-siRNA. (**B**) The detection of cell proliferation capacity in H929 cells.

FAM46C-siRNA-2 was transfected into H929 cells and Cell Counting Kit-8 (CCK8) assay was used to detect cell proliferation capacity. We found that the reduction of FAM46C expression could enhance cell proliferation capacity significantly, particularly at 96h (Figure 7B, P = 3.2E-4, ANOVA test) and 120 h (Figure 7B, P = 3.0E-4, ANOVA test).

DISCUSSION

Cancers originate from a single cell according to Nowell's theory of clonal evolution [27]. The enrichment of the most aggressive clones has been developed and the heterogeneity arises for genomic instability and environmental selection [27]. In our research, the results of single cell RNA sequencing show that the FAM46C-meditated tumor heterogeneity predicts extramedullary metastasis and poorer survival in multiple myeloma by influencing RNA stability.

By analyzing the single cell RNA sequencing data of BMMCs from 6 multiple myeloma patients, we found the BMMCs from 3 patients could be divided into a significantly FAM46C-low and a FAM46C-high group. The other 3 patients trended to be one of the FAM46C-low and FAM46C-high group. There are 4 points that support the heterogeneity in BMMCs: (a) Unsupervised clustering of BMMCs of patients were conducted, which supports that FAM46C is a statistically dominate factor in distinguishing between clusters of BMMCs and the heterogeneity in BMMCs. (b) The immunoglobulin genes which are known as the hallmarks of clonal plasma cells were differently expressed among clusters of BMMCs. (c) Low expression of transcriptome level was observed in the FAM46C-low cluster which supports a functional cluster mediated by FAM46C. (d) The single cell metastasis model showed the high metastasis contribution of FAM46C low expression group in BMMCs. Thus, the tumor heterogeneity of BMMCs mediated by FAM46C exits in the multiple myeloma.

FAM46C-meditated tumor heterogeneity in BMMCs is functional. The following 3 points support the functional FAM46C-meditated tumor heterogeneity: (a) FAM46C low expression cluster contribute much more than FAM46C high expression cluster. (b) Low expression of FAM46C is closely related to the poorer survival and therapeutic response in multiple myeloma. (c) FAM46C is one of the highest frequently mutated genes with many frameshift and nonsense mutations [18, 29]. Thus, FAM46C-meditated tumor heterogeneity in BMMCs is functional.

What is the role of FAM46C in the FAM46C-meditated heterogeneity cluster? In our research, we can uncover the role of FAM46C layer-by-layer. (a) FAM46C-meditated RNA stability is a distinct role of FAM46C in the heterogeneity cluster. Low expression of transcriptome was observed in the BMMCs vs. CPCs comparison (single cell), FAM46C-low cluster vs. FAM46C-high cluster comparison (single cell), FAM46C low group vs. FAM46C high group comparison (population comparison). (b) FAM46C is closely related to the LDH level of multiple myeloma patients. (c) CITED2 is the top 1 down regulated gene in FAM46C-low vs. FAM46C-high group (population comparison). CITED2 is an efficient brake hypoxic response [30]. Low expression of CITED2 will cause the cell into hypoxia survival [30]. Thus, FAM46C-meditated RNA stability and CITED2 hypoxic response may play a role in the FAM46C-meditated tumor heterogeneity.

METHODS

Patients and sample preparation

The participants enrolled in this study were 4 EMP+ (EMP, Extramedullary plasmacytoma) and 2 EMP- primarily untreated MM patients at the Peking University Third Hospital and Beijing Chao-Yang Hospital during 2014-10 to 2015-06 (Supplementary Table 2). All the patients were diagnosed according to the 2009 IMWG (International Myeloma Working Group) consensus criteria [31]. This research was approved by the Peking University Third Hospital ethics committee. Informed consent was obtained in this study in accordance with the Declaration of Helsinki.

4 ml peripheral blood and bone marrow were drawn from untreated MM patients. Peripheral blood mononuclear cells (PBMCs) or bone marrow cells (BMCs) for Single cell RNA sequencing were isolated and purified through Ficoll-Paque Plus (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) and human CD138 MACS Microbeads (Miltenyi Biotec, Germany) according to the manufacturer's instructions. Single cell suspension was then analyzed or sorted by flow cytometry with antibodies (CD38, CD138) (Biolegend, San Diego, CA, USA) through UltraComp beads (eBioscience/ThermoFisher, USA), Aria II SORP flow cytometer (BD Biosciences, San Jose, CA, USA).

Single cell RNA sequencing and raw data analysis

Single cell libraries were constructed according to Smartseq2 protocol [32] and sequenced on HiSeq 2500 platform (Illumina, San Diego, CA, USA). Then RNA-seq reads were mapped to the human reference (hg19) using TopHat software (version 2.0.9) allowed for at most 2 mismatches and at least 20nt read length [33,

34]. "*accepted_hits.bam" files of TopHat results were kept for the subsequent analysis. RNA expression was calculated by Cufflinks software (version 2.1.1) using the input "*accepted_hits.bam" files of each sample and quantified as Fragments Per Kilobase of transcript per Million mapped reads (FPKM). The annotation file of each transcript was using "Homo_sapiens.GRCh37.70.gtf" file (Ensembl). FPKM value larger than 1 at least in one cell were taken as stably expressed transcripts [35]. The cells with > 1E5 mapping read and > 1E3 genes numbers were selected. 190 single cells were used for final data analysis after filtering (Supplementary Table 2).

Single cell RNA sequencing analysis

Immunoglobulin genes of 6 patients were used for One-way analysis of variance. The *P*-value < 1E-6 was defined as specifically expressed immunoglobulin genes in each patient at the single cell transcriptome level.

The transcriptome of clonal circulating plasma cells (CPCs) versus myeloma cells in bone marrow (BMMCs) in P17 and P20 patient were compared using unpaired *t* test. The different expressed genes (DEGs) were defined using the criterion: (a) *P* < 0.05 in unpaired *t* test; (b) foldchange (FC, log2) > 1 or < −1; (c) The criterion of (a and b) should be meet in both P17 and P20 patient. These DEGs may be related to extramedullary metastasis in multiple myeloma.

Unsupervised clustering of BMMCs single cells was done using the DEGs and immunoglobulin genes related genes. Silhouette method was used for cluster selecting. Average silhouette width was calculated for one to ten cluster. The cluster number with the biggest average silhouette width was chosen.

The different expressed genes were used for the whole transcriptome level comparison. *P* < 0.05 in unpaired *t* test and foldchange (FC, log2) > 1 or < −1 was used to defined the different expressed genes if no special designation. For C1 cluster and C2 cluster of P14 BMMCs comparison, *P* < 0.0001 in unpaired *t* test and foldchange (FC, log2) > 1 or < −1 was used to defined the different expressed genes.

The model of extramedullary metastasis in EMP+ patient P17 and P20 at single cell transcriptome level was conducted and showed using Sankey diagram. The Pearson correlations were calculated between each CPCs and BMMCs. Only the BMMC with the biggest Pearson correlation value was chosen for a certain CPC. And this BMMC was defined as a possible contributor for a certain CPC. So, a BMMC can contribute 0, 1 or more CPCs. The metastasis rate means a BMMC contributes to how many CPCs. The metastasis rate (M_j) was calculated using the method in (1).

$$M_j = C_j / C_p \qquad (1)$$

Where M_j indicates the metastasis rate for the jth BMMC.
Where C_j indicates the number of CPC contributed with the jth BMMC.
Where C_p indicates the number of CPC analyzed for the certain patient. 30 CPCs for P17 patient and 47 CPCs for P20 patient.

Gene expression profiling analysis in 2280 multiple myeloma samples

To further analysis the gene expression profiling of multiple myeloma and the relationship between FAM46C expression and clinical characteristics of multiple myeloma, we integrated 7 independent datasets including totally 2280 samples (2072 patients) from the NCBI GEO database. Probeset measures of all the arrays were calculated by RMA (robust multiarray averaging) method. Relative RNA expression value was log-transformed using log2. Data were analyzed with the unpaired *t* test and a *P*-value of < 0.05 was considered to be statistically significantly. Only genes with foldchange (log2) > 1 or < −1 were defined as different expression genes.

GSE24080 with 559 samples (559 patients) were retrieved from the NCBI GEO database. The gene expression array was Affymetrix Human Genome U133 Plus 2.0 Array [36]. The relationship between FAM46C expression and survival (EFS and OS), ISS stages, immunophenotype, FISH 1q21 amplification, molecular subtypes were analyzed.

GSE9782 with 238 samples (238 patients) were retrieved from the NCBI GEO database. The 238 samples were tested by Affymetrix Human Genome U133A Array and Affymetrix Human Genome U133B Array respectively (totally 476 arrays) [37]. FAM46C expression in different therapeutic response with bortezomib or dexamethasone (Dex) was analyzed.

GSE19784 with 308 samples (308 patients) were retrieved from the NCBI GEO database. The 308 samples were tested by Affymetrix Human Genome U133 Plus 2.0 Array [38]. The relationship between FAM46C expression and molecular subtypes were analyzed.

GSE39754 with 136 samples (136 patients) were retrieved from the NCBI GEO database [39]. The 136 samples were tested by Affymetrix Human Exon 1.0 ST Array. FAM46C expression in different therapeutic response with Vincristine, Adriamycin, and

Dexamethasone (VAD) and Autologous Stem Cell Transplant (ASCT) were analyzed.

GSE82307 with 66 samples (33 patients) were retrieved from the NCBI GEO database [40]. The 66 samples were tested by Affymetrix Human Genome U133 Plus 2.0 Array. FAM46C expression in presentation (baseline) and relapse paired samples in GSE82307 dataset (totally 33 paired samples) were analyzed. All samples except two were proceeded with at least one autologous hematopoietic stem-cell transplantation (ASCT).

GSE19554 with 36 samples (18 patients) were retrieved from the NCBI GEO database [41]. The 33 samples were tested by Affymetrix Human Genome U133 Plus 2.0 Array. FAM46C expression in baseline and pre-1st (after chemotherapy; pre-1st bone marrow transplant) paired samples in GSE19554 dataset (totally 18 paired samples) were analyzed.

GSE31161 with 937 samples (780 patients) were retrieved from the NCBI GEO database. The 937 samples were tested by Affymetrix Human Genome U133 Plus 2.0 Array [42, 43]. FAM46C expression in baseline and relapse samples with TT2 (autologous hematopoietic stem-cell transplantation and thalidomide therapy) and TT3 (incorporating bortezomib up-front into a tandem transplant regimen) therapy (totally 937 samples) were analyzed.

Western blot analysis

The best FAM46C-siRNA was selected based on the results of western blot. siRNA was purchased from RiboBio (China). Antibody against FAM46C was purchased from Abcam (UK) and were used at a dilution of 1:1000.

Detection of H929 cell proliferation

The proliferation ability of H929 cells was examined by CCK8 assay. H929 cells of different groups were seeded into 96-well plates, and the optical density (OD) value at 450 nm was determined at gradient times (48 h, 72 h, 96 h and 120 h) via using an automatic microplate reader.

Statistical analysis

R software v3.1.3 (ggplot2, ggpubr package) was used in this study. A statistical significance level of $P < 0.05$ was used.

Abbreviations

ASCT: autologous stem cell transplantation; BM: bone marrow; BMCs: bone marrow cells; BMMC: myeloma cells in bone marrow; CPCs: clonal circulating plasma cells; DEG: different expressed genes; Dex: dexamethasone; EMP: Extramedullary plasmacytoma; FC: foldchange; Ig: immunoglobulin; IMWG: the International Myeloma Working Group; ISS: the international staging system; LDH: lactate dehydrogenase; MM: Multiple myeloma; OS: overall survival; PBMCs: peripheral blood mononuclear cells; R-ISS: revised International Classification System; TT2: autologous hematopoietic stem-cell transplantation and thalidomide therapy; TT3: incorporating bortezomib up-front into a tandem transplant regimen; VAD: Vincristine, Adriamycin, and Dexamethasone.

AUTHOR CONTRIBUTIONS

Yanyi Huang and Hongmei Jing conceived the project. Weilong Zhang, Chaoling Wu, Shuang Geng, and Jing Wang analyzed the data. Weilong Zhang, Chaoling Wu, Shuang Geng, Jing Wang, Yanyi Huang, and Hongmei Jing contributed towards the interpretation of the data. All authors wrote and approved the final manuscript.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

ETHICAL STATEMENT AND CONSENT

This research was approved by the Peking University Third Hospital ethics committee. Informed consent was obtained in this study in accordance with the Declaration of Helsinki.

FUNDING

(2014-ME-14, BMU2018MB004), Beijing Advanced Innovation Center for Genomics, and the State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences.

REFERENCES

1. Kyle RA, Rajkumar SV. Multiple myeloma. N Engl J Med. 2004; 351:1860–73.
   https://doi.org/10.1056/NEJMra041875
   PMID:15509819

2. Palumbo A, Anderson K. Multiple myeloma. N Engl J Med. 2011; 364:1046–60.
   https://doi.org/10.1056/NEJMra1011442
   PMID:21410373

3. Kumar SK, Rajkumar SV. The multiple myelomas - current concepts in cytogenetic classification and therapy. Nat Rev Clin Oncol. 2018; 15:409–21.
   https://doi.org/10.1038/s41571-018-0018-y
   PMID:29686421

4. Durie BG, Salmon SE. A clinical staging system for multiple myeloma. Correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. Cancer. 1975; 36:842–54.
   https://doi.org/10.1002/1097-0142(197509)36:3%3c842::aid-cncr2820360303%3e3.0.co;2-u
   PMID:1182674

5. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Bladé J, Boccadoro M, Child JA, Avet-Loiseau H, Kyle RA, Lahuerta JJ, Ludwig H, Morgan G, et al. International staging system for multiple myeloma. J Clin Oncol. 2005; 23:3412–20.
   https://doi.org/10.1200/JCO.2005.04.242
   PMID:15809451

6. Fonseca R, Bergsagel PL, Drach J, Shaughnessy J, Gutierrez N, Stewart AK, Morgan G, Van Ness B, Chesi M, Minvielle S, Neri A, Barlogie B, Kuehl WM, et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. Leukemia. 2009; 23:2210–21.
   https://doi.org/10.1038/leu.2009.174
   PMID:19798094

7. Palumbo A, Avet-Loiseau H, Oliva S, Lokhorst HM, Goldschmidt H, Rosinol L, Richardson P, Caltagirone S, Lahuerta JJ, Facon T, Bringhen S, Gay F, Attal M, et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. J Clin Oncol. 2015; 33:2863–9.
   https://doi.org/10.1200/JCO.2015.61.2267
   PMID:26240224

8. Köhler M, Greil C, Hudecek M, Lonial S, Raje N, Wäsch R, Engelhardt M. Current developments in immunotherapy in the treatment of multiple myeloma. Cancer. 2018; 124:2075–85.
   https://doi.org/10.1002/cncr.31243
   PMID:29409124

9. Scalzulli E, Grammatico S, Vozella F, Petrucci MT. Proteasome inhibitors for the treatment of multiple myeloma. Expert Opin Pharmacother. 2018; 19:375–86.
   https://doi.org/10.1080/14656566.2018.1441287
   PMID:29478351

10. Diener S, Bayer S, Sabrautzki S, Wieland T, Mentrup B, Przemeck GK, Rathkolb B, Graf E, Hans W, Fuchs H, Horsch M, Schwarzmayr T, Wolf E, et al. Exome sequencing identifies a nonsense mutation in Fam46a associated with bone abnormalities in a new mouse model for skeletal dysplasia. Mamm Genome. 2016; 27:111–21.
    https://doi.org/10.1007/s00335-016-9619-x
    PMID:26803617

11. D'Ambrogio A, Nagaoka K, Richter JD. Translational control of cell growth and malignancy by the CPEBs. Nat Rev Cancer. 2013; 13:283–90.
    https://doi.org/10.1038/nrc3485
    PMID:23446545

12. Fernández-Miranda G, Méndez R. The CPEB-family of proteins, translational control in senescence and cancer. Ageing Res Rev. 2012; 11:460–72.
    https://doi.org/10.1016/j.arr.2012.03.004
    PMID:22542725

13. Kim JH, Richter JD. Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. Mol Cell. 2006; 24:173–83.
    https://doi.org/10.1016/j.molcel.2006.08.016
    PMID:17052452

14. Barnard DC, Ryan K, Manley JL, Richter JD. Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. Cell. 2004; 119:641–51.
    https://doi.org/10.1016/j.cell.2004.10.029
    PMID:15550246

15. Kuchta K, Muszewska A, Knizewski L, Steczkiewicz K, Wyrwicz LS, Pawlowski K, Rychlewski L, Ginalski K. FAM46 proteins are novel eukaryotic non-canonical poly(A) polymerases. Nucleic Acids Res. 2016; 44:3534–48.
    https://doi.org/10.1093/nar/gkw222
    PMID:27060136

16. Kuchta K, Knizewski L, Wyrwicz LS, Rychlewski L, Ginalski K. Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. Nucleic Acids Res. 2009; 37:7701–14.
    https://doi.org/10.1093/nar/gkp854
    PMID:19833706

17. Mroczek S, Chlebowska J, Kuliński TM, Gewartowska O, Gruchota J, Cysewski D, Liudkovska V, Borsuk E, Nowis D, Dziembowski A. The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. Nat Commun. 2017; 8:619.
https://doi.org/10.1038/s41467-017-00578-5
PMID:28931820

18. Barbieri M, Manzoni M, Fabris S, Ciceri G, Todoerti K, Simeon V, Musto P, Cortelezzi A, Baldini L, Neri A, Lionetti M. Compendium of FAM46C gene mutations in plasma cell dyscrasias. Br J Haematol. 2016; 174:642–5.
https://doi.org/10.1111/bjh.13793
PMID:26456599

19. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, Hinton JW, Li Y, Tubio JM, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun. 2014; 5:2997.
https://doi.org/10.1038/ncomms3997
PMID:24429703

20. Boyd KD, Ross FM, Walker BA, Wardell CP, Tapper WJ, Chiecchio L, Dagrada G, Konn ZJ, Gregory WM, Jackson GH, Child JA, Davies FE, Morgan GJ, and NCRI Haematology Oncology Studies Group. Mapping of chromosome 1p deletions in myeloma identifies FAM46C at 1p12 and CDKN2C at 1p32.3 as being genes in regions associated with adverse survival. Clin Cancer Res. 2011; 17:7776–84.
https://doi.org/10.1158/1078-0432.CCR-11-1791
PMID:21994415

21. Prideaux SM, Conway O'Brien E, Chevassut TJ. The genetic architecture of multiple myeloma. Adv Hematol. 2014; 2014:864058.
https://doi.org/10.1155/2014/864058
PMID:24803933

22. Barragán I, Borrego S, Abd El-Aziz MM, El-Ashry MF, Abu-Safieh L, Bhattacharya SS, Antiñolo G. Genetic analysis of FAM46A in Spanish families with autosomal recessive retinitis pigmentosa: characterisation of novel VNTRs. Ann Hum Genet. 2008; 72:26–34.
https://doi.org/10.1111/j.1469-1809.2007.00393.x
PMID:17803723

23. Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, Rice CM. A diverse range of gene products are effectors of the type I interferon antiviral response. Nature. 2011; 472:481–5.
https://doi.org/10.1038/nature09907
PMID:21478870

24. Tian M. The molecular cloning and characterization of Fam46c RNA stability factor. Harvard University; PhD dissertation. 2010.

25. Bianchi G, Ghobrial IM. Biological and Clinical Implications of Clonal Heterogeneity and Clonal Evolution in Multiple Myeloma. Curr Cancer Ther Rev. 2014; 10:70–9.
https://doi.org/10.2174/15733947100214112412404
PMID:25705146

26. Lagali PS, Kakuk LE, Griesinger IB, Wong PW, Ayyagari R. Identification and characterization of C6orf37, a novel candidate human retinal disease gene on chromosome 6q14. Biochem Biophys Res Commun. 2002; 293:356–65.
https://doi.org/10.1016/S0006-291X(02)00228-0
PMID:12054608

27. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976; 194:23–8.
https://doi.org/10.1126/science.959840
PMID:959840

28. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, Sougnez C, Knoechel B, Gould J, Saksena G, Cibulskis K, McKenna A, Chapman MA, et al, and Multiple Myeloma Research Consortium. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell. 2014; 25:91–101.
https://doi.org/10.1016/j.ccr.2013.12.015
PMID:24434212

29. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009; 361:947–57.
https://doi.org/10.1056/NEJMoa0810699
PMID:19692680

30. Berlow RB, Dyson HJ, Wright PE. Hypersensitive termination of the hypoxic response by a disordered protein switch. Nature. 2017; 543:447–51.
https://doi.org/10.1038/nature21705
PMID:28273070

31. Palumbo A, Sezer O, Kyle R, Miguel JS, Orlowski RZ, Moreau P, Niesvizky R, Morgan G, Comenzo R, Sonneveld P, Kumar S, Hajek R, Giralt S, et al, and IMWG. International Myeloma Working Group guidelines for the management of multiple myeloma patients ineligible for standard high-dose chemotherapy with autologous stem cell transplantation. Leukemia. 2009; 23:1716–30.
https://doi.org/10.1038/leu.2009.122
PMID:19494840

32. Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014; 9:171–81.
https://doi.org/10.1038/nprot.2014.006
PMID:24385147

33. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009; 25:1105–11.
https://doi.org/10.1093/bioinformatics/btp120
PMID:19289445

34. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14:R36.
https://doi.org/10.1186/gb-2013-14-4-r36
PMID:23618408

35. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010; 28:511–5.
https://doi.org/10.1038/nbt.1621
PMID:20436464

36. Shi L, Campbell G, Jones WD, Campagne F, Wen Z, Walker SJ, Su Z, Chu TM, Goodsaid FM, Pusztai L, Shaughnessy JD Jr, Oberthuer A, Thomas RS, et al, and MAQC Consortium. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nat Biotechnol. 2010; 28:827–38.
https://doi.org/10.1038/nbt.1665
PMID:20676074

37. Mulligan G, Mitsiades C, Bryant B, Zhan F, Chng WJ, Roels S, Koenig E, Fergus A, Huang Y, Richardson P, Trepicchio WL, Broyl A, Sonneveld P, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood. 2007; 109:3177–88.
https://doi.org/10.1182/blood-2006-09-044974
PMID:17185464

38. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A, Bertsch U, Buijs A, Stevens-Kroef M, Beverloo HB, Vellenga E, Zweegman S, Kersten MJ, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. Blood. 2010; 116:2543–53.
https://doi.org/10.1182/blood-2009-12-261032
PMID:20574050

39. Chauhan D, Tian Z, Nicholson B, Kumar KG, Zhou B, Carrasco R, McDermott JL, Leach CA, Fulciniti M, Kodrasov MP, Weinstock J, Kingsbury WD, Hideshima T, et al. A small molecule inhibitor of ubiquitin-specific protease-7 induces apoptosis in multiple myeloma cells and overcomes bortezomib resistance. Cancer Cell. 2012; 22:345–58.
https://doi.org/10.1016/j.ccr.2012.08.007
PMID:22975377

40. Weinhold N, Ashby C, Rasche L, Chavan SS, Stein C, Stephens OW, Tytarenko R, Bauer MA, Meissner T, Deshpande S, Patel PH, Buzder T, Molnar G, et al. Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. Blood. 2016; 128:1735–44.
https://doi.org/10.1182/blood-2016-06-723007
PMID:27516441

41. Zhou W, Yang Y, Xia J, Wang H, Salama ME, Xiong W, Xu H, Shetty S, Chen T, Zeng Z, Shi L, Zangari M, Miles R, et al. NEK2 induces drug resistance mainly through activation of efflux drug pumps and is associated with poor prognosis in myeloma and other cancers. Cancer Cell. 2013; 23:48–62.
https://doi.org/10.1016/j.ccr.2012.12.001
PMID:23328480

42. Barlogie B, Tricot G, Anaissie E, Shaughnessy J, Rasmussen E, van Rhee F, Fassas A, Zangari M, Hollmig K, Pineda-Roman M, Lee C, Talamo G, Thertulien R, et al. Thalidomide and hematopoietic-cell transplantation for multiple myeloma. N Engl J Med. 2006; 354:1021–30.
https://doi.org/10.1056/NEJMoa053583
PMID:16525139

43. Barlogie B, Tricot G, Rasmussen E, Anaissie E, van Rhee F, Zangari M, Fassas A, Hollmig K, Pineda-Roman M, Shaughnessy J, Epstein J, Crowley J. Total therapy 2 without thalidomide in comparison with total therapy 1: role of intensified induction and posttransplantation consolidation therapies. Blood. 2006; 107:2633–8.
https://doi.org/10.1182/blood-2005-10-4084
PMID:16322468

# SUPPLEMENTARY MATERIALS

Supplementary Figures
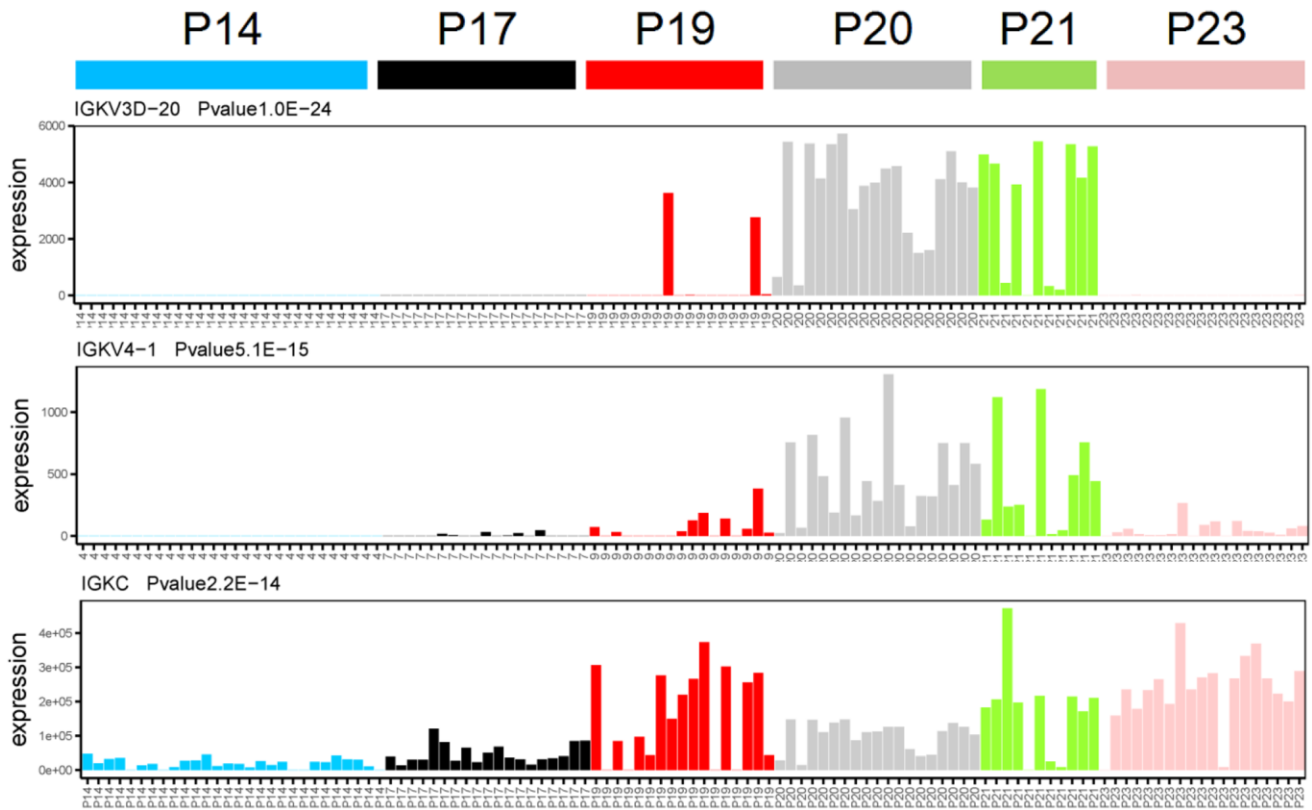
Supplementary Figure 1. Immunoglobulin genes specifically expressed in two to more patients at the single cell transcriptome level. One-way analysis of variance, the gene name and *P*-value were shown in top left corner of each figure.
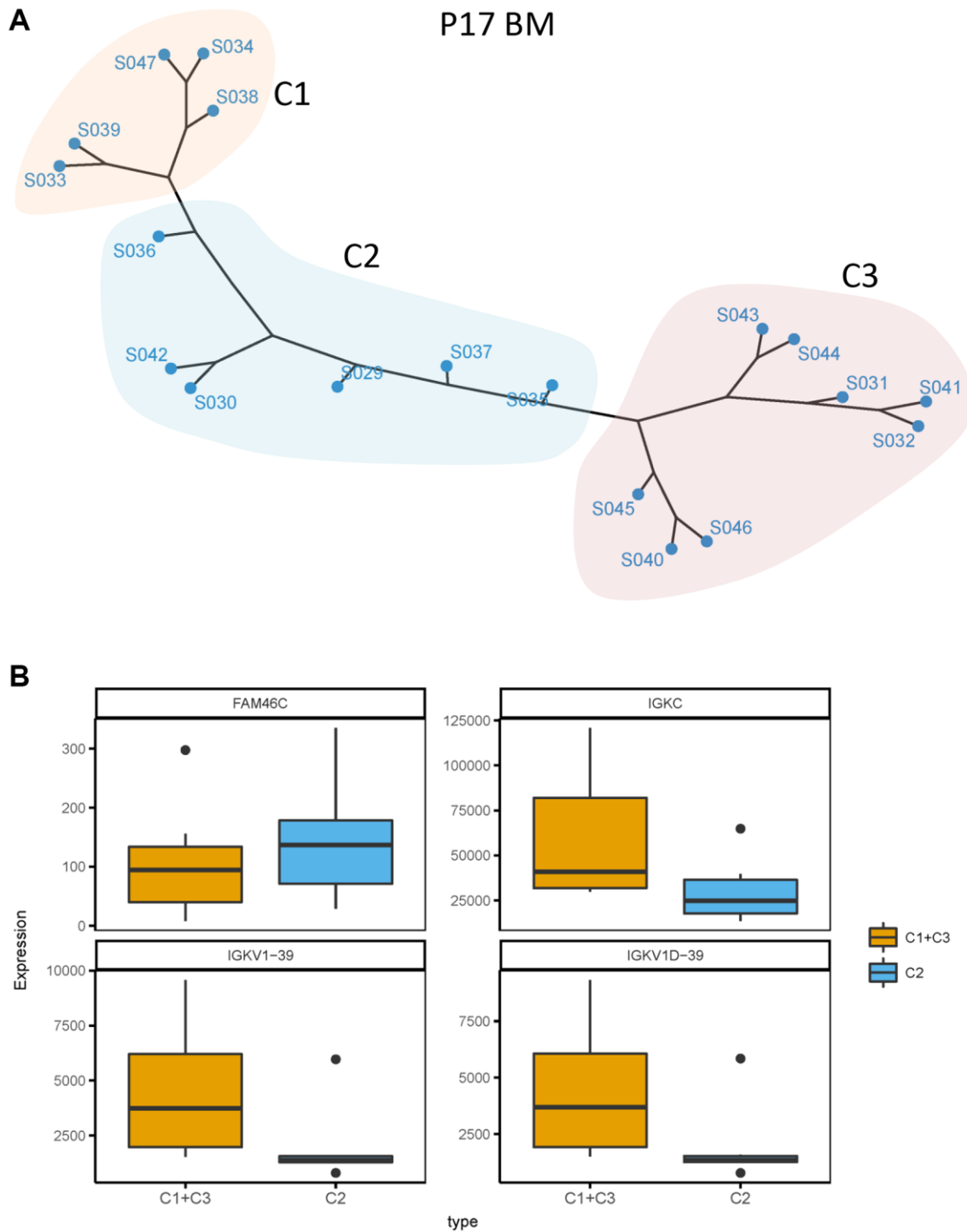
Supplementary Figure 2. Unsupervised clustering of BMMCs of patient P17. (**A**) Unsupervised clustering of BMMCs of patient P17 at the single cell transcriptome level. One point means one single cell of P17 BMMCs. C1, C2, C3 means cluster 1, cluster 2 and cluster 3, showed by 3 different colors. (**B**) FAM46C and immunoglobulin genes trend to differently express in between C1 + C2 and C3. *P* < 0.05, unpaired *t* test, two sided.
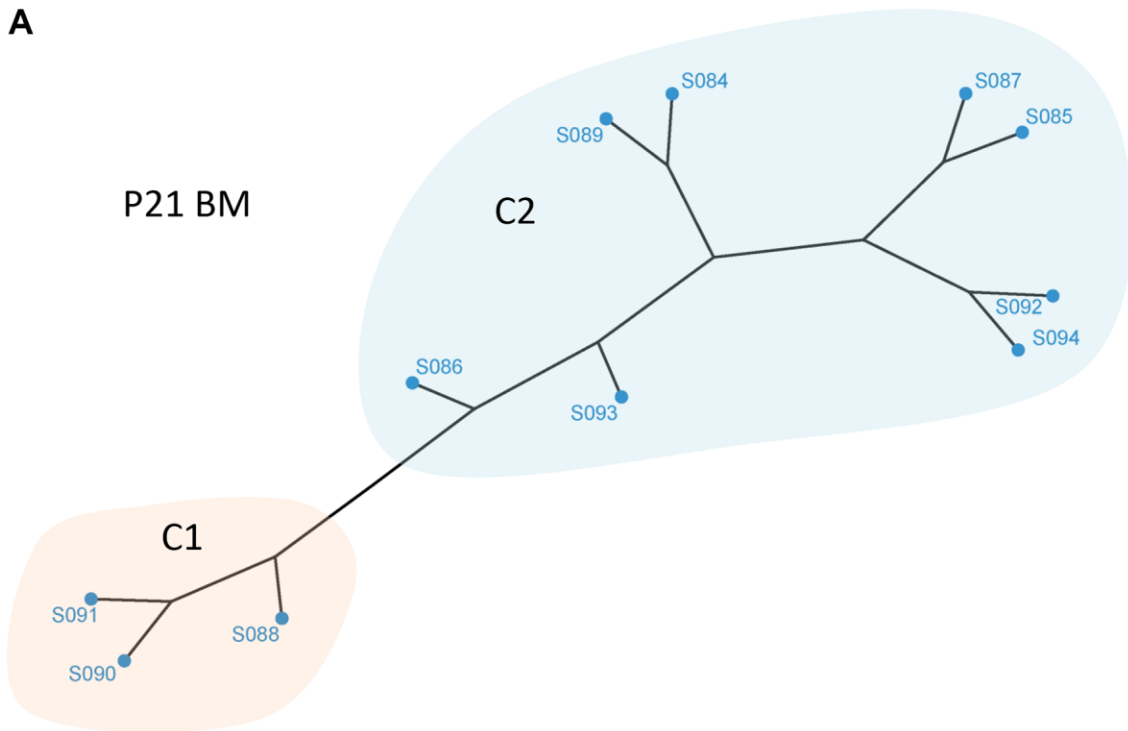
A

B
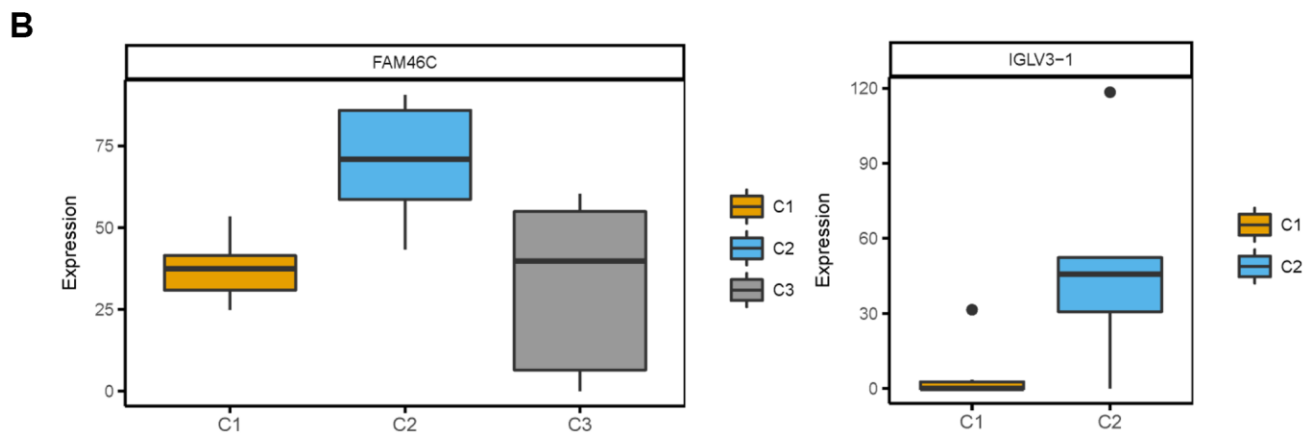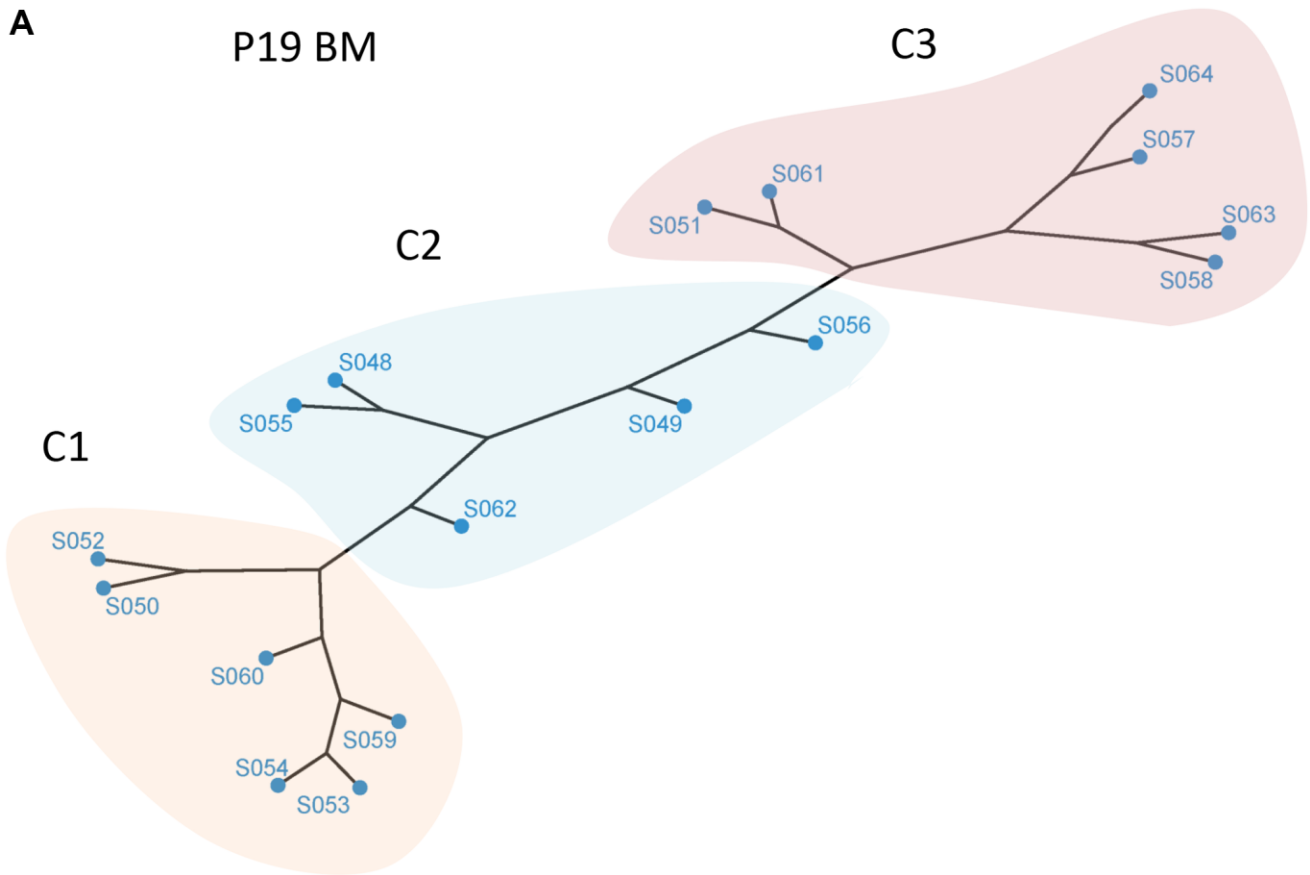
Supplementary Figure 3. Unsupervised clustering of BMMCs of patient P19. (A) Unsupervised clustering of BMMCs of patient P19 at the single cell transcriptome level. One point means one single cell of P19 BMMCs. C1, C2, C3 means cluster 1, cluster 2 and cluster 3, showed by 3 different colors. (B) FAM46C and immunoglobulin genes were differently expressed in between C1, C2 and C3. $P < 0.05$, unpaired t test, two sided.
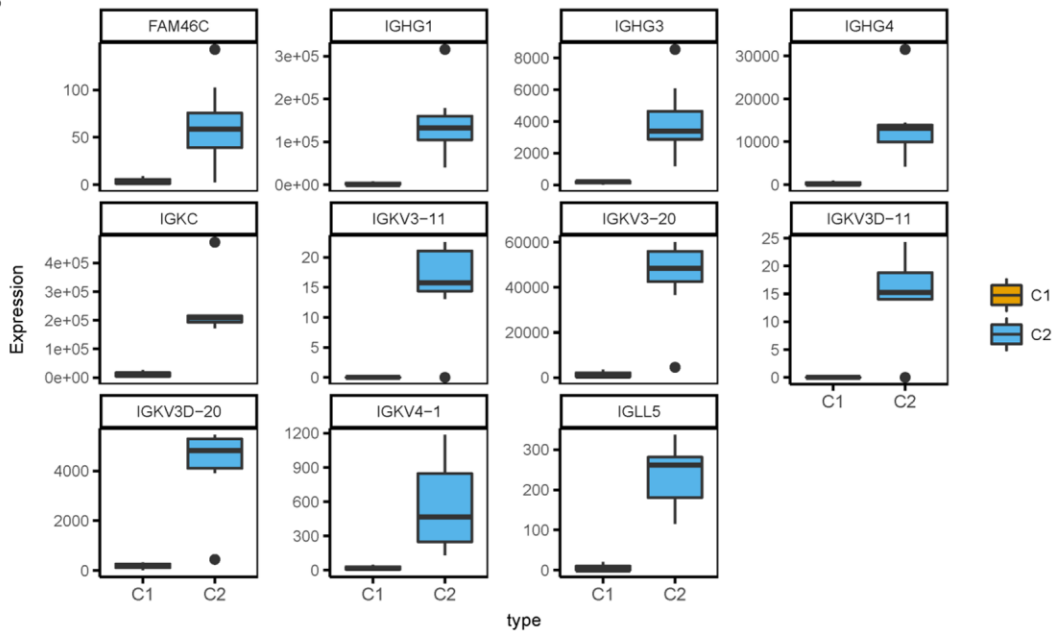
A

P20 BM

B
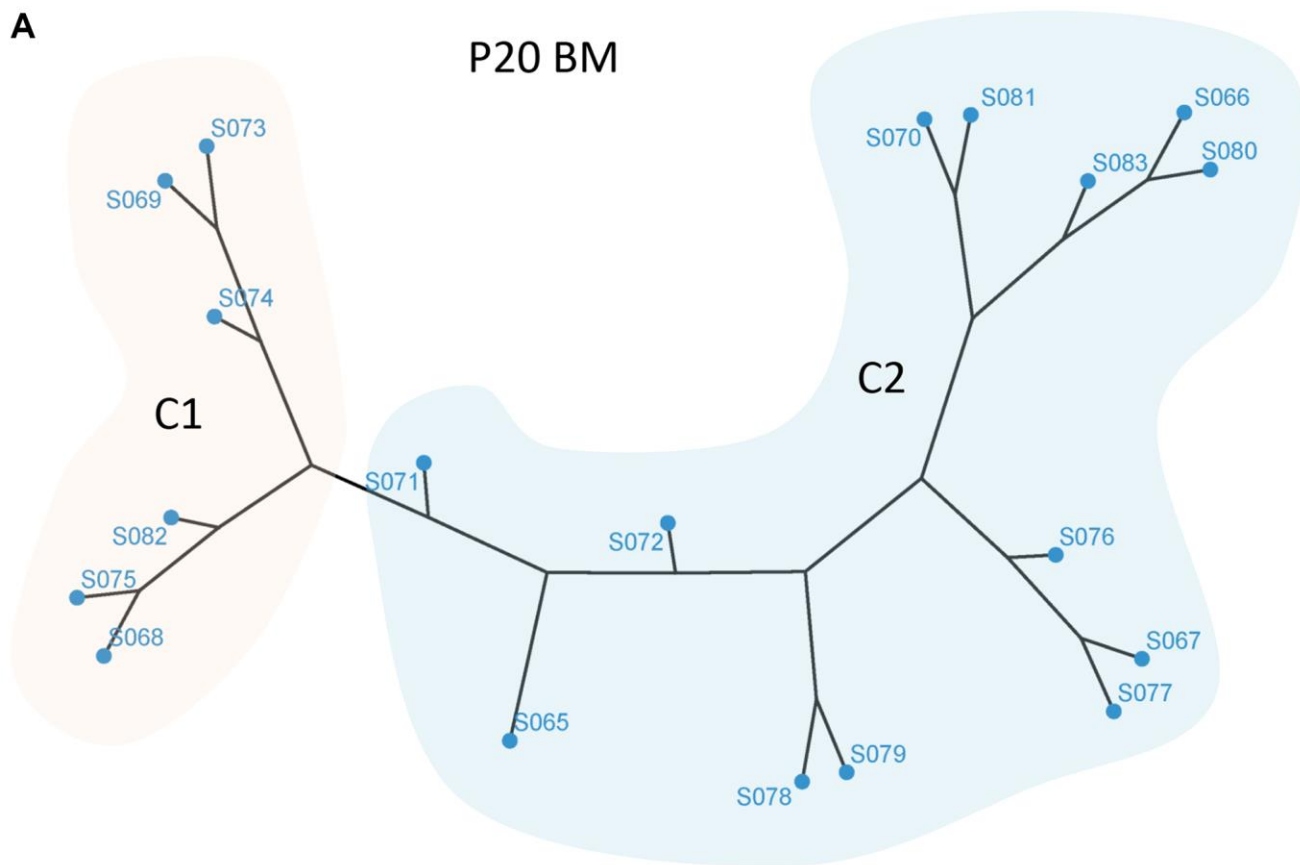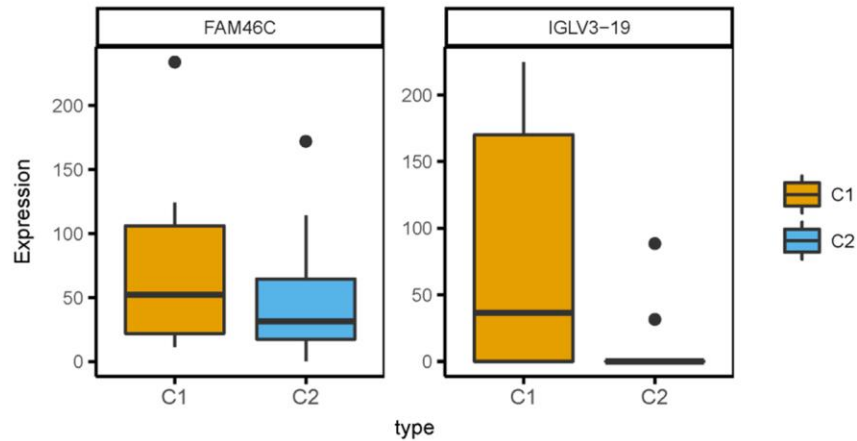
Supplementary Figure 4. Unsupervised clustering of BMMCs of patient P20. (**A**) Unsupervised clustering of BMMCs of patient P20 at the single cell transcriptome level. One point means one single cell of P20 BMMCs. C1, C2 means cluster 1 and cluster 2, showed by 2 different colors. (**B**) FAM46C and immunoglobulin trend to differently express in between C1 and C2. $P < 0.05$, unpaired t test, two sided.
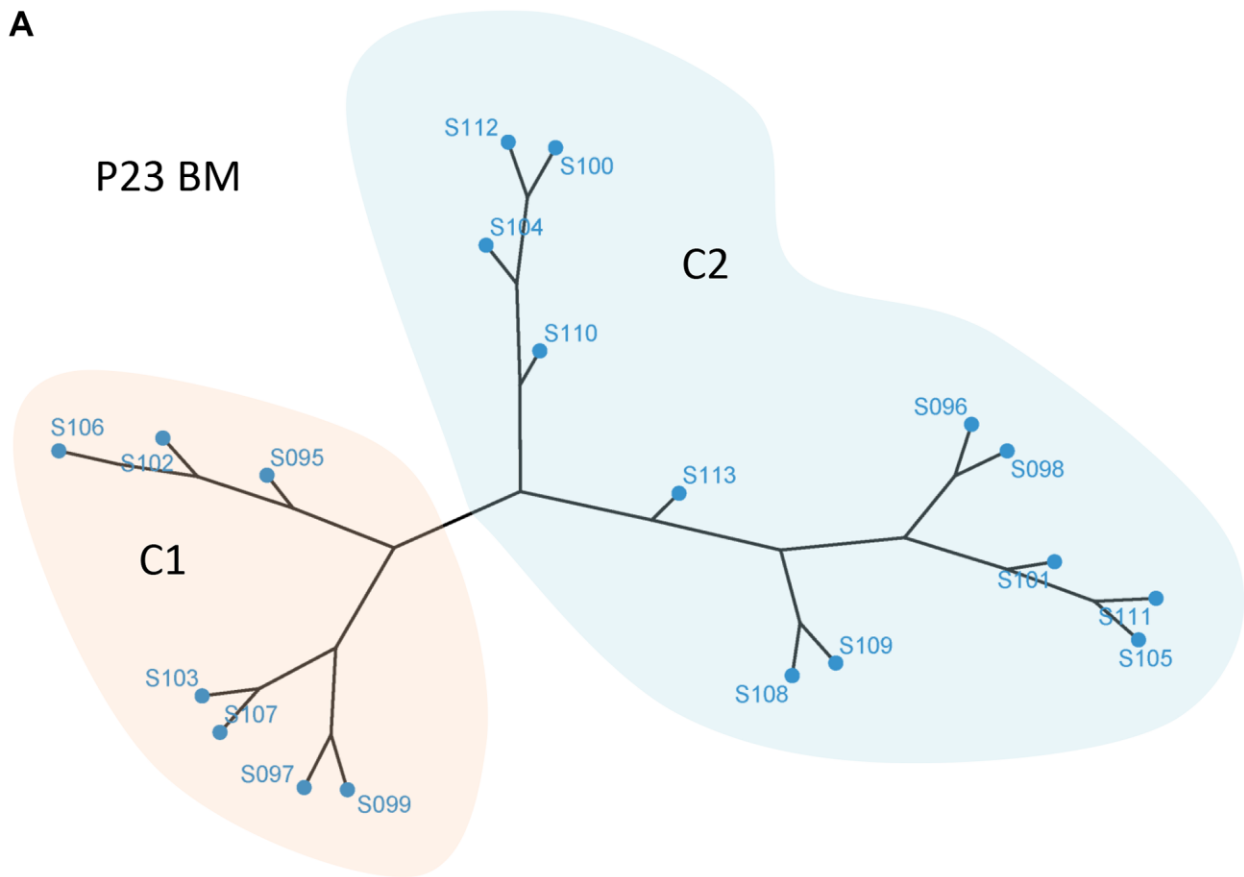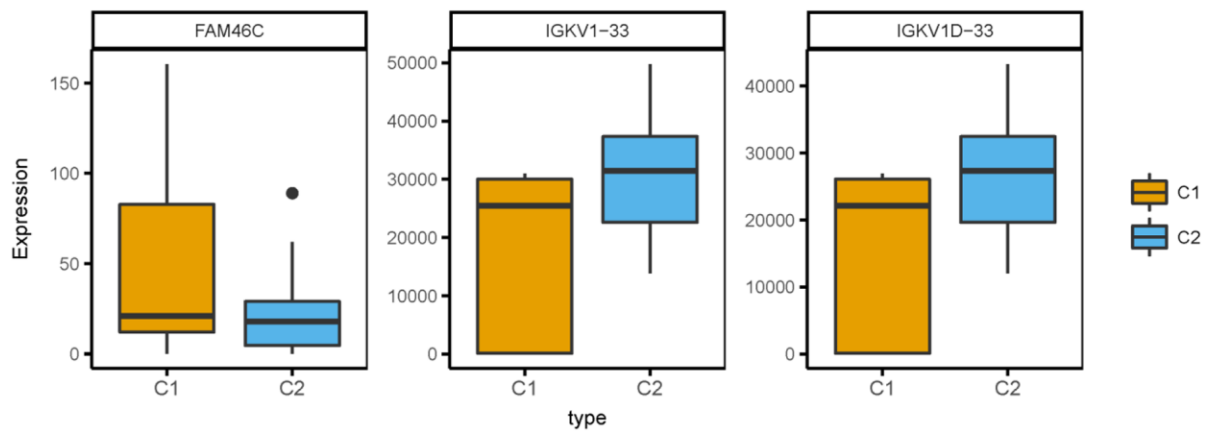
Supplementary Figure 5. Unsupervised clustering of BMMCs of patient P21. (**A**) Unsupervised clustering of BMMCs of patient P21 at the single cell transcriptome level. One point means one single cell of P21 BMMCs. C1, C2 means cluster 1 and cluster 2, showed by 2 different colors. (**B**) FAM46C and immunoglobulin were differently expressed in between C1 and C2. $P < 0.05$, unpaired t test, two sided.

Supplementary Figure 6. Unsupervised clustering of BMMCs of patient P23. (**A**) Unsupervised clustering of BMMCs of patient P23 at the single cell transcriptome level. One point means one single cell of P23 BMMCs. C1, C2 means cluster 1 and cluster 2, showed by 2 different colors. (**B**) FAM46C and immunoglobulin trend to differently express in between C1 and C2. $P < 0.05$, unpaired t test, two sided.
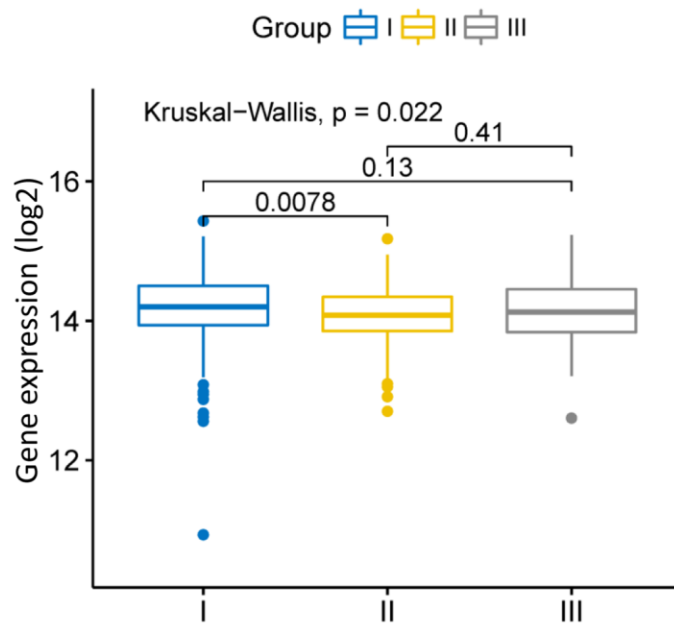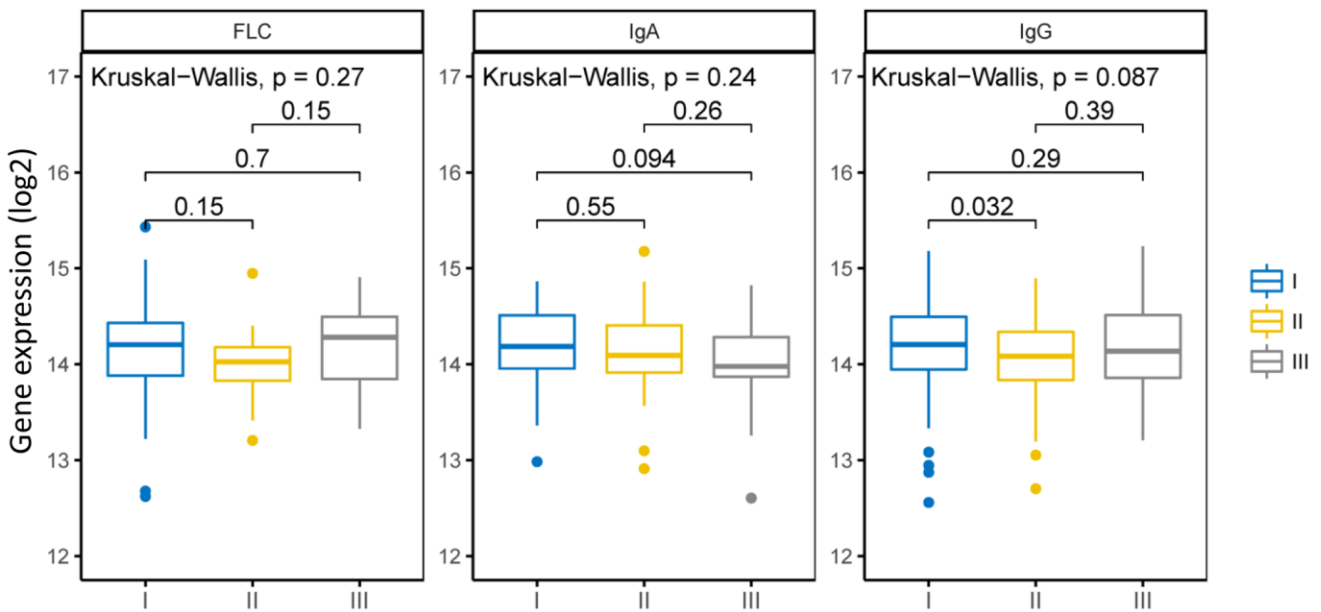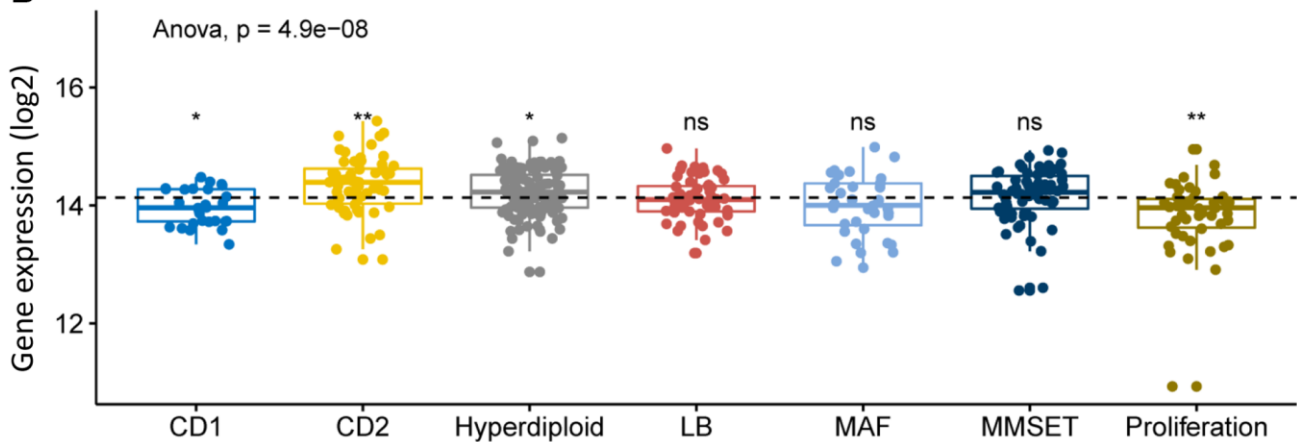
Supplementary Figure 7. FAM46C expression in different ISS stage of 599 multiple myeloma patients. (**A**) FAM46C expression in different ISS stage of 599 multiple myeloma patients. Kruskal-Wallis test. (**B**) FAM46C expression in different ISS stage of FLC, IgG, IgA immunophenotype in 599 multiple myeloma patients. Abbreviation: FLC: free light chain. Kruskal-Wallis test.
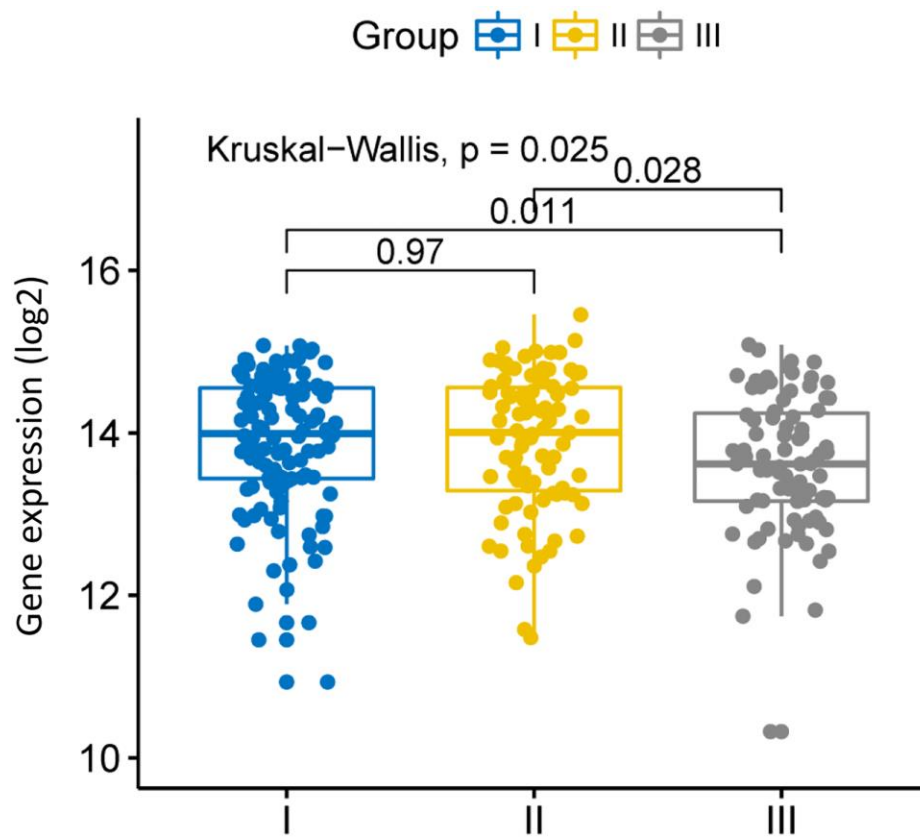
Supplementary Figure 8. FAM46C expression in different ISS stage of another dataset 293 multiple myeloma patients. FAM46C expression in different ISS stage of another dataset (GSE19784) including 293 multiple myeloma patients. Kruskal-Wallis test.
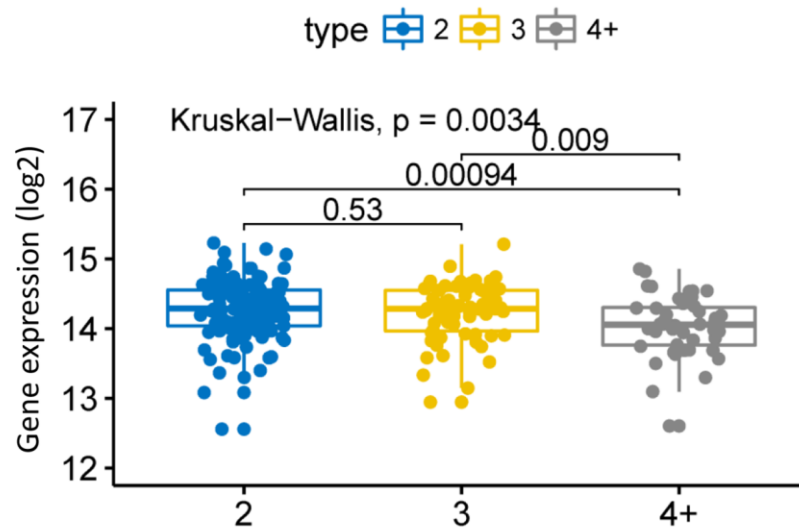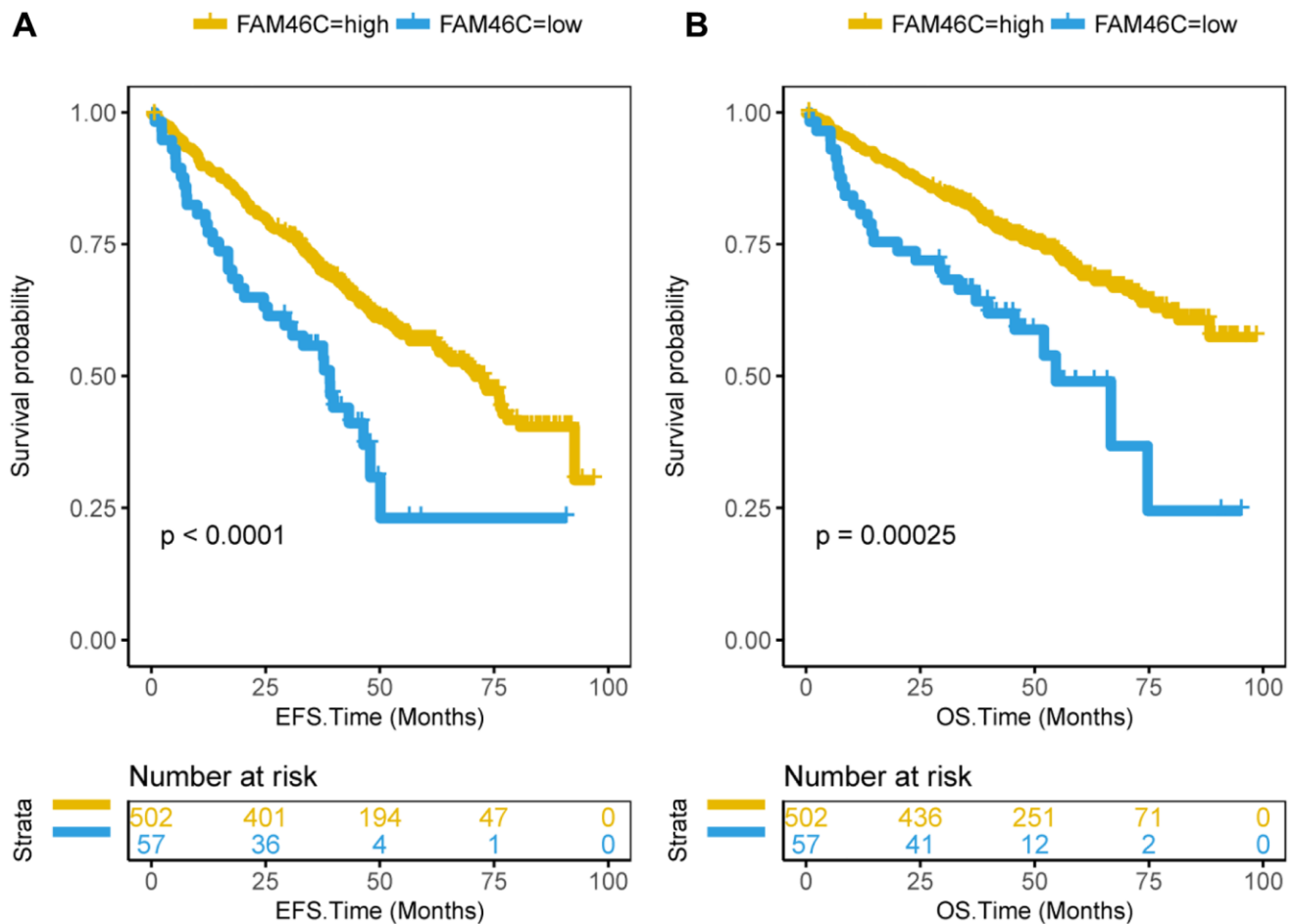
Supplementary Figure 9. FAM46C expression in different molecular subtypes in multiple myeloma patients. (**A**) FAM46C expression in different FISH 1q21 amplification in multiple myeloma patients. Kruskal-Wallis test. (**B**) FAM46C expression in different molecular subtypes in multiple myeloma patients. Kruskal-Wallis test. Symbols indicating statistical significance: ns, $p > 0.05$; $^*p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$. ANOVA (Analysis of variance) test. The mean of the whole dataset was used as reference group. Each of the group levels were compared to the reference group. Horizontal dotted line was added at base mean.
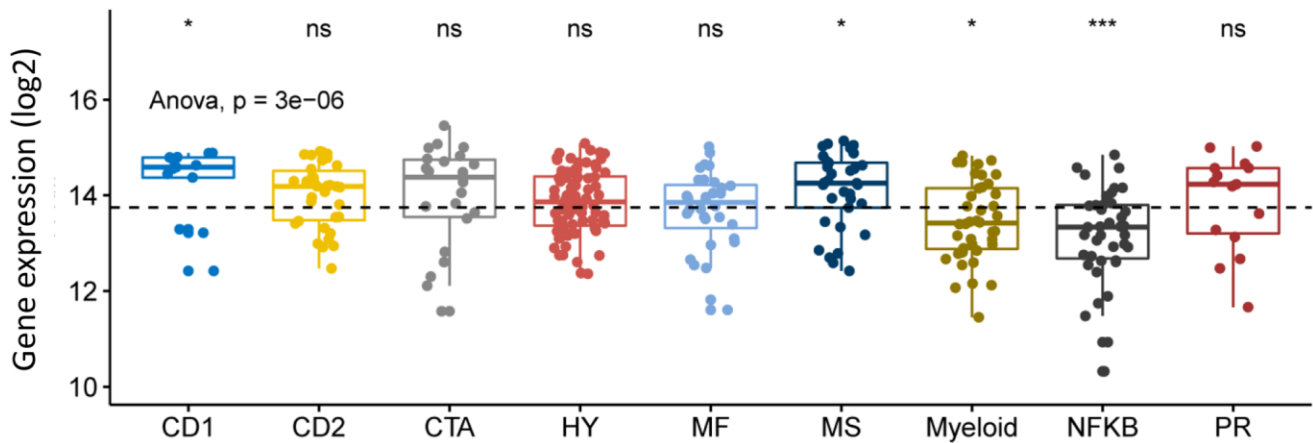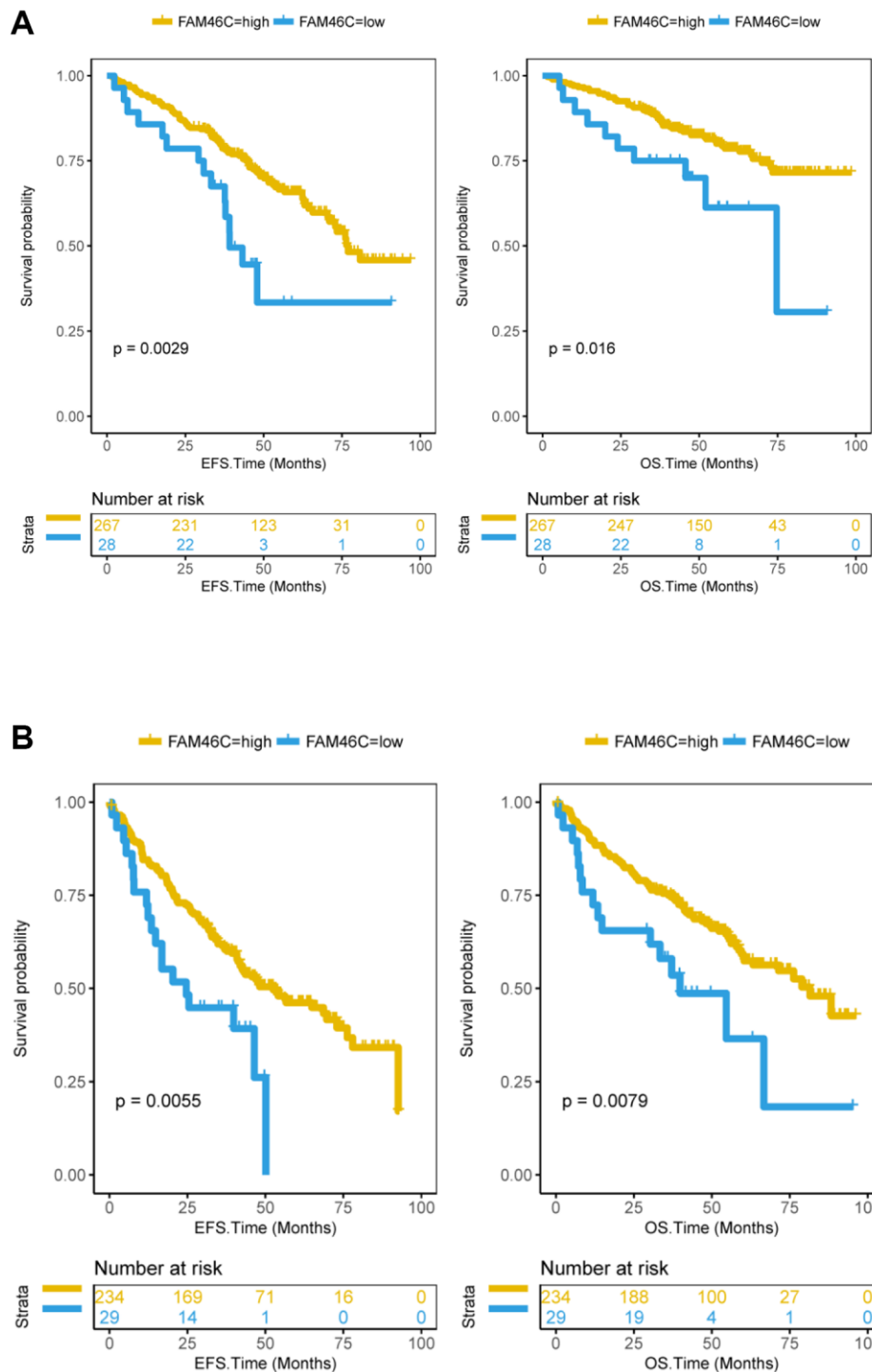
Supplementary Figure 10. FAM46C expression in different molecular subtypes in multiple myeloma patients. FAM46C expression in different molecular subtypes in 308 multiple myeloma patients in another dataset (GSE19784). Kruskal-Wallis test. Symbols indicating statistical significance: ns, $p > 0.05$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$. ANOVA (Analysis of variance) test. The mean of the whole dataset was used as reference group. Each of the group levels were compared to the reference group. Horizontal dotted line was added at base mean.

Supplementary Figure 11. The lower expression of FAM46C predicts poorer survival in multiple myeloma. (**A**) Kaplan-Meier curves for EPS of 559 multiple myeloma patients of FAM46C expression. The log-rank test was used to compare Kaplan–Meier curves. (**B**) Kaplan-Meier curves for OS of 559 multiple myeloma patients of FAM46C expression. The log-rank test was used to compare Kaplan–Meier curves.

Supplementary Figure 12. The lower expression of FAM46C predicts poorer survival in different ISS stage in multiple myeloma. (**A**) Left plot, Kaplan-Meier curves for EPS of 295 ISS stage I multiple myeloma patients of FAM46C expression. The log-rank test was used to compare Kaplan–Meier curves. Right plot, Kaplan-Meier curves for OS of 295 ISS stage I multiple myeloma patients of FAM46C expression. The log-rank test was used to compare Kaplan–Meier curves. (**B**) Left plot, Kaplan-Meier curves for EPS of 263 ISS stage II and III multiple myeloma patients of FAM46C expression. The log-rank test was used to compare Kaplan–Meier curves. Right plot, Kaplan-Meier curves for OS of 263 ISS stage II and III multiple myeloma patients of FAM46C expression. The log-rank test was used to compare Kaplan–Meier curves.
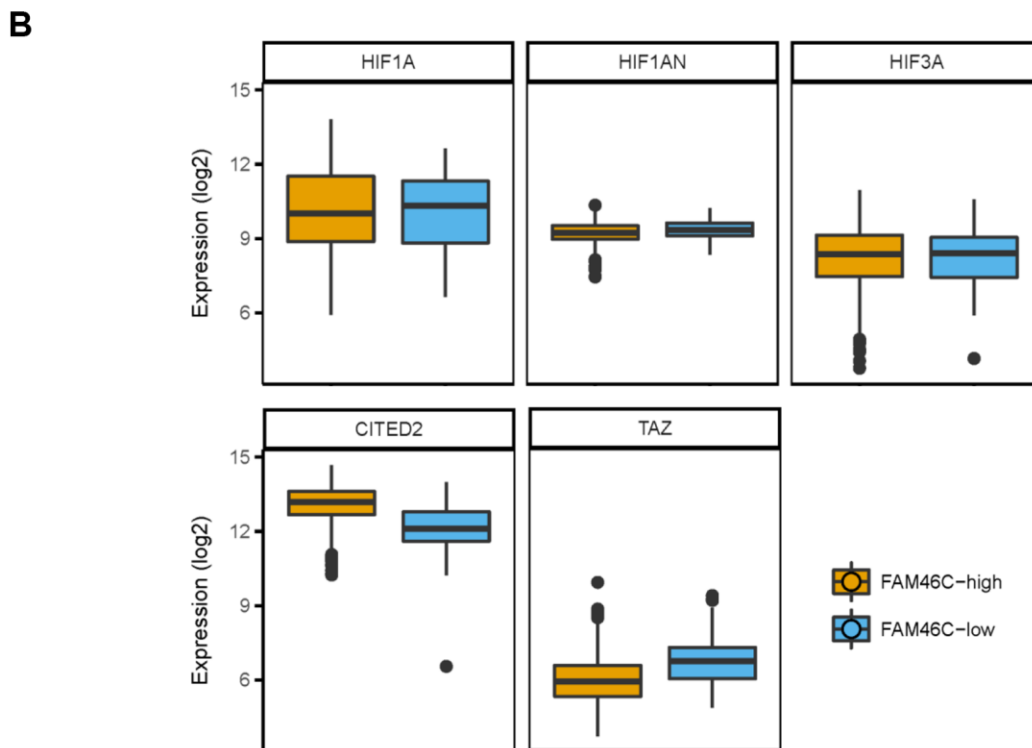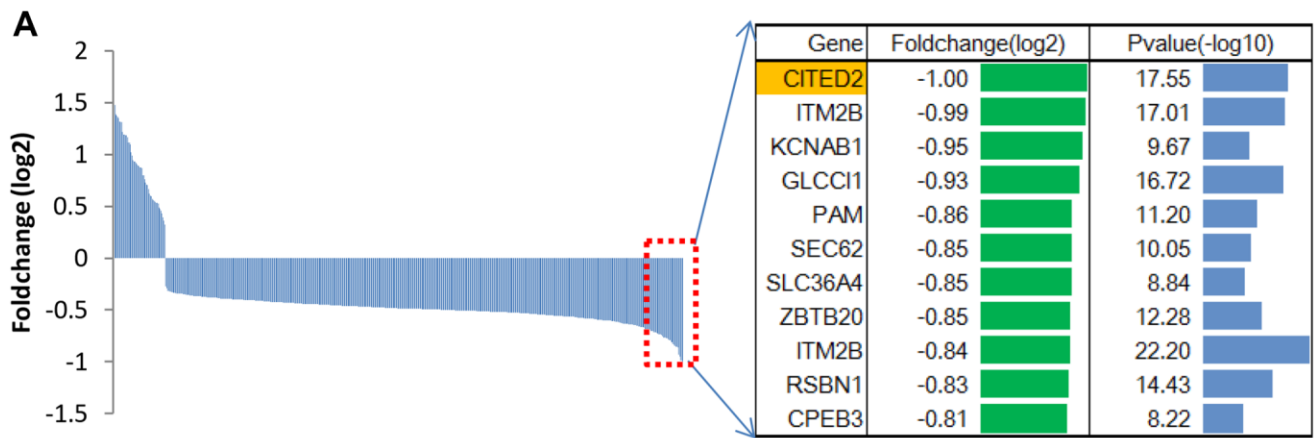
Supplementary Figure 13. The different expression genes in FAM46C-low vs. FAM46C-high group of 559 multiple myelomas. (**A**) All the different expression genes (left) and top 11 (right) in FAM46C-low vs. FAM46C-high group of 559 multiple myelomas. (**B**) Hypoxia related genes expression in FAM46C-low vs. FAM46C-high group of 559 multiple myelomas. Unpaired *t* test, two sided.
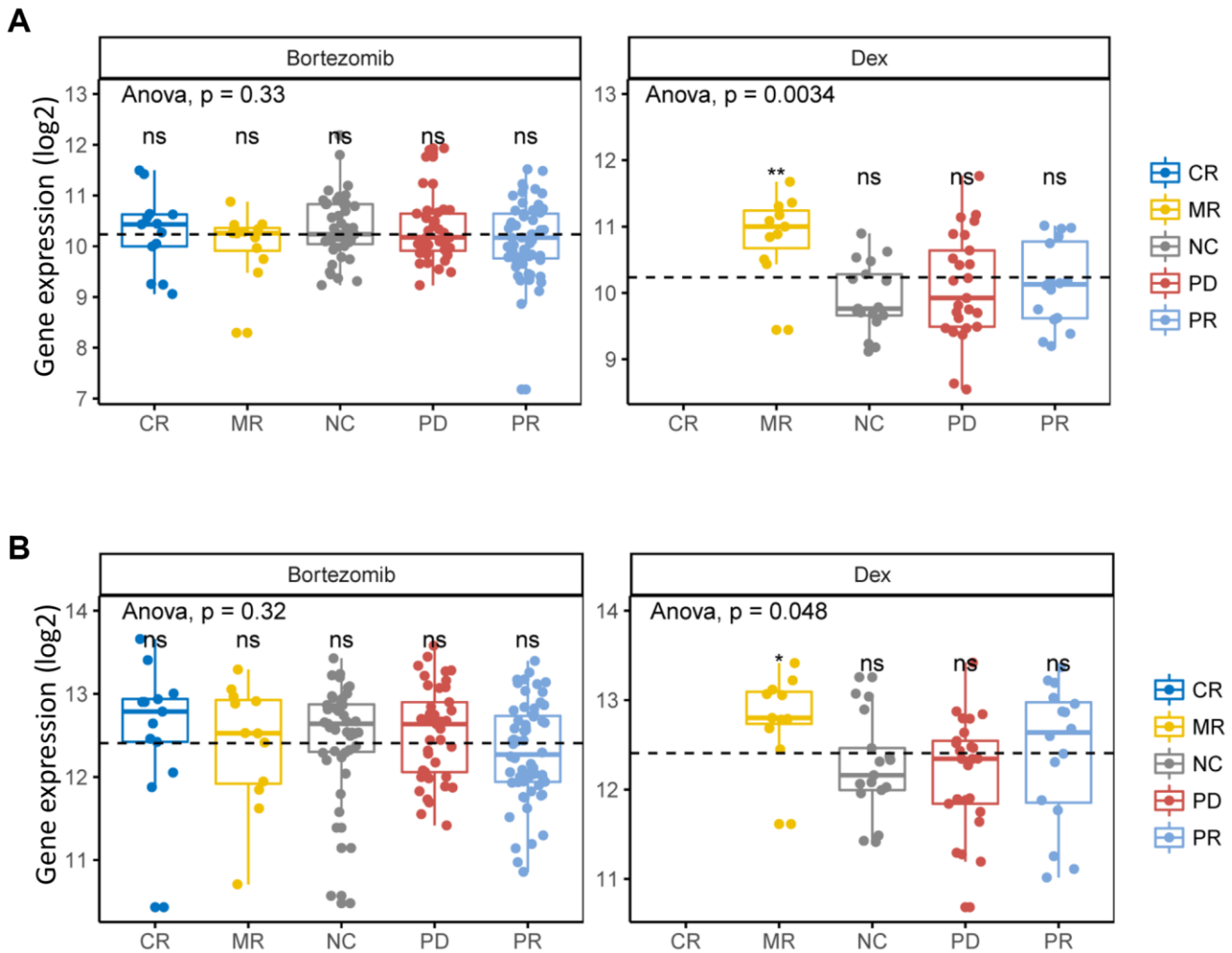
Supplementary Figure 14. FAM46C expression in different therapeutic response with bortezomib or dexamethasone (Dex). (**A**) FAM46C expression in different therapeutic response with bortezomib in 238 multiple myeloma patients (GSE9782) with [HG-U133A] Affymetrix Human Genome U133A Array. (**B**) FAM46C expression in different therapeutic response with bortezomib in 238 multiple myeloma patients (GSE9782) with [HG-U133B] Affymetrix Human Genome U133B Array. Treatment responses: Abbreviations: CR: Complete Response; PR: Partial Response; NC: No Response; PD: Progressive disease; MR: Mini Response. Symbols indicating statistical significance: ns, $p > 0.05$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$. ANOVA (Analysis of variance) test. The mean of the whole dataset was used as reference group. Each of the group levels were compared to the reference group. Horizontal dotted line was added at base mean.
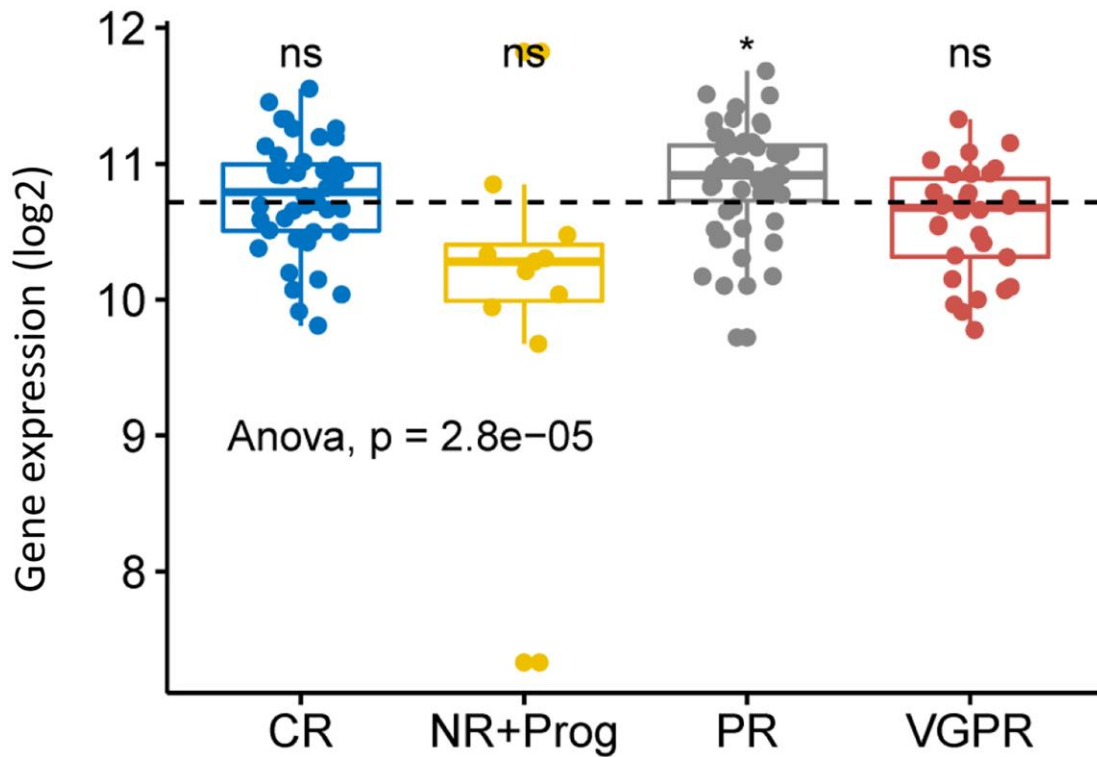
Supplementary Figure 15. FAM46C expression in different therapeutic response. FAM46C expression in different therapeutic response in another dataset (GSE39754) of 136 multiple myeloma patients with Vincristine, Adriamycin, and Dexamethasone (VAD) and Autologous Stem Cell Transplant (ASCT). Treatment responses: Abbreviations: CR: Complete Response; VGPR: Very Good Partial Response; PR: Partial Response; NR: No Response; Prog: Progressive disease. Symbols indicating statistical significance: ns, $p > 0.05$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$. ANOVA (Analysis of variance) test. The mean of the whole dataset was used as reference group. Each of the group levels were compared to the reference group. Horizontal dotted line was added at base mean.
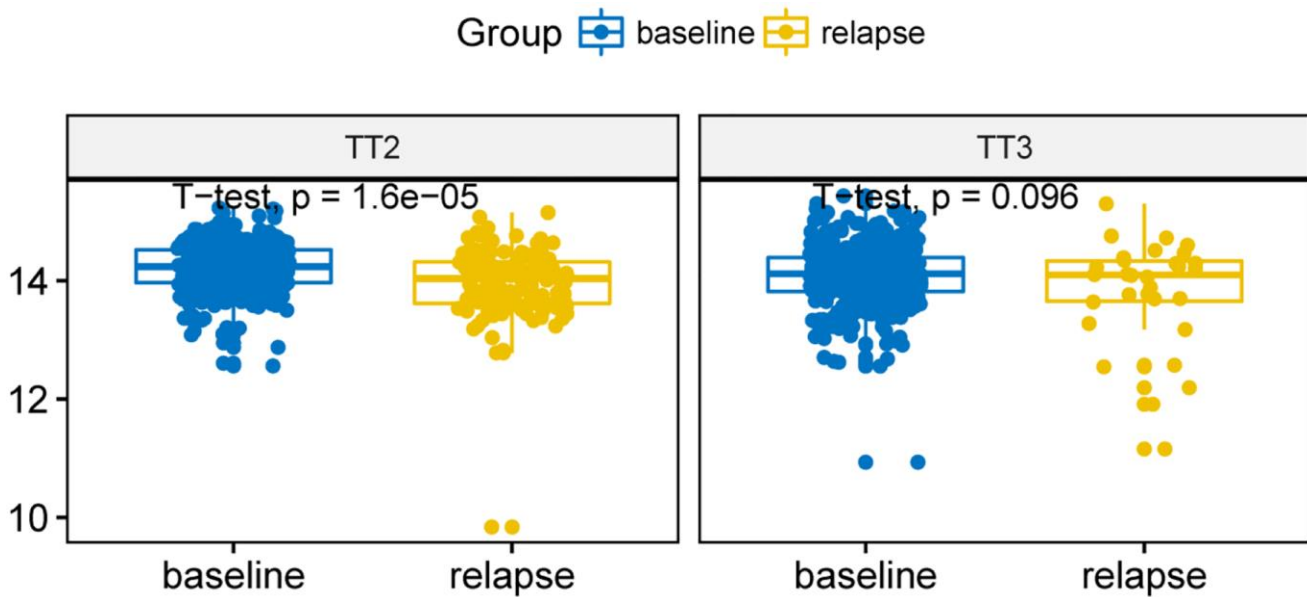
Supplementary Figure 16. FAM46C expression in baseline and relapse samples. FAM46C expression in baseline and relapse samples in GSE31161 dataset with TT2 (left, autologous hematopoietic stem-cell transplantation and thalidomide therapy) and TT3 (right, incorporating bortezomib up-front into a tandem transplant regimen) therapy (totally 937 samples). Unpaired *t* test, two sided.
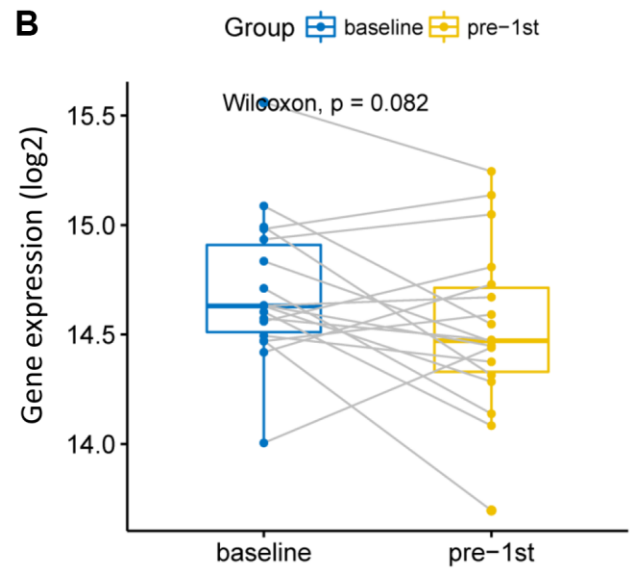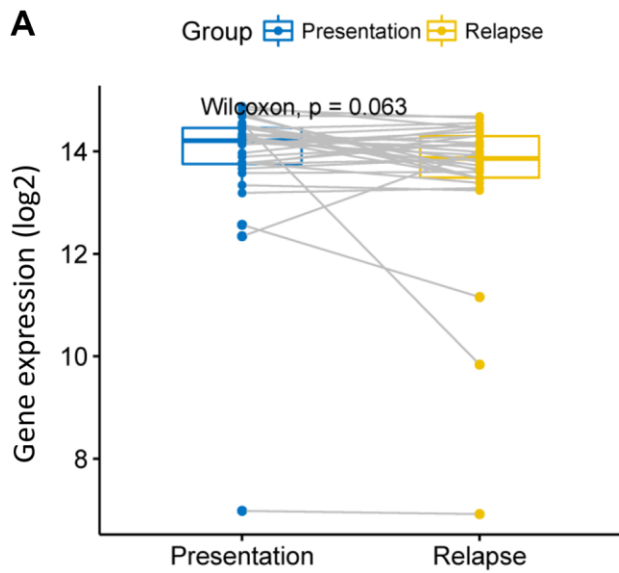
Supplementary Figure 17. FAM46C expression in paired samples of before and after therapy. (**A**) FAM46C expression in presentation (baseline) and relapse paired samples in GSE82307 dataset (totally 33 paired samples). All samples except two proceeded with at least one autologous hematopoietic stem-cell transplantation (ASCT). (**B**) FAM46C expression in baseline and pre-1st (after chemotherapy; pre-1st bone marrow transplant) paired samples in GSE19554 dataset (totally 18 paired samples). Paired *t* test, two sided.

Supplementary Tables

Supplementary Table 1. 95 significantly differently expressed genes were found between BMMCs and cPCs in both P17 and P20 patients.

	PBMC vs. BM	PBMC vs. BM	PBMC vs. BM	PBMC vs. BM
ID	**P17**	**P17**	**P20**	**P20**
Type	**FC**	***P* value**	**FC**	***P* value**
SHC1	−1.7289	0.015	−5.7546	1.3E-14
GANAB	−1.5774	0.00227	−4.233	2.1E-13
SLAMF7	−2.126	0.00029	−3.4001	7.4E-12
IRF4	−1.2338	0.00861	−5.4166	8.8E-11
MAN2A1	−2.2262	0.00874	−4.7581	3.6E-10
MTRNR2L8	2.95762	0.01872	4.85684	7.9E-08
TXNDC5	−1.0376	0.00371	−2.5755	1.3E-07
BTG2	−1.3812	0.00197	−3.3703	1.7E-07
POU2AF1	−2.1768	7E-05	−3.2444	2.7E-07
FAM46C	−1.8532	9.4E-05	−3.5267	2.9E-07
KCNN3	−2.0226	0.01251	−2.9021	4.3E-07
MZB1	−1.0757	0.01224	−2.5107	7.8E-07
SNORD3A	−1.496	0.04556	−2.8061	1.6E-06
NOMO2	−1.5124	0.03537	−2.546	2.8E-06
ALOX5AP	−3.0225	0.0021	−2.9896	3.2E-06
TGOLN2	−2.8476	0.02116	−3.3939	3.2E-06
ZBED6	−1.9243	0.00747	−3.5722	4E-06
CCR2	−3.2103	0.00144	−3.9043	6E-06
NEK9	−2.6786	0.03075	−2.9977	1.1E-05
FNDC3A	−2.5735	0.04964	−3.6116	1.6E-05
FUT8	−2.2422	0.02726	−2.8357	1.6E-05
BSCL2	−2.0737	0.00369	−2.5232	1.7E-05
VOPP1	−1.7465	0.01159	−2.4817	1.9E-05
SULF2	−2.3643	0.00513	−2.7472	2.2E-05
SLC38A2	−1.8483	0.00016	−2.728	3.4E-05
NOMO3	−2.2734	0.00591	−3.2195	4.8E-05
RAC2	−1.8562	0.02608	−2.2361	5.3E-05
SHMT2	−1.9309	0.03362	−3.4989	7.3E-05
SRCAP	−2.1374	0.0006	−2.4755	8E-05
POLR2A	−2.1201	0.01888	−2.4396	9.1E-05
GPR114	−4.1419	0.01344	−4.2574	0.00032
IARS	−2.2027	0.01749	−2.3509	0.00049
ACIN1	−3.2886	0.00304	−3.6991	0.00051
RNF19A	−1.4986	0.01373	−2.2122	0.00053
SIK3	−2.5912	0.01156	−2.7419	0.00063
FASTKD1	−2.9508	0.03705	−3.7125	0.00073
BCOR	−2.6724	0.00168	−3.8076	0.00081
OSBPL3	−2.9286	0.0306	−4.0268	0.00102
IDE	−2.236	0.00328	−3.4248	0.00142
VAC14	−2.1263	0.03366	−3.975	0.00143
AC004447.2	−2.1411	0.00916	−1.8262	0.00154
E2F2	−4.113	0.00813	−3.0402	0.00197
NBEAL2	−3.0182	0.01234	−2.5015	0.00214
ACTR3C	−2.4381	0.01275	−1.5307	0.00253
KANK1	−1.6197	0.04924	−2.9526	0.00264
HIPK2	−2.1466	0.01184	−2.5798	0.00347
IGHV3-23	7.3474	0.0033	3.19899	0.00416
CTD-2619J13.14	−1.3744	0.00127	−2.4502	0.00424

NSF	−1.9021	0.026	−1.7559	0.00521
XRRA1	−2.1513	0.0472	−2.5204	0.0053
DOCK2	−4.5146	0.00152	−2.3046	0.00542
SLC13A3	−1.324	0.03613	−3.1516	0.00638
ZNF460	−1.5418	0.03251	−2.2619	0.00658
TBRG1	−1.9248	0.04335	−3.4894	0.00663
PDS5A	−2.9344	0.00467	−1.8754	0.00668
ADAR	−1.5924	0.01068	−1.8041	0.00705
BAZ2A	−1.1882	0.04463	−2.2956	0.00721
ATAD3C	−1.6287	0.04862	−1.7795	0.00802
MAP4K1	−1.6358	0.04597	−2.2153	0.00814
LRPAP1	−1.8963	0.0166	−1.6959	0.00826
ZBP1	−1.9287	0.04307	−1.8112	0.00928
ELMO1	−3.7704	0.00702	−2.3356	0.00955
FOS	−1.703	0.01442	−2.5704	0.00989
PDXDC2P	−2.5354	0.02918	−2.7736	0.01023
MMACHC	−1.8819	0.02738	−2.0092	0.0103
SEL1L	−1.6216	0.00489	−1.686	0.01078
CNOT4	−2.5722	0.04389	−1.9608	0.01082
CASP10	−3.3902	0.00267	−2.2017	0.01244
SETDB1	−1.9238	0.02903	−2.3407	0.01256
SMARCA2	−3.2094	0.03214	−2.9755	0.01451
NCOA2	−4.0617	0.04477	−2.2601	0.01523
IGLV1-44	7.57289	0.00405	1.93373	0.01524
ST7	−2.0131	0.04718	−3.5356	0.0156
SLC41A1	−2.8949	0.04532	−2.9416	0.01566
RUNX1	−3.0116	0.0217	−2.1007	0.01606
SPTY2D1	−4.126	0.01245	−2.4695	0.01859
CDC73	−1.2557	0.03782	−1.7525	0.01976
PCNX	−3.4407	0.02546	−3.0223	0.02084
DENND1A	−2.3612	0.0236	−1.7931	0.02089
SLC35A4	−1.8958	0.04615	−1.6721	0.02194
HERC2P2	−3.4307	0.01852	−1.7991	0.02406
AP2B1	−2.1699	0.02866	−1.8198	0.02503
DDR2	−3.1547	0.00888	−4.1288	0.02524
NCAPG2	−2.9299	0.02515	−1.298	0.02529
IGHA2	5.23023	0.0025	4.78814	0.02556
BRPF3	−1.9583	0.04551	−1.6205	0.02569
IGHV3-43	3.59169	0.00423	3.96938	0.02608
FTH1	2.61631	0.00019	1.61329	0.02681
C2orf18	−2.383	0.02543	−2.0605	0.0275
UBAP2	−2.1149	0.00816	−2.0553	0.03014
ANGEL1	−3.4752	0.0181	−2.2153	0.03352
PHIP	−1.5215	0.03552	−1.1834	0.04014
RP11-867G23.4	−1.5004	0.0359	−1.2281	0.04141
R3HDM1	−2.0567	0.04427	−2.0878	0.04221
RP11-325F22.2	−2.5729	0.02075	−1.5841	0.04905

Supplementary Table 2. Sample collection and cell detection.

	Patient	BM collected	BMMC detected	PB collected	cPC detected
EMP$^+$	P14	+	28	+	−
($n = 4$)	P17	+	19	+	30
	P20	+	19	+	47
	P23	+	19	N.C.	
EMP$^-$	P19	+	17	+	−
($n = 2$)	P21	+	11	N.C.	

Abbreviations: EMP: Extramedullary plasmacytoma; BM: bone marrow; PB: peripheral blood; +: collected or detected; N.C.: not collected; −: not detected. Number in bracket means number of cells sequenced and used for final data analysis.