A MULTIPLE MYELOMA CLASSIFICATION SYSTEM THAT ASSOCIATES NORMAL B-CELL SUBSET PHENOTYPES WITH PROGNOSIS

Authors:

*Julie Støve Bødker[1]; *Rasmus Froberg Brøndum[1]; *Alexander Schmitz[1]; Anna Amanda Schönherz[12]; Ditte Starberg Jespersen[1]; Mads Sønderkær[1]; Charles Vesteghem[12]; Hanne Due Rasmussen[1]; Caroline Holm Nørgaard[1]; Martin Perez-Andres[1,2]; Mehmet Kemal Samur[3]; Faith Davies[4]; Brian Walker[4]; Charlotte Pawlyn[5]; Martin Kaiser[5]; David Johnson[5]; Uta Bertsch[6]; Annemiek Broyl[7]; Mark van Duin[7]; Rajen Shah[8]; Preben Johansen[9]; Martin Agge Nørgaard[10]; Richard Samworth[8]; Pieter Sonneveld[7]; Hartmut Goldschmidt[6]; Gareth J. Morgan[4]; Alberto Orfao[2]; Nikhil Munshi[3]; Tarec El-Galaly[1,11,12]; **Karen Dybkær[1,11,12]; and **Martin Bøgsted[1,11,12]

*Shared first authorship **Shared senior authorship.

Affiliations:

[1]Department of Haematology, Aalborg University Hospital, Aalborg, Denmark; [2]Institute of Biomedical Research of Salamanca, University of Salamanca, Salamanca, Spain; Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA; [4]UAMS Myeloma Institute, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA; [5]Division of Molecular Pathology, Institute of Cancer Research, London, United Kingdom; [6]National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany; [7]Department of Haematology, Erasmus MC, Rotterdam, Netherlands; [8]Centre for Mathematical Sciences, University of Cambridge, Cambridge, United Kingdom; [9]Department of Haematopathology, Aalborg University Hospital, Aalborg, Denmark; [10]Department of Cardiothoracic Surgery, Aalborg University Hospital, Denmark; [11]Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark, [12]Department of Clinical Medicine, Aalborg University, Denmark.

ABSTRACT

Diagnostic tests for multiple myeloma reflect the criteria from the updated WHO classification based on biomarkers and clinicopathologic heterogeneity. Here we propose a new subtyping of myeloma plasma cells from diagnostic samples, assigned by normal B-cell subset associated gene signatures (BAGS).

For this purpose, we combined fluorescence-activated cell sorting and gene expression profiles from normal bone marrow PreBI, PreBII, immature, naïve, memory, and plasma cell subsets to generate BAGS for assignment of normal bone marrow subtypes in diagnostic samples. The impact of the subtypes was analyzed in eight available data set from 1772 patients' myeloma plasma cell samples.

The resulting tumor assignments in available clinical datasets exhibited similar BAGS subtype frequencies in four cohorts from *de novo* multiple myeloma patients across 1296 individual cases. The BAGS subtypes were significantly associated with progression-free and overall survival in a meta-analysis of 916 patients from three prospective trial cohorts with high-dose melphalan as first line therapy. The major impact was observed within the PreBII and memory subtypes, which had a significantly inferior prognosis compared to other subtypes. A multiple Cox proportional hazard analysis documented that BAGS subtypes added significant, independent prognostic information to the TC classification. BAGS subtype analysis of patient cases identified transcriptonal differences, including a number of differentially spliced genes.

We identified subtype differences in myeloma at diagnosis, with prognostic impact, supporting an acquired B-cell trait and phenotypic plasticity as a pathogenetic hallmark of multiple myeloma.

INTRODUCTION

Despite the extensive insight into multiple myeloma (MM) pathogenesis, as outlined in the WHO classification[1,2], a number of questions remain unanswered regarding the origin, initiation, and developing myeloma cells, including its association with the normal B-cell hierarchy in the bone marrow (BM)[3–6]. We hypothesize that considering MM as a disease of differentiation by identifying it's cell of origin (COO) could lead to novel biological insight and development of new treatment options as described by Boise and coauthors[7].

MM develops from a pre-malignant monoclonal gammopathy of unknown significance (MGUS), by a stepwise oncogenesis to intramedullary early smoldering or evolving *de novo* myeloma because of acquired genetic deregulation[8–10]. The primary translocations implicating the 14q32 locus involve a series of promiscuous target genes, with *CCND1* and *FGFR3/MMSET* being the most frequently present at the MGUS stage[11]. Furthermore, the larger part of breakpoints occurs in the switch regions, suggesting the early translocation happens during IgH class-switch recombination in the germinal center[12–14]. The existence of early translocations and the overexpression of *CCND* genes form the translocations and cyclin D (TC) classification generated from early events[9,11]. Later incidences include a spectrum of mutations and dysregulations occurring in advanced disease with poor prognosis[14–19].

Myeloma plasma cells (PCs) are class-switched, freezing the initiating cell at the post-germinal B-cell maturation stage, refuting that the disease is initiated in earlier B-cell subsets, as have been proposed before[20]. The earliest IgH clonotypic cell we have identified with a class-switched isotype is in the myeloma memory B-cell compartment[21,22], but its clonogenic and malignant potential is a controversial issue[23–26]. Recent studies have concluded that they are remainings of a neoplastic cell with no malignant potential[27,28], contrasting the myeloma PC

compartments.

The myeloma stem cell concept has been reviewed in detail by us and others[29,30]. We proposed an operational definition of COO to allow for acquisition of data supporting the existence of the myeloma stem cell, where a normal B-cell that achieve the first myeloma initiating mutation is not necessarily linearly connected to the myeloma stem cell. These results underpin the hypothesis that myeloma generating cells are present in the malignant PC compartment, but the COO is a normal counterpart of a germinal-center B cell that evolves via differentiation into a premalignant PC compartment already present in MGUS populations.

The plasticity potential of myeloma cells, perhaps caused by interaction with the tumor microenvironment, also plays an important role in development and maintance of MM[30].

The present study takes a COO approach, where we refer to an expanding compartment initiated by a differentiation specific oncogene hit[31]. The terms COO and cancer stem cells have been used interchangeably. However, it is important to differentiate between them as in contrast to our phenotypic COO studies, it is our perception that cancer stem cell research depends on single cell studies in the frame of the classical stem cell definition[29].

The deregulated myeloma cells are under influence by the host as well as the microenvironment may be key in the emergence of myeloma and it's related phenotypic changes. This phenomenon coined plasticity is defined as a changed cellular phenotype or function during deregulated differentiation[32]. More specific, this refers to malignant mature plasma cells that share properties of different maturation steps, including precursors. The phenomenon facilitates a new tool for providing insight into the observed clonal plasticity[33,34] associated with oncogenesis[8–12,17,35–38]. The mechanisms of deregulated differentiation and myeloma-cell plasticity ought to be investigated and their clinical significance assessed.

92 Recently, we have documented a procedure to identify and study gene expression of flow-
93 sorted human B-cell subsets from normal lymphoid tissue[39–42]. These subsets can be profiled
94 and by proper statistical modeling define specific B-cell associated gene signatures (BAGS),
95 recently introduced for DLBCL[43–45]. Here we have applied BAGS from normal BM subsets to
96 assign individual MM subtypes and correlate them with prognosis to delineate their
97 pathogenetic impact.
98

PATIENTS, MATERIAL, AND METHODS

The subtyping method based on normal BM has previously been briefly outlined in Nørgaard et al.[46] and applied to Chronic Lymphocytic Leukemia patients. In this paper we will describe it in more detail and with provide more profound quality controle of the normal samples.

Ethical statement and tissue analysis

All normal tissue samples were collected in accordance with the research protocol (MSCNET, N-20080062MCH) accepted by the North Denmark Regional Committee on Health Research Ethics.

Normal BM was harvested from either sternum (n = 7) during cardiac surgery or taken as aspirates from the iliac crest of healthy volunteers (n = 14). The normal B-cell subsets were phenotyped by multiparametric flow cytometry (MFC) and fluorescence-activated cell sorting (FACS) into six distinct B-cell subsets (PreBI, PreBII, immature (Im), naïve (N), memory (M) B-cells, and PCs) using a monoclonal antibody panel[44]. Gene expression profiles (GEP) for the sternal samples were generated using Affymetrix Exon 1.0 ST while GEP for the iliac crest samples were generated using either Affymetrix Exon 1.0 ST (n = 8) or U133 plus 2.0 (n = 6) arrays. Details on MFC, FACS, and GEP are described in the Supplementary text 1

Clinical myeloma data sets

Data originated from Affymetrix microarray analysis of PC-enriched myeloma samples, from four controlled trials: UAMS, HOVON65/GMMG-HD4, MRC Myeloma IX (all U133 plus 2.0), and APEX (U133A arrays)[35,36,47–50] as well as a preclinical study: IFM-DFCI[37,51] (Human Exon 1.0 ST arrays) as described in Supplementary text 1.

Statistical analysis

All statistical analyses were performed with R version 3.4.3 using Bioconductor packages[52,53].

Below is a summary of the statistical analysis; for full documentation see Supplementary text 1 and detailed session information provided as a Knitr document[54] in the Supplementary text 2. The arrays were cohort-wise background corrected, normalized, and summarized by Robust Multichip Average[55].

BAGS subtyping was build on median-centered gene expression profiles from the sternal BM data summarized at ENSG gene IDs, using regularized multinomial regression, with five discrete outcomes one for each B-cell subset and elastic net penalty[56]. The number of features available in the training of the classifier was prefiltered to only include genes probed by all microarray platforms used in the clinical validation data. Penalization parameters were tuned by leave one out cross validation. To compensate for cohort-wise technical batch effects, each clinical cohort was probe-set-wise median-centered and adjusted to have the same variance as in the sternal BM data. Each patient was BAGS classified according to the class with the highest predicted probability score above 0.40 or otherwise unclassified (**Figure S1**). The robustness of the probability cut-off were thoroughly tested (Supplementary text 2, **Section 11.1**). TC classification was done directely on the RMA normalized samples according to Bergsagel et al.'s algorithm[9].

Kaplan–Meier curves, log-rank tests, and simple and multiple Cox proportional hazards regression were used for survival analsyis. The cohorts ivolving patients from three prospective trial cohorts with high-dose melphalan as first line therapy were amalgamated into a meta-dataset to increase the power of the study. BAGS subtypes as an independent explanatory variable was investigated in the meta-dataset by a Cox proportional hazards regression analysis with BAGS subtypes, TC classes, and cohort as potential confounders. Harell's C-statistic for overall survival was calculated from predicted values from the

multivariate Cox model with and without inclusion of the BAGS classes to asses the prognostic utility.

The samples were assigned resistance probabilities for the drug melphalan by resistance gene signature (REGS) classifiers[57–61].

The significance level was set throughout to 0.05, and effect estimates were provided with 95% confidence intervals. P-values for the differential gene expression and alternative splice analyses were adjusted by Holm's method[62].

RESULTS

BAGS classifier generation and clinical sample assignment

The data quality of the differentiating B-cell subset compartments of the sternal BM was individually validated as illustrated by density plots from MFC analysis (**Figure 1A**) and principal component analysis (**Figure 1B**) of the mean fluorescence intensities (MFIs) of the CD markers used for FACS; unsupervised cluster analysis was also conducted for the gene expression values of the membrane CD markers used for FACS (**Figure 1C**, previously shown in Nørgaard et al.[46]) and 45 classical B-cell markers summarized from a literature search (**Figure 1D**). Subset-specific segregation was further documented by principal component analysis (**Figure S2A-B**).

The BAGS classifiers with the smallest deviance determined by cross validation consisted of 184 genes, for details see Supplementary text 1 and **Figure S3**. Each B-cell subset signature contained 27–54 genes, ensuring comparable gene representation for all subsets in the BAGS classifier. The selected genes and associated coefficents for the BAGS signatures are shown in **Table S1** (previously shown in Nørgaard et al.[46]). The expressed signatures included 51 genes associated with specific B-cell functions, 79 specific genes with more fundamental biological

associations, and 24 probes with unknown gene functions (**Table S2A-F**).

We subsequently validated the BAGS classifier, which was trained using GEP data from Human Exon 1.0 ST Arrays, on independent normal B-cell subsets from BM aspirates from the iliac crest profiled using either Human Exon 1.0 ST or U133 plus 2.0 arrays, and found concordant assignments for both platforms (**Table S3**). This documented the cross-platform validity of the classifier, allowing it's use in clinical data with GEP originating from other platforms.

Microarray data from four independent cohorts (n = 1296) of *de novo* MM patients were assigned for BAGS subtypes. **Table 1** shows the resulting assignment of the tumors and exhibited BAGS subtype frequencies and average percentage for PreBI = 1%, PreBII = 6%, Im = 11%, N = 23%, M = 41%, and PC = 4%, with no significant variation between the cohorts from different geographical regions, time periods, or sampling methods (P = 0.9). We allow 15% of cases within each cohort to be unclassified, resulting in a probability cut-off of approximately 0.40. The distribution of the TC classes within the BAGS subtypes is given in **Table 1**, with significant association identified ($P < 0.001$). There was also significant correlation with ISS staging (P = 0.032), with increased numbers of the PreBII and M subtypes associated with ISS stage III, as shown in **Table 1**. BAGS, proliferation index, and melphalan resistance assignments for all samples used in the analyses are provided in Supplementary data, **Table S4A-H.**

Prognostic impact of assigned BAGS subtypes

Figure 2A-B illustrates the results from a meta-analysis of the 916 patients included in the three prospective trial cohorts with high-dose melphalan as first line therapy (UAMS, HOVON65/GMMG-HD4, and MRC Myeloma IX) with the Affymetrix U133 plus 2.0 microarray data available, documenting that the assigned BAGS subtypes were significantly associated with

progression-free (PFS) and overall survival (OS) (PFS, log-rank test, $P < 0.001$; OS, log-rank test $P < 0.001$). Major impact was observed within patient cohorts with the PreBII and M subtypes, which had a significantly inferior prognosis compared to the patients with Im, N, and PC subtypes.

The robustness of the BAGS association with outcome was successfully evaluated for a wide range of probability cut-offs for the percentage of unclassified cases (Supplementary text 2, **Section 11.1**). The BAGS-assigned MM subtypes in the individual clinical trial data sets UAMS/TT2&3, HOVON/GMMG-HD4, and MRCIX, all including HDM and a variety of new drugs, were also separately analyzed for outcome following treatment as illustrated in **Figure S4A-H**. Results from the individual datasets were in accordance with the above described meta-analysis illustrated in **Figure 2A-B**.

Cox proportional hazard meta-analysis results, as shown in **Table 2** and Harell's C-statistic, giving the concordance between observed survival and predicted risk scores from the a multivariate cox model with ($C = 0.65$) and without ($C = 0.59$) BAGS classes, demonstrated that the BAGS subtypes added significant and independent prognostic information to the already well-established TC classification. In addition, we found significant correlation between the BAGS subtypes and the proliferation index (PI) risk profiling ($P < 0.001$), melphalan resistance probability ($P < 0.001$), and beta-2 microglobulin plasma level ($P < 0.001$) as illustrated in **Figure 3A-C**, respectively. Results in these figures are done on a combined dataset adjusted for differences in individual datasets, while results for individual datasets may be found in Supplementary text 1, **Figure S5A-C**.

BAGS assignment of MGUS, smoldering myeloma, MM, extramedullary MM, and

myeloma cell lines

Available data sets were used for BAGS assignment of associated myeloma diseases, as shown in **Table S5A**. Of interest, five of six plasma cell leukemia cases were M subtypes, indicating a subtype evolution or selection for advanced disease. In contrast, MGUS cases had a significantly high frequency (> 50%) of N subtypes, which was different from smoldering myeloma and newly diagnosed and relapsed MM. The distribution of M component isotypes showed no significant differences across BAGS subtypes, except for a tendency for LCD to be overrepresented in the post germinal subtypes as shown in **Table S5B**. Frequencies of BAGS subtypes in relapsed MM patients from the APEX dataset **(Table S5C)** were similar to the frequencies in first line patients shown in **Table 1**. Finally, we observed that 9 out of 12 human myeloma cancer cell lines were classified as PC subtypes (**Table S6**).

Characterization of BAGS subtypes

Differential expression analysis of BAGS subtypes with poor prognosis (Pre-BII or Memory vs the rest) identified hundreds of genes with a highly significant differential expression as given in **Table S7A-B** for the Pre-BII and Memory subtypes, respectively. GO enrichment of significant genes showed that the Pre-BII subtype myelomas are enriched for the categories mitotic cell cycle, nuclear division and DNA-dependent DNA replication **(Table S8A)**, and Memory subtype myelomas are enriched for the categories cell-cell signaling**,** synaptic transmission**,** multicellular organismal process **(Table S8B)**. For more detail see Supplementary text 1.

Finally, in order to detect whether the Pre-BII and Memory subtypes showed alternative splicing patterns associated with oncogenesis, we investigated alternative exon usage in the IFM-DFCI data set. Results suggested Pre-BII-specific alternative exon usage for 16 genes

(**Table S9A**), which were especially associated with biological processes involved in cell cycle regulation (**Table S10A**). In Memory subtype cases, we only identified 2 candidate genes with potential alternative exon usage (**Table S9B**) associated with basic cell functions including regulation of programmed cell death, metabolism, and signaling transduction (**Table S10B**). Comparison to alternative exon usage patterns detected in non-malignant samples (**Table S11**) indicated that the majority of events were specific to malignant samples, suggesting association to oncogenesis.

DISCUSSION

We have phenotyped distinct cellular subsets of B-cells in the normal BM to generate a BAGS classifier and have documented that the assigned subtypes have prognostic impact.

A probability estimate for each sample to be assigned to each of the six BAGS subtypes was provided. Samples with very low classification probabilities were labeled as unclassified. The frequency of unclassified samples in other gene expression–based COO classifications is around 15%[17]. A pragmatic probability cut-off of 0.40 was used, which is well above the random assignment probability of one out of six, to ensure that 85% of the samples would be BAGS-subtypes. The robustness of the BAGS association with outcome was successfully assessed for a wide series of probability cut-offs.

The present study was whenever possible conducted according to guidelines of –omics-directed medicine, e.g. McShane et al.[63], REMARK[64], and MIAME[65]. However, it is worth noting that the BAGS classifier used cohort-based normalization, which implies that it cannot practically be used in a clinical set-up where patients show up one at a time. Remedies to this problem have been proposed elsewhere[66] and were not further pursued here.

The assignment of BAGS subtypes to MM may explain an inter-individual disease heterogeneity, which could reflect the association between cellular differentiation and oncogenesis[27,67–69]. A standardized flow cytometry immunophenotyping of hematological malignancies, illustrates the potential clinical application of surface expressed markers to identify diagnostic tumor clones[70]. Such a strategy has allowed new studies of normal PC heterogeneity by differentiation[6,10,71].

MM is an example of a malignant disease that has been studied intensively with microarrays. Many peer-reviewed papers have documented new classification systems based

on gene expression profiles to correlate with biology and prognosis[1,8–12]. The present work addresses a need to study a new diagnostic platform defined by the molecular classification of BAGS in MM, as for DLBCL[43,72]. In accordance with our studies in DLBCL, where we applied the whole lymphoid differentiation compartment from tonsils or normal lymph nodes, we prospectively analyzed the lymphoid subset–defined compartments from normal BM to generate six BAGS for MM assignment. The idea was that the COO concept would hold true also for MM and assign subtypes from the post-germinal differentiation pathway. To our surprise, a major fraction of patient tissues were assigned a PreBII, Im, or N subtype, disproving the subtyping to be true reminiscence of the origin from a germinal or post germinal phenotype. Given the phenotypic variation among MGUS, smoldering myeloma, MM, MM relapse, extramedullary MM, plasma cell leukemia, and human myeloma cancer cell lines, it is more likely that BAGS assignment does classify MM cases based on reversible phenotypic plasticity[33].

The BAGS classification is correlated to the well-established TC classification; however, we found that the poor prognosis for PreBII and memory subtypes correlated with the myeloma cell PI and the beta-2 microglobulin plasma, but not gene expression level. PI and beta-2 microglobulin is historically the most important and persistent biomarker in different trials, independent of the evolving therapy. The mechanisms behind these prognostically useful markers are unknown but should now be studied to understand their pathogenetic impact.

Our detection of alternative exon usage suggested subtype specific patterns, supporting that the BAGS phenotyping is based on biological processes. Alternatively spliced candidate genes detected in the Pre-BII subtype revealed an overrepresentation of genes involved in cell cycle regulation and increased proliferation, such as *GTSE1, PKMYT1, BIRC5, and AURKB* , suggesting an association with altered cell cycle regulation and proliferation. However,

289 detected alternative exon usage of candidate genes needs to be experimentally validated and
290 confirmed.

291 HDM forms the basis of MM treatment[73]. However, patients with refractory or relapsed
292 diseases represent a large unmet need for drug-specific predictive tests and precise companion
293 diagnostics[57–61]. This need can be exemplified by REGS and BAGS classification with predictive
294 information to guide therapy. The current analyses indicate that such information is available
295 at diagnosis (**Figure 3B and Figure S5B**) and could be used for identification of candidates for
296 more precise strategies. Collectively, this result indicates BAGS subtypes experiences different
297 clinical tracks and drug resistant mechanisms, and maybe even different molecular
298 pathogenesis. We believe our results support the future inclusion of gene expression profiling
299 in randomized prospective clinical trials aimed at improving MM treatment.

300 BAGS classification divided *de novo* MM patients into so-far-unrecognized, differentiation-
301 dependent prognostic groups. These prognostic analyses and observations support the idea
302 that BAGS classification in MM may contribute with pathogenetic information, especially in
303 attempts to understand the biology behind the classical and still meaningful biomarkers PI and
304 beta-2 microglobulin. Most importantly, the classification included pregerminal subtypes,
305 pointing at a reversible phenotypic plasticity in myeloma PCs. Prospective future studies are
306 needed to prove the concept using clinical endpoints, including prediction of therapeutic
307 outcome.

308 **Acknowledgments**

309 We acknowlegde the tremendous perseverance by the late Professor Hans Erik Johnsen (May
310 13, 1948 – May 17, 2018) in driving the project behind this publication. Ufortunately, he did
311 not experience the final outcome of the work.

This study was supported by research funding from the EU 6th FP to MSCNET (LSHC-CT-2006-037602), the Danish Cancer Society, the Danish Research Agency to CHEPRE (#2101-07-0007), and the KE Jensen Foundation (2006-2010) to HEJ and KD.

Technicians Louise Hvilshøj Madsen and Helle Høholt participated in several aspects of the laboratory work.

Author contributions

M.B. and K.D. designed, collected data, and lead the study; J.S.B., R.F.B., A.S., M.B., and K.D. wrote the manuscript; R.F.B. and M.B. performed statistical analysis; Biological and clinical interpretation was provided by J.S.B., K.D., A.A.S., D.S.J, T.E.G. and H.E.J.; M.P.A., A.O., A.S., R.F.B., A.A.S., M.S., C.V., H.D.R., C.H.N., D.S.J., P.J., M.A.N., R.S., and R.S. contributed vital technology or analytical tools; M.K.S., N.M., F.D., B.W., C.P., M.K., D.J., G.J.M., U.B., H.G., A.B., M.v.D., and P.S. delivered clinical and expression data. All authors read and approved the final version of the manuscript before submission. Hans Erik Johnsen, Department of Hematology, Aalborg University Hospital, Denmark lead the project before his death and therefore contributed in designing the study and drafting the manuscript.

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

Correspondance: Martin Bøgsted, Department of Clinical Medicine, Aalborg University, Sdr. Skovvej 15, DK-9000 Aalborg, Denmark; E-mail: martin.boegsted@rn.dk

REFERENCES

1. Campo E, Swerdlow SH, Harris NL, et al. The 2008 WHO classification of lymphoid neoplasms and beyond : evolving concepts and practical applications. *Blood*. 2011;117(19):5019–32.
2. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classi fi cation of lymphoid neoplasms. *Blood*. 2016;127(20):2375–2391.
3. Rasmussen T, Honoré L, Johnsen HE. Identification and characterisation of malignant cells using RT-PCR on single flow-sorted cells. *Med. Oncol.* 1998;15(2):96–102.
4. Rasmussen T, Haaber J, Dahl IM, et al. Identification of translocation products but not K-RAS mutations in memory B cells from patients with multiple myeloma. *Haematologica*. 2010;95(10):1730–1737.
5. Kim D, Park CY, Medeiros BC, Weissman IL. CD19-CD45 low/- CD38 high/CD138+ plasma cells enrich for human tumorigenic myeloma cells. *Leukemia*. 2012;26(12):2530–7.
6. Cenci S, van Anken E, Sitia R. Proteostenosis and plasma cell pathophysiology. *Curr. Opin. Cell Biol.* 2011;23(2):216–22.
7. Boise LH, Kaufman JL, Bahlis NJ, Lonial S, Lee KP. The Tao of myeloma. *Blood*. 2014;124(12):1873–9.
8. Hallek M, Bergsagel PL, Anderson KC. Multiple myeloma: increasing evidence for a multistep transformation process. *Blood*. 1998;91(1):3–21.
9. Bergsagel PL, Kuehl WM, Zhan F, et al. Cyclin D dysregulation: an early and unifying pathogenic event in multiple myeloma. *Blood*. 2005;106(1):296–303.
10. Zhan F, Hardin J, Kordsmeier B, et al. Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood*. 2002;99(5):1745–57.
11. Bergsagel PL, Kuehl WM. Chromosome translocations in multiple myeloma. *Oncogene*. 2001;20(40):5611–5622.
12. Bergsagel PL, Chesi M, Nardini E, et al. Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc. Natl. Acad. Sci. U. S. A.* 1996;93(24):13931–6.
13. Chesi M, Nardini E, Brents LA, et al. Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of

fibroblast growth factor receptor 3. *Nat. Genet.* 1997;16(3):260–4.

14. Rasmussen T, Theilgaard-Mönch K, Hudlebusch HR, et al. Occurrence of dysregulated oncogenes in primary plasma cells representing consecutive stages of myeloma pathogenesis: indications for different disease entities. *Br. J. Haematol.* 2003;123(2):253–62.
15. Davies FE, Dring AM, Li C, et al. Insights into the multistep transformation of MGUS to myeloma using microarray expression analysis. *Blood*. 2003;102(13):4504–11.
16. Rasmussen T, Kuehl M, Lodahl M, Johnsen HE, Dahl IMS. Possible roles for activating RAS mutations in the MGUS to MM transition and in the intramedullary to extramedullary transition in some plasma cell tumors. *Blood*. 2005;105(1):317–23.
17. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467–472.
18. Egan JB, Shi CX, Tembe W, et al. Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood*. 2012;120(5):1060–1066.
19. Leich E, Weißbach S, Klein H-U, et al. Multiple myeloma is affected by multiple and heterogeneous somatic mutations in adhesion- and receptor tyrosine kinase signaling molecules. *Blood Cancer J.* 2013;3(2):e102–e102.
20. Vicente-Dueñas C, Romero-Camarero I, González-Herrero I, et al. A novel molecular mechanism involved in multiple myeloma development revealed by targeting MafB to haematopoietic progenitors. *EMBO J.* 2012;31(18):3704–3717.
21. Rasmussen T, Lodahl M, Hancke S, Johnsen HE. In multiple myeloma clonotypic CD38-/CD19+ / CD27+ memory B cells recirculate through bone marrow, peripheral blood and lymph nodes. *Leuk. Lymphoma*. 2004;45(7):1413–7.
22. Matsui W, Wang Q, Barber JP, et al. Clonogenic multiple myeloma progenitors, stem cell properties, and drug resistance. *Cancer Res.* 2008;68(1):190–7.
23. Paino T, Ocio EM, Paiva B, et al. CD20 positive cells are undetectable in the majority of multiple myeloma cell lines and are not associated with a cancer stem cell phenotype. *Haematologica*. 2012;97(7):1110–1114.
24. Yaccoby S, Epstein J. The proliferative potential of myeloma plasma cells manifest in the SCID-hu host. *Blood*. 1999;94(10):3576–3582.

25. Guikema JEJ, Vellenga E, Bakkus MHC, Bos NA. Myeloma clonotypic B cells are hampered in their ability to undergo B-cell differentiation in vitro. *Br. J. Haematol.* 2002;119(1):54–61.
26. Rasmussen T, Jensen L, Johnsen HE. The clonal hierachy in multiple myeloma. *Acta Oncol.* 2000;39(7):765–70.
27. Pfeifer S, Perez-Andres M, Ludwig H, Sahota SS, Zojer N. Evaluating the clonal hierarchy in light-chain multiple myeloma: implications against the myeloma stem cell hypothesis. *Leukemia*. 2011;25(7):1213–6.
28. Van Valckenborgh E, Matsui W, Agarwal P, et al. Tumor-initiating capacity of CD138- and CD138+ tumor cells in the 5T33 multiple myeloma model. *Leukemia*. 2012;26(6):1436–9.
29. Johnsen HE, Bøgsted M, Schmitz A, et al. The myeloma stem cell concept, revisited: from phenomenology to operational terms. *Haematologica*. 2016;101(12):1451–1459.
30. Hajek R, Okubote SA, Svachova H. Myeloma stem cell concepts, heterogeneity and plasticity of multiple myeloma. *Br. J. Haematol.* 2013;163(5):551–564.
31. Visvader JE. Cells of origin in cancer. *Nature*. 2011;469(7330):314–22.
32. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011;144(5):646–674.
33. Yaccoby S. The phenotypic plasticity of myeloma plasma cells as expressed by dedifferentiation into an immature, resilient, and apoptosis-resistant phenotype. *Clin. Cancer Res.* 2005;11(21):7599–606.
34. Bam R, Khan S, Ling W, et al. Primary myeloma interaction and growth in coculture with healthy donor hematopoietic bone marrow. *BMC Cancer*. 2015;15(1):864.
35. Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. *Blood*. 2006;108(6):2020–8.
36. Mulligan G, Mitsiades C, Bryant B, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. *Blood*. 2007;109(8):3177–3188.
37. Avet-Loiseau H, Li C, Magrangeas F, et al. Prognostic Significance of Copy-Number Alterations in Multiple Myeloma. *J. Clin. Oncol.* 2009;27(27):4585–4590.
38. Samur MK, Shah PK, Wang X, et al. The shaping and functional consequences of the dosage

effect landscape in multiple myeloma. *BMC Genomics*. 2013;14(1):672.

39. Johnsen HE, Bergkvist KS, Schmitz A, et al. Cell of origin associated classification of B-cell malignancies by gene signatures of the normal B-cell hierarchy. *Leuk. Lymphoma*. 2014;55(6):1251–1260.
40. Kjeldsen MK, Perez-Andres M, Schmitz A, et al. Multiparametric flow cytometry for identification and fluorescence activated cell sorting of five distinct B-cell subpopulations in normal tonsil tissue. *Am. J. Clin. Pathol.* 2011;136(6):960–9.
41. Bergkvist KS, Nyegaard M, Bøgsted M, et al. Validation and implementation of a method for microarray gene expression profiling of minor B-cell subpopulations in man. *BMC Immunol.* 2014;15(1):3.
42. Rasmussen SM, Bilgrau AE, Schmitz A, et al. Stable phenotype of B-cell subsets following cryopreservation and thawing of normal human lymphocytes stored in a tissue biobank. *Cytom. Part B - Clin. Cytom.* 2015;88(1):40–49.
43. Dybkær K, Bøgsted M, Falgreen S, et al. Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. *J. Clin. Oncol.* 2015;33(12):1379–88.
44. Bergkvist KS, Nørgaard MA, Bøgsted M, et al. Characterization of memory B cells from thymus and its impact for DLBCL classification. *Exp. Hematol.* 2016;44(10):982–990.e11.
45. Michaelsen TY, Richter J, Brøndum RF, et al. A B-cell-associated gene signature classification of Diffuse Large B-cell Lymphoma by the NanoString technology. *Under Rev. Blood Adv.* 2019;2(13):1542–1546.
46. Nørgaard CH, Jakobsen LH, Gentles AJ, et al. Subtype assignment of CLL based on B-cell subset associated gene signatures from normal bone marrow – A proof of concept study. *PLoS One*. 2018;13(3):e0193249.
47. Barlogie B, Mitchell A, van Rhee F, et al. Curing myeloma at last: defining criteria and providing the evidence. *Blood*. 2014;124(20):3043–51.
48. Broyl A, Hose D, Lokhorst H, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood*. 2010;116(14):2543–2553.
49. Morgan GJ, Davies FE, Gregory WM, et al. Cyclophosphamide, thalidomide, and dexamethasone as induction therapy for newly diagnosed multiple myeloma patients destined for autologous stem-cell transplantation: MRC Myeloma IX randomized trial

results. *Haematologica*. 2012;97(3):442–450.

50. Richardson PG, Sonneveld P, Schuster MW, et al. Bortezomib or High-Dose Dexamethasone for Relapsed Multiple Myeloma. *N. Engl. J. Med.* 2005;352(24):2487–2498.

51. Chauhan D, Tian Z, Nicholson B, 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(3):345–358.

52. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;5(10):R80.

53. R Core Team. R : A Language and Environment for Statistical Computing. Vienna, Austria, Austria: R Foundation for Statistical Computing, Vienna, Austria; 2017.

54. Xie Y. knitr: A General-Purpose Package for Dynamic Report Generation in R. 2017;

55. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4(2):249–64.

56. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* 2010;33(1):1–22.

57. Boegsted M, Holst JM, Fogd K, et al. Generation of a predictive melphalan resistance index by drug screen of B-cell cancer cell lines. *PLoS One*. 2011;6(4):e19322.

58. Bøgsted M, Bilgrau AE, Wardell CP, et al. Proof of the concept to use a malignant B cell line drug screen strategy for identification and weight of melphalan resistance genes in multiple myeloma. *PLoS One*. 2013;8(12):e83252.

59. Laursen MB, Falgreen S, Bødker JS, et al. Human B-cell cancer cell lines as a preclinical model for studies of drug effect in diffuse large B-cell lymphoma and multiple myeloma. *Exp. Hematol.* 2014;42(11):927–938.

60. Falgreen S, Laursen MB, Bødker JS, et al. Exposure time independent summary statistics for assessment of drug dependent cell line growth inhibition. *BMC Bioinformatics*. 2014;15(1):168.

61. Falgreen S, Dybkær K, Young KH, et al. Predicting response to multidrug regimens in cancer patients using cell line experiments and regularised regression models. *BMC Cancer*. 2015;15:235.

62. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.*

1979;6(2):65–70.

63. McShane LM, Cavenagh MM, Lively TG, et al. Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration. *BMC Med.* 2013;11:220.

64. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies. *J. Clin. Oncol.* 2005;23(36):9067–72.

65. Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 2001;29(4):365–71.

66. Falgreen S, Ellern Bilgrau A, Brøndum RF, et al. hemaClass.org: Online One-By-One Microarray Normalization and Classification of Hematological Cancers for Precision Medicine. *PLoS One*. 2016;11(10):.

67. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481(7381):306–13.

68. Corre J, Munshi N, Avet-Loiseau H. Genetics of multiple myeloma: another heterogeneity level? *Blood*. 2015;125(12):1870–1876.

69. Kuehl WM, Bergsagel PL. Multiple myeloma: evolving genetic events and host interactions. *Nat. Rev. Cancer*. 2002;2(3):175–87.

70. van Dongen JJM, Lhermitte L, Böttcher S, et al. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. *Leukemia*. 2012;26(9):1908–75.

71. Paiva B, Perez-Andres M, Gutierrez ML, et al. CD56+ Clonal Plasma Cells In Multiple Myeloma Are Associated With Unique Disease Characteristics and Have a Counterpart Of CD56+ Normal Plasma Cells With Increased Maturity. *Blood*. 2013;122(21):.

72. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503–11.

73. Engelhardt M, Terpos E, Kleber M, et al. European Myeloma Network recommendations on the evaluation and treatment of newly diagnosed patients with multiple myeloma. *Haematologica*. 2014;99(2):232–242.

515 # TABLES AND FIGURE LEGENDS

516

Table 1 BAGS defined subtype analysis.

Frequencies across data sets (P= 0.90).								
Group	**PreBI (%)**	**PreBII (%)**	**Im (%)**	**N (%)**	**M (%)**	**PC (%)**	**UC (%)**	**Sum**
UAMS	3 (1)	29 (5)	58 (10)	143 (26)	219 (39)	23 (4)	84 (15)	559
Hovon 65	2 (1)	19 (6)	45 (14)	61 (19)	134 (42)	11	48 (15)	320
Myeloma IX	1 (0)	14 (6)	23 (9)	59 (24)	105 (43)	8 (3)	37 (15)	247
IFM-DFCI	2 (1)	11 (6)	21 (12)	38 (22)	68 (40)	4 (2)	26 (15)	170
Sum*	8 (1)	75 (6)	147 (11)	302 (23)	528 (41)	46 (4)	196 (15)	1296

Association with the TC classification (P< 0.001).								
Group	**PreBI (%)**	**PreBII (%)**	**Im (%)**	**N (%)**	**M (%)**	**PC (%)**	**UC (%)**	**Sum**
4p16	1 (1)	21 (12)	7 (4)	45 (25)	78 (44)	3 (2)	22 (12)	177
MAF	2 (2)	4 (5)	12 (15)	25 (30)	31 (38)	1 (1)	7 (9)	82
6p21	1 (1)	2 (2)	12 (14)	19 (22)	36 (42)	2 (2)	14 (16)	86
11q13	0 (0)	3 (2)	11 (7)	46 (29)	81 (51)	4 (2)	15 (9)	160
D1	2 (0)	11 (3)	63 (15)	89 (21)	157 (37)	19 (5)	81 (19)	422

D1plusD2	0 (0)	4 (20)	2 (10)	2 (10)	9 (45)	0 (0)	3 (15)	20
D2	1 (1)	16 (13)	12 (10)	15 (12)	53 (44)	7 (6)	17 (14)	121
Unclassified	1 (0)	12 (5)	28 (12)	60 (26)	81 (36)	10 (4)	36 (16)	228
Sum	8 (1)	73 (6)	147 (11)	301 (23)	526 (41)	46 (4)	195 (15)	1296

ISS stage III with increased frequencies in the PreBII and M subtypes. (P= 0.032).

Group	**PreBI (%)**	**PreBII (%)**	**Im (%)**	**N (%)**	**M (%)**	**PC (%)**	**UC (%)**	**Sum**
Stage I	2 (0)	8 (2)	51 (12)	115 (28)	149 (36)	18 (4)	72 (17)	415
Stage II	1 (0)	18 (7)	25 (9)	65 (24)	105 (39)	15 (6)	39 (15)	268
Stage III	2 (1)	19 (10)	22 (11)	33 (17)	89 (46)	5 (3)	23 (12)	193
NA	1 (0)	17 (7)	28 (11)	50 (20)	115 (46)	4 (2)	35 (14)	250
Sum	6 (1)	62 (6)	126 (11)	263 (23)	458 (41)	42 (4)	169 (15)	1126

The BAGS-defined subtype analysis was performed across 4 different clinical cohorts (N = 1296 cases) following assignment of the data sets according to the restricted multinomial classifier.

520

Table 2 Cox proportional hazards regression analysis.

Combined PFS, n = 642, number of events = 311						
	Hazard ratio	**95% CI**	**P value**	**Hazard ratio**	**95% CI**	**P value**
Pre-BII	1	-	-	1	-	-
Immature	0.45	(0.26-0.76)	0.0032	0.48	(0.28-0.83)	0.0085
Naive	0.41	(0.25-0.66)	< 0.001	0.39	(0.24-0.63)	< 0.001
Memory	0.61	(0.39-0.95)	0.027	0.58	(0.37-0.92)	0.02
Plasmacell	0.28	(0.13-0.6)	< 0.001	0.3	(0.14-0.64)	0.0018
4p16	1	-	-	1	-	-
MAF	0.47	(0.29-0.77)	0.0026	0.53	(0.32-0.86)	0.011
6p21	0.23	(0.07-0.73)	0.013	0.17	(0.05-0.54)	0.0027
11q13	0.36	(0.26-0.52)	< 0.001	0.31	(0.22-0.45)	< 0.001
D1	0.35	(0.26-0.47)	< 0.001	0.3	(0.22-0.41)	< 0.001
D1plusD2	0.27	(0.12-0.63)	0.0024	0.27	(0.12-0.62)	0.0021
D2	0.38	(0.25-0.57)	< 0.001	0.25	(0.16-0.39)	< 0.001
Hovon65	1	-	-	1	-	-
MyelomaIX	1.16	(0.84-1.6)	0.38	1.15	(0.83-1.59)	0.41
UAMS	0.37	(0.29-0.48)	< 0.001	0.33	(0.26-0.42)	< 0.001
Combined OS, n= 642, number of events= 236						
	Hazard ratio	**95% CI**	**P value**	**Hazard ratio**	**95% CI**	**P value**
Pre-BII	1	-	-	1	-	-
Immature	0.31	(0.18-0.51)	< 0.001	0.36	(0.21-0.62)	< 0.001
Naive	0.19	(0.12-0.3)	< 0.001	0.19	(0.12-0.31)	< 0.001
Memory	0.33	(0.22-0.5)	< 0.001	0.34	(0.23-0.52)	< 0.001

Plasmacell	0.1	(0.03-0.28)	< 0.001	0.11	(0.04-0.32)	< 0.001
4p16	1	-	-	1	-	-
MAF	0.53	(0.3-0.93)	0.027	0.59	(0.34-1.05)	0.072
6p21	0.43	(0.13-1.36)	0.15	0.51	(0.16-1.65)	0.26
11q13	0.48	(0.33-0.71)	< 0.001	0.51	(0.35-0.76)	< 0.001
D1	0.37	(0.26-0.51)	< 0.001	0.39	(0.28-0.55)	< 0.001
D1plusD2	0.41	(0.17-1.02)	0.056	0.31	(0.12-0.78)	0.013
D2	0.53	(0.34-0.83)	0.0056	0.47	(0.3-0.75)	0.0014
Hovon65	1	-	-	1	-	-
MyelomaIX	1.23	(0.81-1.85)	0.33	1.24	(0.82-1.88)	0.31
UAMS	0.91	(0.68-1.21)	0.5	0.87	(0.65-1.17)	0.35

Cox proportional hazards regression analysis in the meta data set for BAGS subtypes based on PFS and OS, demonstrating added independent significance to the TC classification staging system. Columns on the left show results for a univariate analysis with each of the covariates, while columns on the right show results from the multivariate model. The Pre-BI class was dropped from the analysis due to, too few observations in this group.

FIGURE LEGENDS

Figure 1A-D Expression of membrane markers, transcription factors, and B-cell subset-specific genes in normal BM tissue.

A) B-cells of the BM were defined by flow cytometry as CD19+, CD45+, and CD3− and were additionally divided by surface marker expression of CD10, CD20, CD27, CD38, and CD34, published in detail previously[44].

The data quality of the differentiating B-cell subset compartments was validated as illustrated by normalized histograms of (**A**) the mean fluorescence intensities (MFIs) CD markers based on merged MFC reanalysis of pure sorted populations resulting from seven independent sorting procedures. Broken lines represents MFI values for each sorted B-cell subset.

B) Principal component analysis of the MFI values for each sorted cell in all samples. The cells are coded with a color according to their original subset. The dots represents mean values for each sorted B-cell subset.

C) The most variable probe sets were used in unsupervised hierarchical clustering analysis of the surface markers MME = CD10, CD34, CD38, CD27, PTPRC = CD45, MS4A = CD20 and CD19 used for FACS.

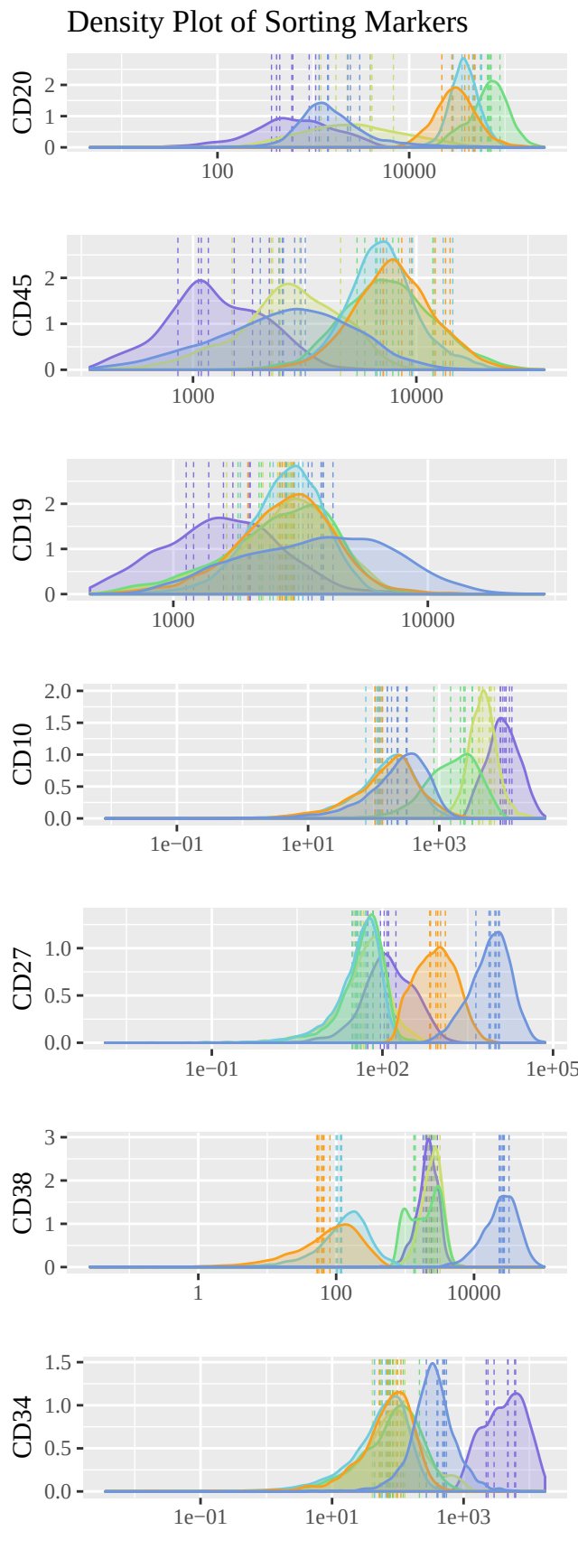
D) B-cell differentiation–specific genes (n = 45), summarized from a literature review of transcriptional regulation of B lymphopoiesis. The colors at the top of D indicate the relative gene expression for each sample, with blue representing high and brown representing low.

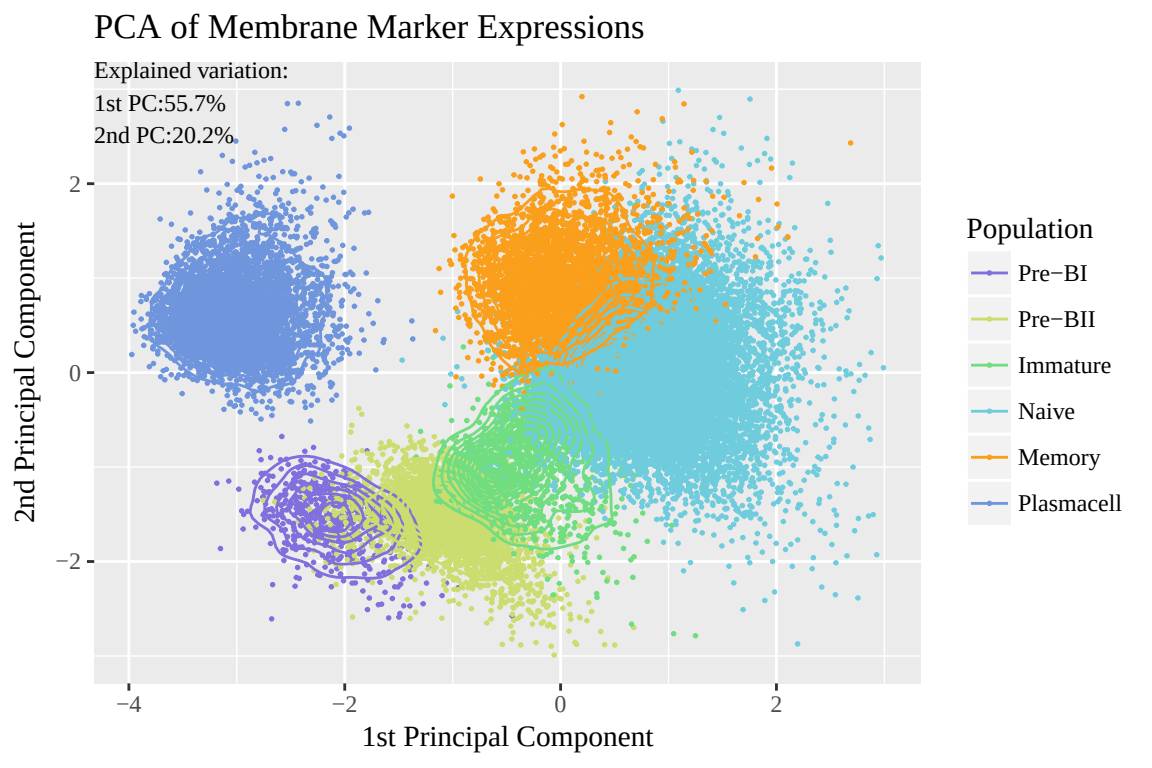
Figure 2A-B Meta-analysis of the prognostic impact of assigned BAGS subtypes.

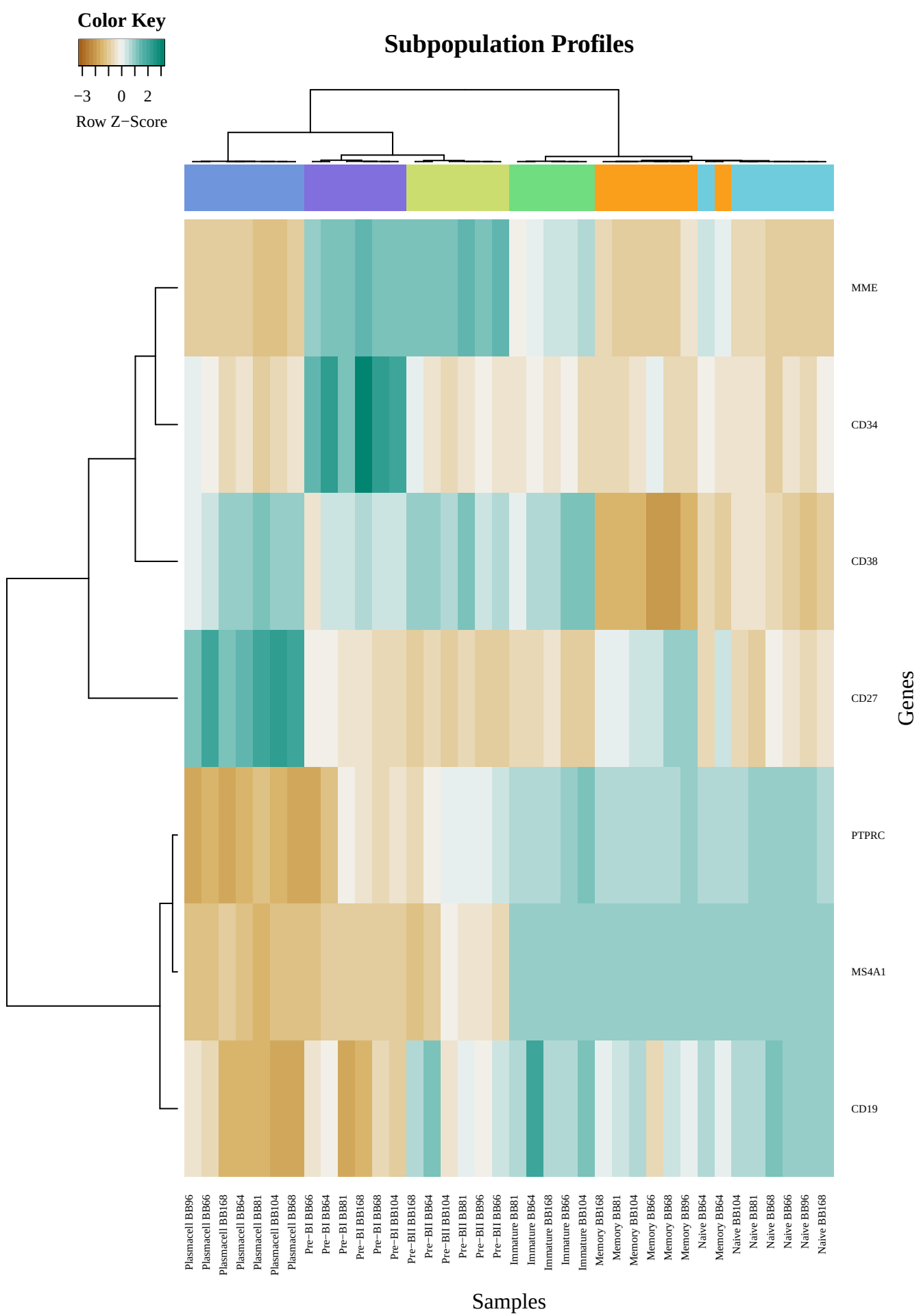
Progression-free **(A)** and overall survival **(B)** were compared between BAGS subtypes for high-dose melphalan–treated patients in published clinical trials. P-values are results from log-rank tests. The subtypes numbers given as n are the numbers of events/number of assigned patients with the subtypes in the meta data set. The BAGS subtypes are color coded as in **Figure 1A-D**.

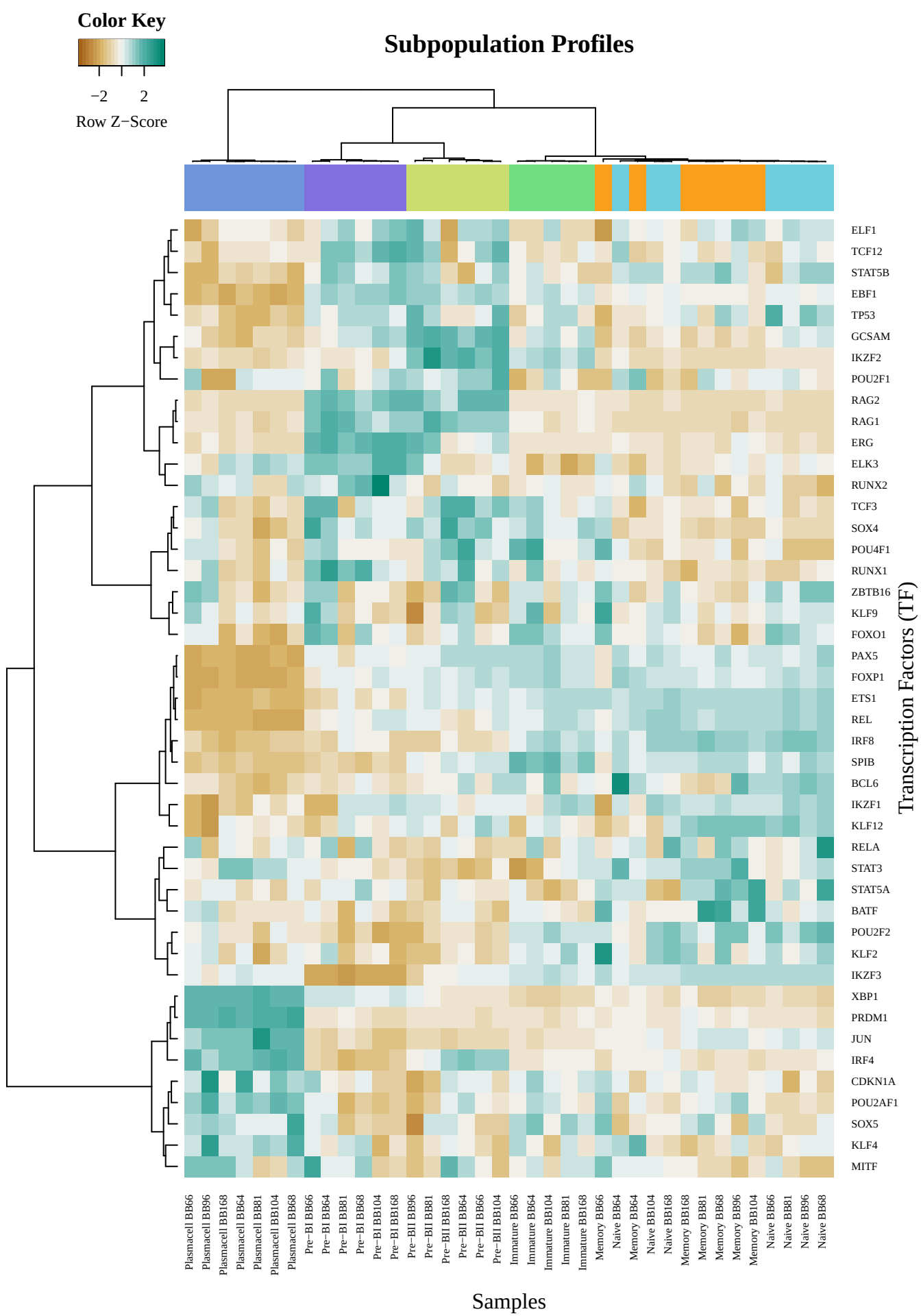
Figure 3A-C: BAGS subtype boxplots with correlation to proliferation, melphalan resistance, and beta-2 microglobulin.

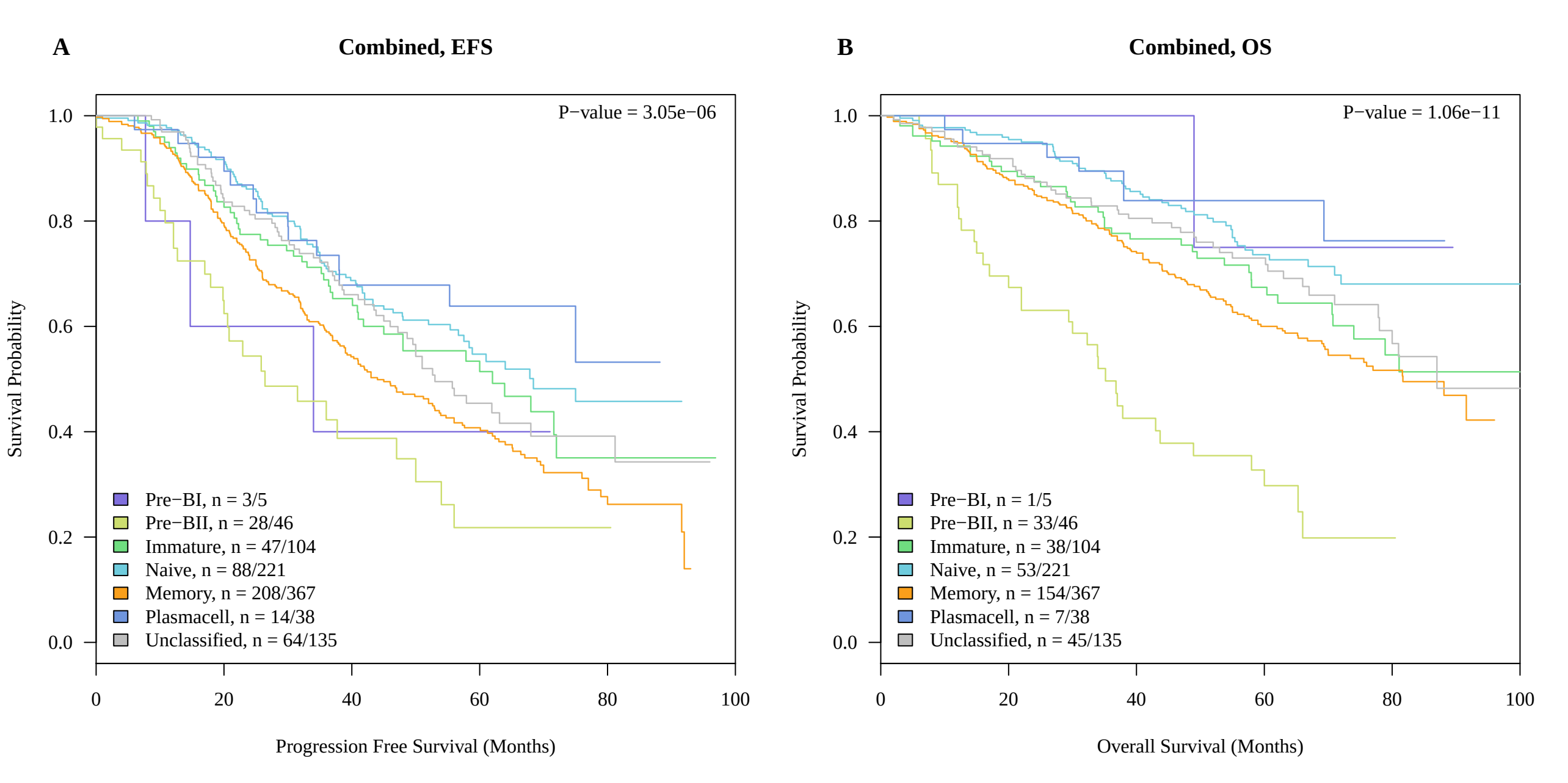
The individual adjusted proliferation index (PI) risk profiling **(A)**, melphalan drug resistance probability (index) **(B)**, and beta-2 microglobulin plasma level **(C)**, respectively per BAGS subtype cases from analysis of the meta dataset. The BAGS subtypes are color coded as in **Figure 1A-D**.

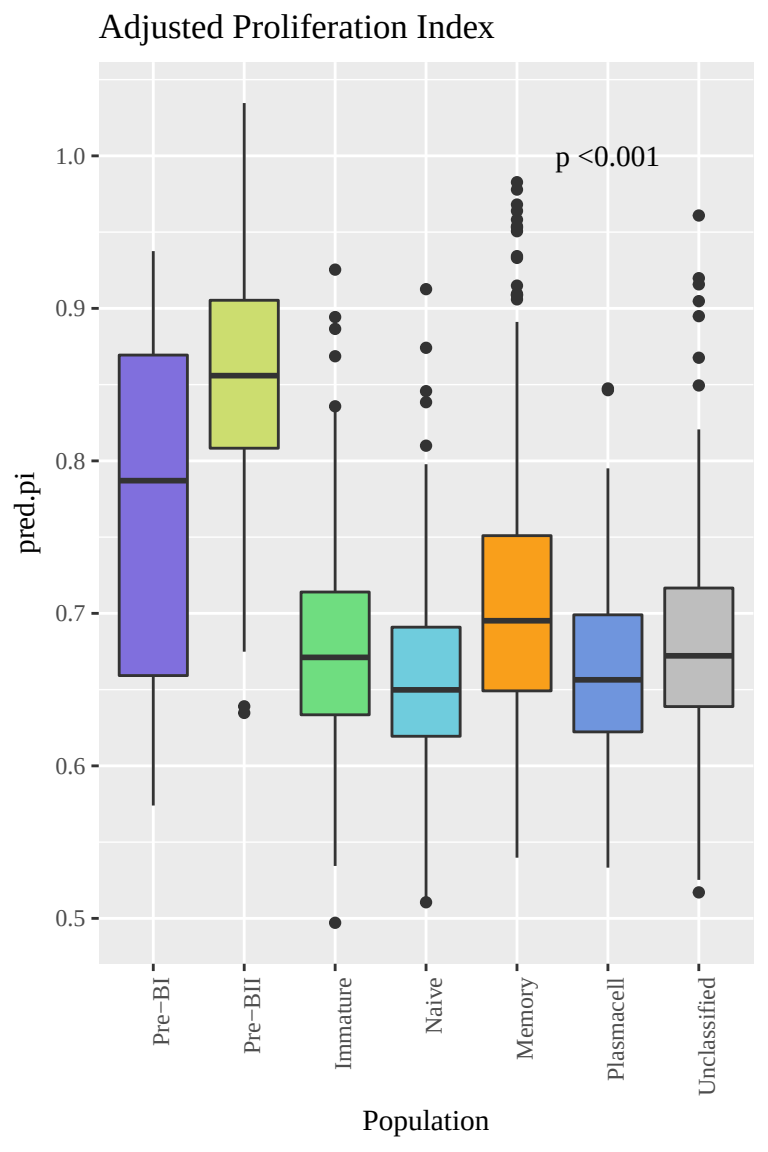

Density Plot of Sorting Markers
CD20
2
1
0
100
10000
CD45
2
1
0
1000
10000
CD19
2
1
0
1000
10000
CD10
2.0
1.5
1.0
0.5
0.0
1e−01
1e+01
1e+03
CD27
1.0
0.5
0.0
1e−01
1e+02
1e+05
CD38
3
2
1
0
1
100
10000
CD34
1.5
1.0
0.5
0.0
1e−01
1e+01
1e+03

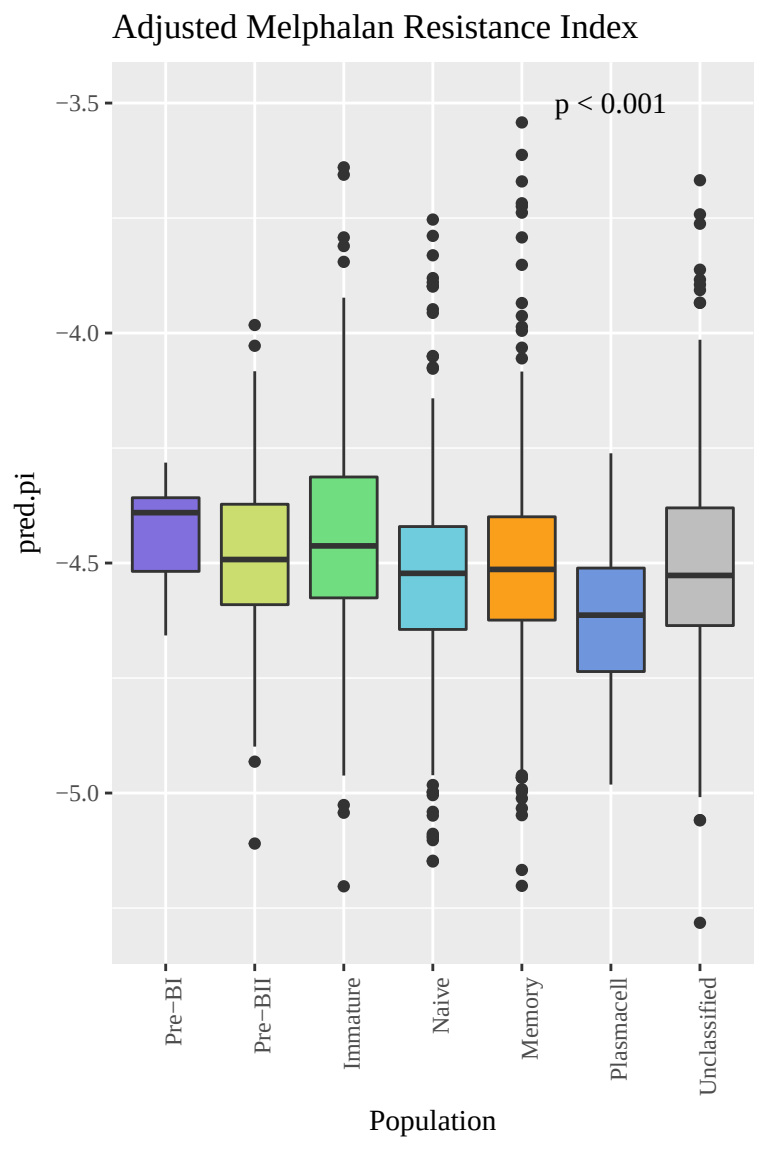

PCA of Membrane Marker Expressions
Explained variation:
1st PC:55.7%
2nd PC:20.2%
2nd Principal Component
1st Principal Component
2
0
−2
−4
−2
0
2
Population
Pre−BI
Pre−BII
Immature
Naive
Memory
Plasmacell

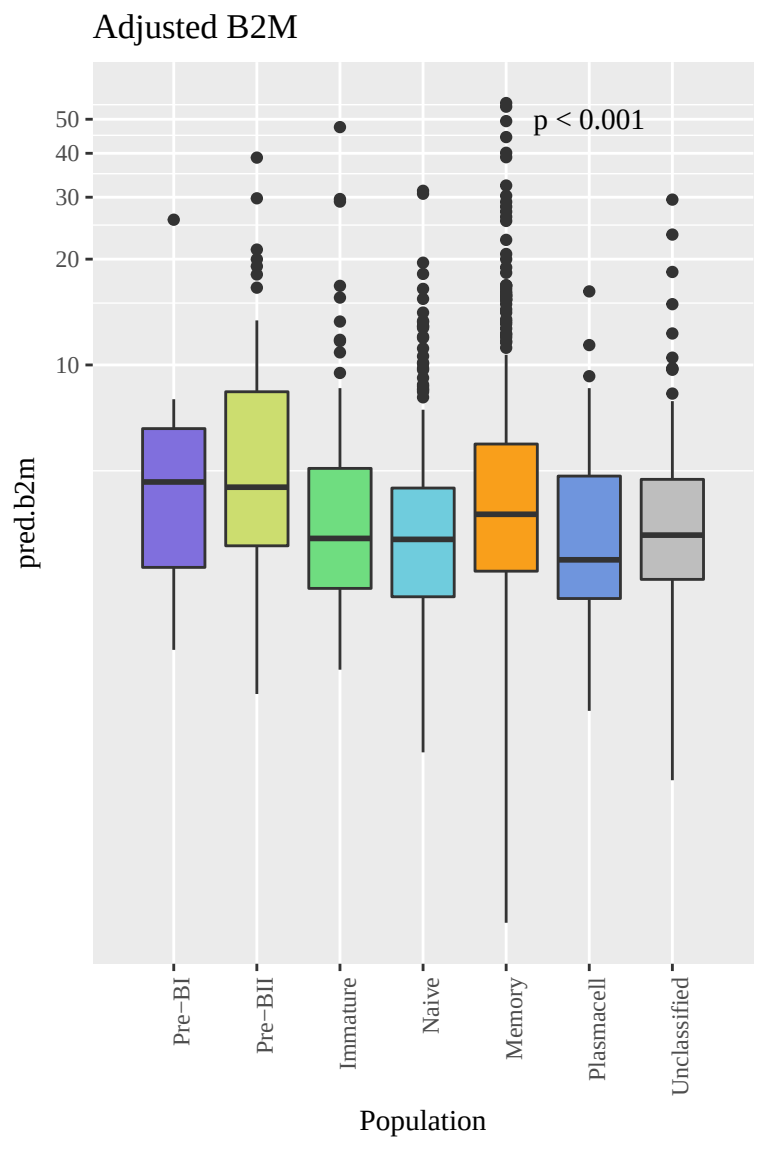

Color Key
−3
0
2
Row Z−Score
Subpopulation Profiles
MME
CD34
CD38
CD27
PTPRC
MS4A1
CD19
Genes
Plasmacell BB96
Plasmacell BB66
Plasmacell BB168
Plasmacell BB64
Plasmacell BB81
Plasmacell BB104
Plasmacell BB68
Pre−BI BB66
Pre−BI BB64
Pre−BI BB81
Pre−BI BB168
Pre−BI BB68
Pre−BI BB104
Pre−BII BB168
Pre−BII BB64
Pre−BII BB104
Pre−BII BB81
Pre−BII BB96
Pre−BII BB66
Immature BB81
Immature BB64
Immature BB168
Immature BB66
Immature BB104
Memory BB168
Memory BB81
Memory BB104
Memory BB66
Memory BB68
Memory BB96
Naive BB64
Memory BB64
Naive BB104
Naive BB81
Naive BB68
Naive BB66
Naive BB96
Naive BB168
Samples


Color Key
−2
2
Row Z−Score
Subpopulation Profiles
ELF1
TCF12
STAT5B
EBF1
TP53
GCSAM
IKZF2
POU2F1
RAG2
RAG1
ERG
ELK3
RUNX2
TCF3
SOX4
POU4F1
RUNX1
ZBTB16
KLF9
FOXO1
PAX5
FOXP1
ETS1
REL
IRF8
SPIB
BCL6
IKZF1
KLF12
RELA
STAT3
STAT5A
BATF
POU2F2
KLF2
IKZF3
XBP1
PRDM1
JUN
IRF4
CDKN1A
POU2AF1
SOX5
KLF4
MITF
Transcription Factors (TF)
Plasmacell BB66
Plasmacell BB96
Plasmacell BB168
Plasmacell BB64
Plasmacell BB81
Plasmacell BB104
Plasmacell BB68
Pre−BI BB66
Pre−BI BB64
Pre−BI BB81
Pre−BI BB68
Pre−BI BB104
Pre−BI BB168
Pre−BII BB96
Pre−BII BB81
Pre−BII BB168
Pre−BII BB64
Pre−BII BB66
Pre−BII BB104
Immature BB66
Immature BB64
Immature BB104
Immature BB81
Immature BB168
Memory BB66
Naive BB64
Memory BB64
Naive BB104
Naive BB168
Memory BB168
Memory BB81
Memory BB68
Memory BB96
Memory BB104
Naive BB66
Naive BB81
Naive BB96
Naive BB68
Samples


A
Combined, EFS
P−value = 3.05e−06
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Progression Free Survival (Months)
Pre−BI, n = 3/5
Pre−BII, n = 28/46
Immature, n = 47/104
Naive, n = 88/221
Memory, n = 208/367
Plasmacell, n = 14/38
Unclassified, n = 64/135
B
Combined, OS
P−value = 1.06e−11
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Overall Survival (Months)
Pre−BI, n = 1/5
Pre−BII, n = 33/46
Immature, n = 38/104
Naive, n = 53/221
Memory, n = 154/367
Plasmacell, n = 7/38
Unclassified, n = 45/135


Adjusted Proliferation Index
p <0.001
pred.pi
1.0
0.9
0.8
0.7
0.6
0.5
Pre−BI
Pre−BII
Immature
Naive
Memory
Plasmacell
Unclassified
Population


Adjusted Melphalan Resistance Index
p < 0.001
pred.pi
−3.5
−4.0
−4.5
−5.0
Pre−BI
Pre−BII
Immature
Naive
Memory
Plasmacell
Unclassified
Population


Adjusted B2M
p < 0.001
50
40
30
20
10
pred.b2m
Pre−BI
Pre−BII
Immature
Naive
Memory
Plasmacell
Unclassified
Population