Howard Hughes Medical Institute

Published as: *Cancer Immunol Res*. 2017 September ; 5(9): 744–754.

Clonal expansion and interrelatedness of distinct B-lineage compartments in multiple myeloma bone marrow

Leo Hansmann[1,5,6], **Arnold Han**[1], **Livius Penter**[5,6], **Michaela Liedtke**[2], and **Mark M. Davis**[1,3,4]

[1]Department of Microbiology and Immunology, Stanford University, Stanford, CA 94305

[2]Division of Hematology, Department of Medicine, Stanford University, Stanford, CA 94305

[3]Institute for Immunity, Transplantation, and Infection, Stanford University, Stanford, CA 94305

[4]The Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305

[5]Department of Hematology, Oncology, and Tumor Immunology, Charité – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany

[6]Berlin Institute of Health (BIH), Berlin, Germany

Abstract

Multiple myeloma is characterized by the clonal expansion of malignant plasma cells in the bone marrow. But the phenotypic diversity and the contribution of less predominant B-lineage clones to the biology of this disease have been controversial. Here we asked whether cells bearing the dominant multiple myeloma immunoglobulin rearrangement occupy phenotypic compartments other than that of plasma cells. To accomplish this, we combined 13-parameter FACS index sorting and t-Stochastic Neighbor Embedding (t-SNE) visualization with high-throughput single-cell immunoglobulin sequencing to track selected B-lineage clones across different stages of human B-cell development. As expected, the predominant clones preferentially mapped to aberrant plasma cell compartments, albeit phenotypically altered from wild type. Interestingly, up to 1.2 % of cells of the predominant clones co-localized with B-lineage cells of a normal phenotype. In addition, minor clones with distinct immunoglobulin sequences were detected in up to 9 % of sequenced cells but only 2 out of 12 of these clones showed aberrant immune phenotypes. The majority of these minor clones showed intraclonal silent nucleotide differences within the CDR3s and varying frequencies of somatic mutations in the immunoglobulin genes. Therefore the phenotypic range of multiple myeloma cells in the bone marrow is not confined to aberrant-phenotype plasma cells but extends to low frequencies of normal-phenotype B cells in line with the recently reported success of B cell-targeting cellular therapies in some patients. The majority of minor clones result from parallel non-malignant expansion.

Correspondence: Leo Hansmann, MD, Charité – Universitätsmedizin Berlin (CVK), Department of Hematology, Oncology, and Tumor Immunology, Augustenburger Platz 1, 13353 Berlin, Germany, leo.hansmann@charite.de, phone: +49-(0)30-450-665238, fax: +49-(0)30-450-553914; Mark M. Davis, Ph.D. Beckman Center, B221, Stanford University School of Medicine, 279 Campus Drive, Stanford, CA 94305 – 5323, USA, mmdavis@stanford.edu, phone: +1-650-725-4755, fax: +1-650-498-7771.

Conflict of interest statement: The authors declare no potential conflicts of interest.

Conflict of interest

L.H., A.H., L.P., M.L., M.M.D. declare no competing financial interests.

Keywords

Multiple Myeloma; human B-cell development; single-cell tracking; clonal relatedness; phenotypic diversity

Introduction

Multiple myeloma is a malignancy of the B lineage characterized by the accumulation of clonal plasma cells in the bone marrow. Most, if not all, cases develop from a pre-cancer called monoclonal gammopathy of undetermined significance (MGUS) (1,2). Due to the presence of identical somatic mutations of the immunoglobulin genes and completed class-switch recombination, the disease is assumed to originate at the germinal center or post-germinal center stage of B-cell development (3–5). Current thinking holds that the origin of this malignancy includes alterations in the bone marrow microenvironment (6,7), vasculogenesis (8), the T-, B-, NK-, and dendritic cell compartments (9–11), as well as genetic aberrations within the malignant cells themselves (12–16), resulting in variable degrees of intraclonal heterogeneity.

Intraclonal heterogeneity usually refers to accumulating genetic and epigenetic aberrations that drive the progression from MGUS to symptomatic multiple myeloma (17,18). Clonal heterogeneity has also been detected at the immune phenotype level and certain phenotypes have been shown to be partly associated with genetic aberrations, chemotherapy-resistance, and proliferation advantages (16). Nevertheless, identification of phenotypic diversity has been challenging. Although flow cytometry can detect high-dimensional immune phenotypes on the single-cell level, clonality has usually been determined by immunofixation or immunoglobulin sequencing in bulk populations resulting in the loss of single-cell resolution and accuracy.

Neither RNA- nor DNA-based bulk immunoglobulin gene sequencing can quantify clone frequencies or assigned immune phenotypes to the identified clones could be inaccurate as a result of contamination with other cells (19). Yet, the success of B cell-targeting cellular therapeutics in subsets of patients (20,21) suggests the involvement of cells phenotypically different from the majority of aberrant-phenotype plasma cells in multiple myeloma pathophysiology.

Knowledge of clonal diversity and phenotype is critical to understanding the pathogenesis of multiple myeloma and may aid identification of targets for therapeutic intervention. We combined phenotypic tracking of B-lineage clones using single-cell multiparameter FACS index sorting with single-cell next-generation sequencing (NGS) of immunoglobulin genes in order to assess the phenotypic diversity of any B-lineage clone of choice in multiple myeloma. Index sorting allows the isolation of hundreds to thousands of single cells without up-front selection of specific immune phenotypes, yet fluorescence intensities for all markers included in the panel are recorded and can be read out for each individual cell retrospectively. We developed a methodology that combines DNA-barcoding and high-throughput NGS of immunoglobulin light chain genes from hundreds of single B-lineage

cells with 13-parameter FACS index sorting to phenotypically track individual clones on the single-cell level in human bone marrow.

We find that the immune phenotypes of these molecularly defined, predominant multiple myeloma clones range from aberrant-phenotype multiple myeloma cells to normal-phenotype B-lineage cells. The parallel expansion of a variety of less predominant B-lineage clones follows the same pattern.

The majority of the less predominant B-lineage clones phenotypically clustered with normal-phenotype B cells and showed low somatic mutation rates in the light chain genes and silent nucleotide differences within the CDR3, suggesting that their expansion is driven by antigen-dependent physiologic rather than malignant processes.

Materials and methods

Subjects and bone marrow cell preparation

Fresh heparin-anticoagulated bone marrow from three multiple myeloma patients (MM1-3, Table 1, Supplementary Table S1) and peripheral blood from one healthy volunteer were acquired after they gave written informed consent in accordance with federal and local human subjects regulations (Institutional Review Board ID 25310 to LH). Mononuclear cells were isolated using Ficoll-Paque PLUS (GE Healthcare) centrifugation, resuspended in cell culture medium (Supplementary Table S2) containing 10 % dimethyl sulfoxide and 50 % fetal bovine serum (both Sigma Aldrich), and cryopreserved.

FACS staining and sorting

Cells were thawed and stained for FACS analysis/sorting (Supplementary Table S3) according to the manufacturer's instructions. For the detailed gating strategy, see Supplementary Fig. S1. Plates were pre-loaded with 1.4×first strand buffer for Superscript III RT (Invitrogen) and chilled during single-cell index sorting on a FACS Aria II or Aria Fusion cell sorter (BD). Plates were immediately transferred to dry ice after sorting.

Single-cell Reverse Transcription, light chain amplification, barcoding, and sequencing

Superscript III (Invitrogen) reverse transcription (RT) was done on index-sorted single cells in a final volume of 14 µl/reaction. Octylphenoxypolyethoxyethanol (IGEPAL CA-630, Sigma) was added at 1 % final concentration and RT primers (Suppl. Tab. S4) were added at 257 nM final concentration. Plates were incubated at 65 °C for 3 min., 25 °C for 3 min., and placed on ice to add 7.9 mM DTT (Invitrogen), 429 µM dNTPs (Thermo Scientific) (all final concentrations), and 0.3 µl Superscript III RT. RT was done at 37 °C for 1 h and 70 °C for 15 min. First and second amplification reactions were done using Phusion High-Fidelity DNA Polymerase (NEB). Final PCR volumes were 15 µl using 4 µl cDNA or 1 µl first amplification products as templates. dNTPs were used at 433 µM and DMSO was added at 5 % final concentration to the second amplification PCR. First amplification and second amplification (barcoding) primers were added to a final concentration of 333 nM each. Cycling conditions: first amplification: 98 °C-60 sec., 3×(98 °C-45 sec., 45 °C-45 sec., 72 °C-105 sec), 30×(98 °C-30 sec., 50 °C-45 sec., 72 °C-90 sec.), 72 °C-7 min; second

amplification: 98 °C-30 sec., 30×(98 °C-30 sec., 50 °C-45 sec., 72 °C-45 sec.), 72 °C-5 min. Primers for RT, first, and second amplification were adapted from Wang et al. (22) and modified for molecular barcoding and Illumina sequencing. The barcoding primers add a nucleotide sequence specific for plate, column, and row to each single amplified transcript (Suppl. Tab. S4). Barcodes were designed with differences in at least two nucleotides when compared to all other barcodes in the panel.

Third amplification reactions were done in 15 μl final volume using a multiplex PCR kit (Qiagen) and 1ul of second amplification product as template. Paired-end primers adding the remaining sequences of the Illumina adapters were used at 400 nM final concentration and cycling conditions were: 95 °C-15 min., 28×(95 °C-30 sec., 60 °C-45 sec., 72 °C-90 sec.), 72 °C-10 min. After the third amplification, PCR products from all wells were combined, gel purified and sequenced on an Illumina MiSeq (Illumina) using Illumina V3 kits (Illumina).

Sequencing analysis

Illumina sequencing reads were assembled using PEAR (23) and a cut-off of at least 200 nucleotides for assembled reads. Python scripts identified barcodes in every single read (only perfect matches were accepted), filtered sequences by aligning against germline-encoded V genes using bowtie2 (24), built consensus sequences with MUSCLE (25) to upload them to IMGT/HighV-QUEST (26) for V, J gene assignment, detection of somatic mutations, and CDR3 identification. Read number cut-offs were determined based on background reads in empty wells in analogy to our methodology for T-cell receptor sequencing (27) (Supplementary Fig. S2). Sequence alignments were done running MUSCLE out of MEGA (28). Nucleotide sequences of individual cells of the predominant clones of all three patients have been made publicly available in GenBank, accession number: KY177746-KY178269.

Sequence logos were created with WebLogo 3 (29).

Data visualization

For t-SNE (30) analyses, compensated FACS data and index sort information were exported from flowjo (treestar) and DiVA (BD) software and combined using R (31). The Barnes-Hut implementation of t-SNE was run using the arguments "–v –d 2 –p 10" on combined FACS data exported from gates shown in Suppl. Fig. S1D and index sort data. t-SNE visualization was done using the ggplot2 package in R. For the visualization of individual marker expressions, fluorescence values were trimmed at the 1st and 99th percentiles, scaled, log10-transformed, and color-coded with yellow representing the lowest and red representing the highest expression levels.

For two-dimensional FACS plots of single clonal cells, FACS data from the whole sort dataset (contours in gray) were overlaid with fluorescence data of single clonal cells (filled black circles) of the indicated clones in R. In two-dimensional FACS plots of individual clones, negative fluorescence values were set to 0 for visualization purposes.

Results

Multidimensional visualization of phenotypic diversity in multiple myeloma plasma cells

In multiple myeloma, flow cytometry can identify malignant cells through their aberrant antigen expression profiles (16,32,33). Although a variety of markers characterize this disease, variation is considerable between and within individual patients. To better capture this variation, we used a nonlinear principal component analysis algorithm, t-SNE (30) for the visualization of multidimensional (13 markers and forward-/side-scatter characteristics, Supplementary Table S3) flow cytometry data of multiple myeloma bone marrow cells from three heterogeneous patients (MM1-3, Fig. 1, Table 1). Single data points in the t-SNE maps represent single cells. Coordinates were calculated taking into account all pairwise distances in the multidimensional datasets. Malignant plasma cells in MM were $CD38^{hi}CD27^{+}$ with aberrant co-expression of CD56, but lacked CD45 and CD20 expression. Plasma cells in MM2 were $CD38^{hi}$ but did not co-express CD27, CD56, CD45, or CD20. The aberrantly differentiated cells in MM3 were $CD38^{hi}CD27^{lo}CD56^{lo}$. In each of the three multiple myeloma cases the plasma cell populations occupied multiple compartments in the t-SNE maps, underlining their phenotypic heterogeneity even within individual patients (Fig. 1). MM3 received treatment until two months before the bone marrow aspiration done for this study.

Current flow cytometry approaches to multiple myeloma can identify diversity in the malignant cell populations through their aberrant marker profiles, even in the absence of information about the individual cells' immunoglobulin sequences. However, these approaches do not allow visualization of the phenotypic distribution of single cells belonging to a particular, molecularly defined B-lineage clone.

Phenotypic tracking of individual B-lineage clones

Phenotypes and clonal relatedness of individual cells can be identified at the single-cell level. Therefore we combined single-cell FACS index sorting with DNA-barcoding and high-throughput single-cell immunoglobulin light chain sequencing to determine the clonal relatedness of individual B-lineage cells and to define their position on high-dimensional t-SNE phenotypic maps (Fig. 2). We applied a 13-marker FACS panel to bone marrow cells from three multiple myeloma patients and index sorted single B-lineage cells identified by CD19 and CD38 expression characteristics (Fig. 2A left panel) into multiwell plates. Index sorting allowed assignment of fluorescence intensities for each fluorchrome in the FACS panel and forward/side scatter characteristics to every cell sorted independently from the sorting gates. In addition to other B-lineage cells, we sorted phenotypically defined multiple myeloma plasma cells (Supplementary Fig. S1) to identify the multiple myeloma immunoglobulin rearrangement. Light chain RNA of each single sorted cell was reverse transcribed, amplified, barcoded with plate- and well-specific DNA barcodes, and sequenced using the Illumina MiSeq® system. For the PCR amplifications, we used a nested PCR approach (Fig. 2B) resulting in the amplification of κ and λ light chain genes spanning from the framework region 1 to the constant region. Cells showing identical VJ junction amino acid sequences were considered clonal. Clonality was confirmed by comparing amplified nucleotide sequences (frame work region 1 to constant region, approximately 344

nucleotides), including somatic mutations. In parallel, t-SNE maps were generated from the corresponding FACS data and index sort technology allowed the precise mapping of each cell (Fig. 2A, right panel, clonal cells of the predominant clone were colored in black). Our sequencing approach led to the identification of light chain sequences with a median efficiency of 71 % of the sorted cells (range: 32 % – 84 %). Efficiency varied with the stringency of the applied sorting gates. We sought to include all possible B-lineage cells in the sorting gate at the expense of overall sequencing efficiency by allowing selection of a variable percentage of non-B cells (Fig. 2A, left panel, and Supplementary Fig. S1).

To obtain an experimental estimate of the accuracy of our index sorting, we randomly sorted single T and B cells into 96-well plates and used our sequencing technology to identify B cells that were wrongly assigned a T-cell phenotype. A wrong assignment did not occur in even a single instance and therefore the probability for an index sort mistake was below the experimental detection limit of 0.0045 (Supplementary Fig. S3).

Clonal multiple myeloma cells are not only present in plasma cell compartments

Hundreds of single B-lineage cells from the same multiple myeloma patients whose phenotypes were presented in Fig. 1 were index sorted and sequenced as outlined in Fig. 2. Cells of the predominant clones were overlaid in black filled circles over the corresponding t-SNE plots (Fig. 3A). The monoclonal cells mapped to plasma cell compartments but with aberrant marker expression patterns. In MM1, most cells of the predominant B-lineage clone co-localized with $CD38^{hi}CD27^{+}CD56^{+}$ cells. In MM2, the cells of the predominant clone were $CD38^{hi}CD45^{-}$ and in MM3 cells of the predominant clone mapped to the $CD38^{hi}CD27^{lo}CD56^{lo}$ compartments (Fig. 3A).

In addition to aberrant-phenotype plasma cells, we detected cells that had the same characteristic immunoglobulin nucleotide sequence (frame work region 1 to constant region, including somatic mutations, Supplementary Fig. S4) but mapped to other phenotype compartments. For example, single cells from MM1, when plotted in two dimensions with a selected set of markers (Fig. 3B), showed expression of CD19, CD20, and CD45 alone or in combination without co-expression of CD56. A few single cells of the predominant clones showed phenotypic characteristics similar to normal B rather than multiple myeloma cells in up to 1.2 % of all sequenced B-lineage cells (up to 6 cells per patient; for detailed FACS plots of the predominant clones of all multiple myeloma patients see Supplementary Fig. S5). Each single B cell of normal phenotype that belonged to the multiple myeloma clonotype yielded on average 1137 sequencing reads (range: 89–3228 reads) which was > 8-fold more than background (Supplementary Fig. S2). Although multiple myeloma cells have undergone class switch recombination and lack CD20 as well as surface immunoglobulin expression, one B-lineage cell in MM2 shared the characteristic light chain rearrangement and was $CD19^{+}CD20^{+}CD27^{-}$ in combination with weak surface IgD expression, presenting an immune phenotype similar to naïve B cells (Supplementary Fig. S5).

To confirm the identity of the light chain sequences representing the multiple myeloma clones (predominant clones in Table 2), we sequenced the amplified fragments (framework region 1 to constant region, approx. 344 nucleotides). The sequences, including those nucleotides deriving from somatic mutation (Supplementary Fig. S4), were identical. This

suggests that these multiple myeloma clones did not arise in or re-enter germinal centers, where variation due to somatic mutation, might be expected.

Phenotype of rare B-lineage clones from multiple myeloma bone marrow

Less predominant clones accompanying the predominant multiple myeloma clones have been identified by immunofixation and on the phenotypic (FACS) or molecular (bulk sequencing) levels (16, 34–36). However, sequencing analyses in bulk are prone to PCR bias and the phenotypic relationships of less predominant clones to the multiple myeloma clones cannot be determined by bulk sequencing. Our methodology allows for quantification of less predominant clones in combination with their immune phenotypes on the single-cell level.

In MM1, we detected a total of 45 expanded B-lineage clones (their CDR3 sequences occurred in two or more sorted cells), 38 in MM2, and 21 in MM3 (Supplementary Table S5). CDR3 sequences and frequencies of the 5 most predominant clones are summarized in Table 2. The majority of the less predominant clones (clones 2–5 in MM1 and MM2, clones 4–5 in MM3) did not map to multiple myeloma plasma cell compartments and showed phenotypic characteristics of normal B cells ($CD45^{+}CD19^{+}CD20^{+}CD27^{+/-}CD56^{-}CD138^{-}$, surface $IgD^{+/-}$) as shown for the second predominant clone in MM2 as an example in Fig. 4A. The complete phenotypes of less predominant clones are shown in Supplementary Fig. S6. However, the second and third predominant clones in MM3 showed mixed characteristics of B cells ($CD45^{+}$, partially $CD19^{+}CD20^{+}$) and plasma cells ($CD38^{hi}$, surface $IgD^{-}IgG^{-}$, Fig. 4B, Supplementary Fig. S6). These clones mapped to the same phenotypic compartments that were occupied with the corresponding multiple myeloma cells (see Fig. 3A) and the partial co-expression of CD56 characterized these cells as aberrantly differentiated B-lineage cells (Fig. 4B, detailed phenotypes in Supplementary Fig. S6).

The second most predominant clones in MM2 and MM3 showed identical CDR3 sequences (Table 2). Due to less diversity in immunoglobulin light chain sequences when compared to heavy chains, overlapping CDR3 sequences between individuals is not surprising. In fact, this particular nucleotide sequence of MM2 and MM3 has been previously identified and deposited in the NCBI GenBank database (accession number EU789132). The expansion of this particular light chain in the context of multiple myeloma has to be confirmed in larger cohorts and the underlying cues will be identified in future studies.

Clones expanded by antigen-driven convergent expansion

The finding of less predominant B-lineage clones in multiple myeloma bone marrow raises the question of whether these clones are the result of a parallel expansion of malignant B-lineage cells or part of a normal antigen-directed response. The light chain gene sequences of all cells of the most predominant clones showed nucleotide sequence identity across the entire VJ gene segment including somatic mutations (see Fig. 5A for numbers of somatic mutations and Supplementary Fig. S4 for sequence logos), indicating their malignant clonal relationship. In contrast, 8 out of the 12 less predominant clones listed in Table 2 showed unique CDR3 sequences and a nucleotide diversity within the CDR3 of individual clonal cells that did not affect amino acid sequences (example in Fig. 5B). Similar patterns could also be observed in peripheral blood B cells from healthy individuals (Suppl. Fig. S7)

suggesting that less predominant clones in multiple myeloma bone marrow could arise as part of a normal, antigen-driven expansion (Supplementary Fig. S4 for sequence logos of the five most predominant B-lineage clones from all multiple myeloma samples). Less predominant clones harbored varying (and in 10 out of 12 cases significantly lower, $P < 0.05$, Fig. 5A) amounts of somatic mutations in their light chain sequences when compared to the predominant clones (Fig. 5A). In line with a lower number of somatic mutations, 7 out of 12 less predominant clones showed surface IgD expression and were $CD45^{+}CD20^{+}$, underlining their phenotypic and molecular difference from the predominant clones (shown for clones 1–3 of MM2 as an example in Fig. 5C).

Taken together, the expansion of the most predominant multiple myeloma clones, despite their phenotypic diversity, is part of the malignant monoclonal expansion and shows its phenotypic range. The minor clones in most cases do not show plasma cell phenotypes and seem characteristic of a normal, antigen-driven process.

Discussion

Estimation of B cell clonal frequencies and identification of clonal phenotypes in multiple myeloma require the efficient and reliable combination of single-cell technologies.

The application of single-cell methods is especially useful here, as it overcomes the bulk sequencing bias due to the variable number of immunoglobulin gene transcripts per cell. Especially when analyzing bone marrow cells of the entire B lineage, where plasma cells can contain 10–300 times more immunoglobulin RNA than mature B cells (37), bulk sequencing approaches using immunoglobulin mRNA as a template are apt to be especially biased. DNA-based approaches are less affected by varying template copy numbers per cell but are still subject to PCR amplification bias and in general achieve lower efficiencies.

As sequencing efficiency is important for our methodology, we focused on immunoglobulin light chain sequencing, which yields higher efficiencies when compared to heavy chain sequencing. Despite less junctional diversity in light chain than in heavy chain immunoglobulin genes, the substantial amount of somatic mutations in multiple myeloma cells (at average 24 somatic mutations in the most predominant clones in our dataset) allow us to detect clonality (38). Unproductive heavy chain rearrangements can occur in approx. 15 % of multiple myeloma patients (39,40).

The combination of sequencing technology (a median efficiency of 71 %) with multicolor (13 parameters) single-cell FACS index-sorting allowed the high-dimensional phenotypic tracking of even modestly expanded B-lineage clones. To minimize PCR and sequencing errors, we used a high fidelity PCR enzyme (™Phusion High-Fidelity DNA Polymerase (NEB)) with an error rate > 50-fold lower than the typical Taq polymerases (41) and paired-end sequencing (27) that yielded hundreds to thousands of reads per cell allowing for correction of sequencing errors. With respect to sorting accuracy, the sorter identifies whether and to what extent markers are expressed on individual cells. However, fluorescence detection is subject to variance, which adds difficulty to assigning a cell to a population when that cell has characteristics that place it between two populations.

Our methodology integrates high-dimensional (13 fluorescence parameters plus forward-/side-scatter characteristics) phenotypic datasets into maps of multilevel intra- and interclonal diversity, which will support identification of disease-related phenotypes in autoimmune disorders, inflammation, infections, and malignancies beyond multiple myeloma.

Applying this methodology to B-lineage cells in multiple myeloma bone marrow, we also found the dominant multiple myeloma clones in up to 1.2 % of sequenced cells that were phenotypically distinct from the majority. The existence of B-lineage cells other than plasma cells with the major clonotype and their physiological relevance has been discussed (19,42–49). However, the phenotypic range of clonotypic cells could not be studied at the single-cell level in high-dimensional space. We show in multiple myeloma bone marrow that the phenotypes of clonotypic B-lineage cells can be heterogeneous, varying from differences in only 1–2 markers (e.g. expression of CD45 and absence of CD56 where the predominant clone is $CD45^{-}CD56^{+}$) to those similar to mature B cells ($CD19^{+}CD20^{+}CD45^{+}$). Phenotypes of clonal cells were also heterogeneous, which would have given earlier FACS-sorting approaches with upfront-defined gating strategies limited chance of success. Given the diversity within the plasma cell compartment of our study samples and data in the literature (20,21,43,44,46,50–54), low frequencies of different-phenotype B-lineage cells sharing the multiple myeloma rearrangement are not surprising. We would expect the variety of these phenotypes to increase with sample size. Whether these cells are multiple myeloma precursors, stem cells, or result from the phenotypic instability of multiple myeloma cannot be concluded from our data. However, the therapeutic success of anti-CD3 × anti-CD20 bispecific antibody-armed T cells (21) and chimeric antigen receptor T cells against CD19 in some patients with apparently $CD19^{-}$ multiple myeloma (20) suggests that in selected cases, cells that are phenotypically B cells but harbor a plasma cell clonotype may be clinically relevant therapeutic targets. Our methodology could help to identify patients that might benefit from B-cell ablation.

Multiple myeloma patient 3 received treatment until two months before the bone marrow aspiration done for this study. Whether the lower phenotypic diversity of the cells belonging to the predominant multiple myeloma clone in this patient was due to the previous treatment or reflects the phenotypic distribution of the disease cannot be concluded from our data.

Future studies will determine whether clonotypic cells with B-cell immune phenotypes are more characteristic of mature B cells or of malignant plasma cells (13,15,16,55). Our approach to identify these rare clonotypic cells requires sequencing, which in destroying the cells renders them inaccessible for downstream functional analyses.

Our approach also allowed us to discover other B-lineage clones that, although less abundant than the dominant species, were nonetheless amplified (between 1 % and 9 % of sequenced cells). Multiple B-lineage clones are evident in 2 % – 6 % of patients (34,35) based on protein electrophoresis or immunofixation. These techniques have poor resolution, high detection limits (150–500 mg/l monoclonal protein) (56–58), do not reach single-cell resolution, and require the monoclonal protein being present in the serum. With single-cell analysis, we found minor B-lineage clone expansion in all three of the patients analyzed. These minor clones showed phenotypes associated with mature B rather than plasma cell

differentiation and seem to resemble physiologic processes as they also occur in healthy individuals.

Two patients shared the same CDR3 sequence as has been reported before in a patient with activation-induced cytidine deaminase deficiency (GenBank accession number EU789132) (59). This and the fact that the predominant light chain CDR3 sequences from MM2 and MM3 have already been published in patients with multiple myeloma and cast nephropathy (60) suggests that there may be common antigen specificities in at least some patients with this disease.

The analysis of V gene mutation frequencies revealed varying and lower numbers of somatic mutations in cells belonging to 10 out of 12 of the less predominant clones. When comparing CDR3 nucleotide sequences within the same clones, we detected silent nucleotide exchanges, suggesting a convergent antigen-driven expansion of these clones as it also occurs independent form multiple myeloma in peripheral blood B cells from healthy individuals. Our methodology can give insights whether parallel expanding minor clones in a patient are likely to be part of the malignancy (aberrant phenotypes, nucleotide sequence identity including somatic mutations) and might be of potential therapeutic/prognostic relevance or follow disease-unrelated processes (normal B/plasma cell phenotypes, silent nucleotide exchanges).

In conclusion, we present a method that combines the accuracy of multiparameter FACS single-cell index-sorting with high throughput, high efficiency next generation sequencing to link molecular information to immune phenotypes on the single-cell level. The method, applicable to the field of B-cell immunology, allows one to analyze the immune phenotypes that a given B- or plasma-cell clone can occupy, with the immunoglobulin sequence acting as a bar code to identify clonal progeny. We used this methodology for phenotypic tracking of B-lineage clones in multiple myeloma bone marrow and showed the expansion of molecularly defined predominant multiple myeloma clones to include a) aberrant phenotype multiple myeloma cells and b) normal phenotype B-lineage cells. The parallel expansion of a variety of less predominant B-lineage clones includes c) aberrant phenotype, and d) normal phenotype B-lineage cells and in the majority of clones is the result of convergent B-lymphocyte expansion rather than malignant proliferation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Francesco Vallania and Christopher Bolen for helpful discussions and support with bioinformatics, Hans-Peter Rahn at the Preparative Flowcytometry Facility at the Max Delbrück Center for Molecular Medicine for flow cytometry support, Kerstin Dietze for excellent technical assistance, and Alessandra Aquilanti for critically reading the manuscript. We thank Irene Panzer for the sharing of FACS reagents as well as Bernd Dörken and Jörg Westermann for their support. Single-cell sorting was performed in the Stanford Shared FACS Facility using NIH S10 Shared Instrument Grant (S10RR025518-01). Leo Hansmann was supported by a research fellowship by the German Research Foundation (DFG, HA 6772/1-1) and is a Charité – Berlin Institute of Health (BIH) Clinician Scientist. Livius Penter is a BIH Junior Clinician Scientist. Major support for this work also came from the National Institutes of Health (U19 AI 057229 to M.M.D.).

Financial support: Leo Hansmann was supported by a research fellowship by the German Research Foundation (DFG, HA 6772/1-1) and is a Charité – Berlin Institute of Health (BIH) Clinician Scientist. Livius Penter is a BIH Junior Clinician Scientist. Major support for this work also came from the National Institutes of Health (U19 AI 057229 to M.M. Davis) and the Parker Institute for Cancer Immunotherapy.

References

1. Landgren O, Kyle RA, Pfeiffer RM, Katzmann JA, Caporaso NE, Hayes RB, et al. Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. Blood. 2009; 113(22):5412–7. DOI: 10.1182/blood-2008-12-194241 [PubMed: 19179464]
2. Weiss BM, Abadie J, Verma P, Howard RS, Kuehl WM. A monoclonal gammopathy precedes multiple myeloma in most patients. Blood. 2009; 113(22):5418–22. DOI: 10.1182/blood-2008-12-195008 [PubMed: 19234139]
3. 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. DOI: 10.1038/leu.2011.70 [PubMed: 21494259]
4. Kosmas C, Stamatopoulos K, Stavroyianni N, Zoi K, Belessi C, Viniou N, et al. Origin and diversification of the clonogenic cell in multiple myeloma: lessons from the immunoglobulin repertoire. Leukemia. 2000; 14(10):1718–26. [PubMed: 11021746]
5. Bakkus MH, Heirman C, Van Riet I, Van Camp B, Thielemans K. Evidence that multiple myeloma Ig heavy chain VDJ genes contain somatic mutations but show no intraclonal variation. Blood. 1992; 80(9):2326–35. [PubMed: 1421403]
6. Mitsiades CS, Mitsiades N, Munshi NC, Anderson KC. Focus on multiple myeloma. Cancer cell. 2004; 6(5):439–44. DOI: 10.1016/j.ccr.2004.10.020 [PubMed: 15542427]
7. Mitsiades CS, Mitsiades NS, Richardson PG, Munshi NC, Anderson KC. Multiple myeloma: a prototypic disease model for the characterization and therapeutic targeting of interactions between tumor cells and their local microenvironment. Journal of cellular biochemistry. 2007; 101(4):950–68. DOI: 10.1002/jcb.21213 [PubMed: 17546631]
8. Moschetta M, Mishima Y, Kawano Y, Manier S, Paiva B, Palomera L, et al. Targeting vasculogenesis to prevent progression in multiple myeloma. Leukemia. 2016; 30(5):1103–15. DOI: 10.1038/leu.2016.3 [PubMed: 26859080]
9. Hansmann L, Blum L, Ju CH, Liedtke M, Robinson WH, Davis MM. Mass cytometry analysis shows that a novel memory phenotype B cell is expanded in multiple myeloma. Cancer immunology research. 2015; 3(6):650–60. DOI: 10.1158/2326-6066.CIR-14-0236-T [PubMed: 25711758]
10. Favaloro J, Brown R, Aklilu E, Yang S, Suen H, Hart D, et al. Myeloma skews regulatory T and pro-inflammatory T helper 17 cell balance in favor of a suppressive state. Leukemia & lymphoma. 2014; 55(5):1090–8. DOI: 10.3109/10428194.2013.825905 [PubMed: 23865833]
11. Paiva B, Mateos MV, Sanchez-Abarca LI, Puig N, Vidriales MB, Lopez-Corral L, et al. Immune status of high-risk smoldering multiple myeloma patients and its therapeutic modulation under LenDex: a longitudinal analysis. Blood. 2016; 127(9):1151–62. DOI: 10.1182/blood-2015-10-662320 [PubMed: 26668134]
12. Bianchi G, Munshi NC. Pathogenesis beyond the cancer clone(s) in multiple myeloma. Blood. 2015; 125(20):3049–58. DOI: 10.1182/blood-2014-11-568881 [PubMed: 25838343]
13. Keats JJ, Chesi M, Egan JB, Garbitt VM, Palmer SE, Braggio E, et al. Clonal competition with alternating dominance in multiple myeloma. Blood. 2012; 120(5):1067–76. DOI: 10.1182/blood-2012-01-405985 [PubMed: 22498740]
14. Lohr JG, Kim S, Gould J, Knoechel B, Drier Y, Cotton MJ, et al. Genetic interrogation of circulating multiple myeloma cells at single-cell resolution. Science translational medicine. 2016; 8(363):363ra147.doi: 10.1126/scitranslmed.aac7037
15. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer cell. 2014; 25(1):91–101. DOI: 10.1016/j.ccr.2013.12.015 [PubMed: 24434212]

16. Paino T, Paiva B, Sayagues JM, Mota I, Carvalheiro T, Corchete LA, et al. Phenotypic identification of subclones in multiple myeloma with different chemoresistant, cytogenetic and clonogenic potential. Leukemia. 2015; 29(5):1186–94. DOI: 10.1038/leu.2014.321 [PubMed: 25388955]
17. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. Nature reviews Cancer. 2012; 12(5):335–48. DOI: 10.1038/nrc3257 [PubMed: 22495321]
18. Melchor L, Brioli A, Wardell CP, Murison A, Potter NE, Kaiser MF, et al. Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma. Leukemia. 2014; 28(8):1705–15. DOI: 10.1038/leu.2014.13 [PubMed: 24480973]
19. Thiele B, Kloster M, Alawi M, Indenbirken D, Trepel M, Grundhoff A, et al. Next-generation sequencing of peripheral B-lineage cells pinpoints the circulating clonotypic cell pool in multiple myeloma. Blood. 2014; 123(23):3618–21. DOI: 10.1182/blood-2014-02-556746 [PubMed: 24753536]
20. Garfall AL, Maus MV, Hwang WT, Lacey SF, Mahnke YD, Melenhorst JJ, et al. Chimeric Antigen Receptor T Cells against CD19 for Multiple Myeloma. The New England journal of medicine. 2015; 373(11):1040–7. DOI: 10.1056/NEJMoa1504542 [PubMed: 26352815]
21. Lum LG, Thakur A, Kondadasula SV, Al-Kadhimi Z, Deol A, Tomaszewski EN, et al. Targeting CD138-/CD20+ Clonogenic Myeloma Precursor Cells Decreases These Cells and Induces Transferable Antimyeloma Immunity. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2016; 22(5):869–78. DOI: 10.1016/j.bbmt.2015.12.030
22. Wang X, Stollar BD. Human immunoglobulin variable region gene analysis by single cell RT-PCR. Journal of immunological methods. 2000; 244(1–2):217–25. [PubMed: 11033034]
23. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 2014; 30(5):614–20. DOI: 10.1093/bioinformatics/btt593 [PubMed: 24142950]
24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9(4): 357–9. DOI: 10.1038/nmeth.1923 [PubMed: 22388286]
25. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 2004; 32(5):1792–7. DOI: 10.1093/nar/gkh340 [PubMed: 15034147]
26. Alamyar E, Duroux P, Lefranc MP, Giudicelli V. IMGT((R)) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods in molecular biology. 2012; 882:569–604. DOI: 10.1007/978-1-61779-842-9_32 [PubMed: 22665256]
27. Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nature biotechnology. 2014; 32(7):684–92. DOI: 10.1038/nbt.2938
28. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30(12):2725–9. DOI: 10.1093/molbev/mst197 [PubMed: 24132122]
29. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome research. 2004; 14(6):1188–90. DOI: 10.1101/gr.849004 [PubMed: 15173120]
30. van der Maaten L, Hinton G. Visualizing Data using t-SNE. Journal of Machine Learning Research. 2008; 9:2579–605.
31. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
32. Flores-Montero J, de Tute R, Paiva B, Perez JJ, Bottcher S, Wind H, et al. Immunophenotype of normal vs. myeloma plasma cells: Toward antibody panel specifications for MRD detection in multiple myeloma. Cytometry B Clin Cytom. 2015; doi: 10.1002/cyto.b.21265
33. Tembhare PR, Yuan CM, Venzon D, Braylan R, Korde N, Manasanch E, et al. Flow cytometric differentiation of abnormal and normal plasma cells in the bone marrow in patients with multiple myeloma and its precursor diseases. Leukemia research. 2014; 38(3):371–6. DOI: 10.1016/j.leukres.2013.12.007 [PubMed: 24462038]

34. Kyle RA, Robinson RA, Katzmann JA. The clinical aspects of biclonal gammopathies. Review of 57 cases Am J Med. 1981; 71(6):999–1008. [PubMed: 6797297]
35. Mullikin TC, Rajkumar SV, Dispenzieri A, Buadi FK, Lacy MQ, Lin Y, et al. Clinical characteristics and outcomes in biclonal gammopathies. American journal of hematology. 2016; 91(5):473–5. DOI: 10.1002/ajh.24319 [PubMed: 26840395]
36. Paiva B, Corchete LA, Vidriales MB, Puig N, Maiso P, Rodriguez I, et al. Phenotypic and genomic analysis of multiple myeloma minimal residual disease tumor cells: a new model to understand chemoresistance. Blood. 2016; 127(15):1896–906. DOI: 10.1182/blood-2015-08-665679 [PubMed: 26755711]
37. Perry RP, Kelley DE. Immunoglobulin messenger RNAs in murine cell lines that have characteristics of immature B lymphocytes. Cell. 1979; 18(4):1333–9. [PubMed: 117905]
38. Kosmas C, Stamatopoulos K, Stavroyianni N, Belessi C, Viniou N, Yataganas X. Molecular analysis of immunoglobulin genes in multiple myeloma. Leukemia & lymphoma. 1999; 33(3–4): 253–65. DOI: 10.3109/10428199909058425 [PubMed: 10221505]
39. Magrangeas F, Cormier ML, Descamps G, Gouy N, Lode L, Mellerin MP, et al. Light-chain only multiple myeloma is due to the absence of functional (productive) rearrangement of the IgH gene at the DNA level. Blood. 2004; 103(10):3869–75. DOI: 10.1182/blood-2003-07-2501 [PubMed: 14715636]
40. Gonzalez D, van der Burg M, Garcia-Sanz R, Fenton JA, Langerak AW, Gonzalez M, et al. Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. Blood. 2007; 110(9):3112–21. DOI: 10.1182/blood-2007-02-069625 [PubMed: 17634408]
41. Frey B, Suppmann B. Demonstration of the ExpandTM PCR system's greater fidelity and higher yields with a lacl-based PCR fidelity assay. Biochemica. 1995; 2:34–5.
42. Bergsagel PL, Smith AM, Szczepek A, Mant MJ, Belch AR, Pilarski LM. In multiple myeloma, clonotypic B lymphocytes are detectable among CD19+ peripheral blood cells expressing CD38, CD56, and monotypic Ig light chain. Blood. 1995; 85(2):436–47. [PubMed: 7529064]
43. Szczepek AJ, Seeberger K, Wizniak J, Mant MJ, Belch AR, Pilarski LM. A high frequency of circulating B cells share clonotypic Ig heavy-chain VDJ rearrangements with autologous bone marrow plasma cells in multiple myeloma, as measured by single-cell and in situ reverse transcriptase-polymerase chain reaction. Blood. 1998; 92(8):2844–55. [PubMed: 9763569]
44. Taylor BJ, Pittman JA, Seeberger K, Mant MJ, Reiman T, Belch AR, et al. Intraclonal homogeneity of clonotypic immunoglobulin M and diversity of nonclinical post-switch isotypes in multiple myeloma: insights into the evolution of the myeloma clone. Clinical cancer research : an official journal of the American Association for Cancer Research. 2002; 8(2):502–13. [PubMed: 11839670]
45. Thiago LS, Perez-Andres M, Balanzategui A, Sarasquete ME, Paiva B, Jara-Acevedo M, et al. Circulating clonotypic B-cells in multiple myeloma and monoclonal gammopathy of undetermined significance. Haematologica. 2013; doi: 10.3324/haematol.2013.092817
46. Billadeau D, Ahmann G, Greipp P, Van Ness B. The bone marrow of multiple myeloma patients contains B cell populations at different stages of differentiation that are clonally related to the malignant plasma cell. The Journal of experimental medicine. 1993; 178(3):1023–31. [PubMed: 8350044]
47. Hosen N, Matsuoka Y, Kishida S, Nakata J, Mizutani Y, Hasegawa K, et al. CD138-negative clonogenic cells are plasma cells but not B cells in some multiple myeloma patients. Leukemia. 2012; 26(9):2135–41. DOI: 10.1038/leu.2012.80 [PubMed: 22430638]
48. Kapoor P, Greipp PT, Morice WG, Rajkumar SV, Witzig TE, Greipp PR. Anti-CD20 monoclonal antibody therapy in multiple myeloma. British journal of haematology. 2008; 141(2):135–48. DOI: 10.1111/j.1365-2141.2008.07024.x [PubMed: 18318769]
49. Trepel M, Martens V, Doll C, Rahlff J, Gosch B, Loges S, et al. Phenotypic detection of clonotypic B cells in multiple myeloma by specific immunoglobulin ligands reveals their rarity in multiple myeloma. PloS one. 2012; 7(2):e31998.doi: 10.1371/journal.pone.0031998 [PubMed: 22384124]
50. Bergsagel PL, Masellis Smith A, Belch AR, Pilarski LM. The blood B-cells and bone marrow plasma cells in patients with multiple myeloma share identical IgH rearrangements. Current topics in microbiology and immunology. 1995; 194:17–24. [PubMed: 7534667]

51. Matsui W, Huff CA, Wang Q, Malehorn MT, Barber J, Tanhehco Y, et al. Characterization of clonogenic multiple myeloma cells. Blood. 2004; 103(6):2332–6. DOI: 10.1182/blood-2003-09-3064 [PubMed: 14630803]
52. Matsui W, Wang Q, Barber JP, Brennan S, Smith BD, Borrello I, et al. Clonogenic multiple myeloma progenitors, stem cell properties, and drug resistance. Cancer research. 2008; 68(1):190–7. DOI: 10.1158/0008-5472.CAN-07-3096 [PubMed: 18172311]
53. Pilarski LM, Hipperson G, Seeberger K, Pruski E, Coupland RW, Belch AR. Myeloma progenitors in the blood of patients with aggressive or minimal disease: engraftment and self-renewal of primary human myeloma in the bone marrow of NOD SCID mice. Blood. 2000; 95(3):1056–65. [PubMed: 10648422]
54. Pilarski LM, Seeberger K, Coupland RW, Eshpeter A, Keats JJ, Taylor BJ, et al. Leukemic B cells clonally identical to myeloma plasma cells are myelomagenic in NOD/SCID mice. Exp Hematol. 2002; 30(3):221–8. [PubMed: 11882359]
55. Robiou du Pont S, Cleynen A, Fontan C, Attal M, Munshi N, Corre J, et al. Genomics of Multiple Myeloma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017; 35(9):963–7. DOI: 10.1200/JCO.2016.70.6705 [PubMed: 28297630]
56. Barnidge DR, Dasari S, Botz CM, Murray DH, Snyder MR, Katzmann JA, et al. Using mass spectrometry to monitor monoclonal immunoglobulins in patients with a monoclonal gammopathy. J Proteome Res. 2014; 13(3):1419–27. DOI: 10.1021/pr400985k [PubMed: 24467232]
57. Kim HS, Kim HS, Shin KS, Song W, Kim HJ, Kim HS, et al. Clinical comparisons of two free light chain assays to immunofixation electrophoresis for detecting monoclonal gammopathy. BioMed research international. 2014; 2014:647238.doi: 10.1155/2014/647238 [PubMed: 24971342]
58. Mills JR, Barnidge DR, Murray DL. Detecting monoclonal immunoglobulins in human serum using mass spectrometry. Methods. 2015; 81:56–65. DOI: 10.1016/j.ymeth.2015.04.020 [PubMed: 25916620]
59. Longo NS, Satorius CL, Plebani A, Durandy A, Lipsky PE. Characterization of Ig gene somatic hypermutation in the absence of activation-induced cytidine deaminase. Journal of immunology. 2008; 181(2):1299–306.
60. Ying WZ, Allen CE, Curtis LM, Aaron KJ, Sanders PW. Mechanism and prevention of acute kidney injury from cast nephropathy in a rodent model. The Journal of clinical investigation. 2012; 122(5):1777–85. DOI: 10.1172/JCI46490 [PubMed: 22484815]

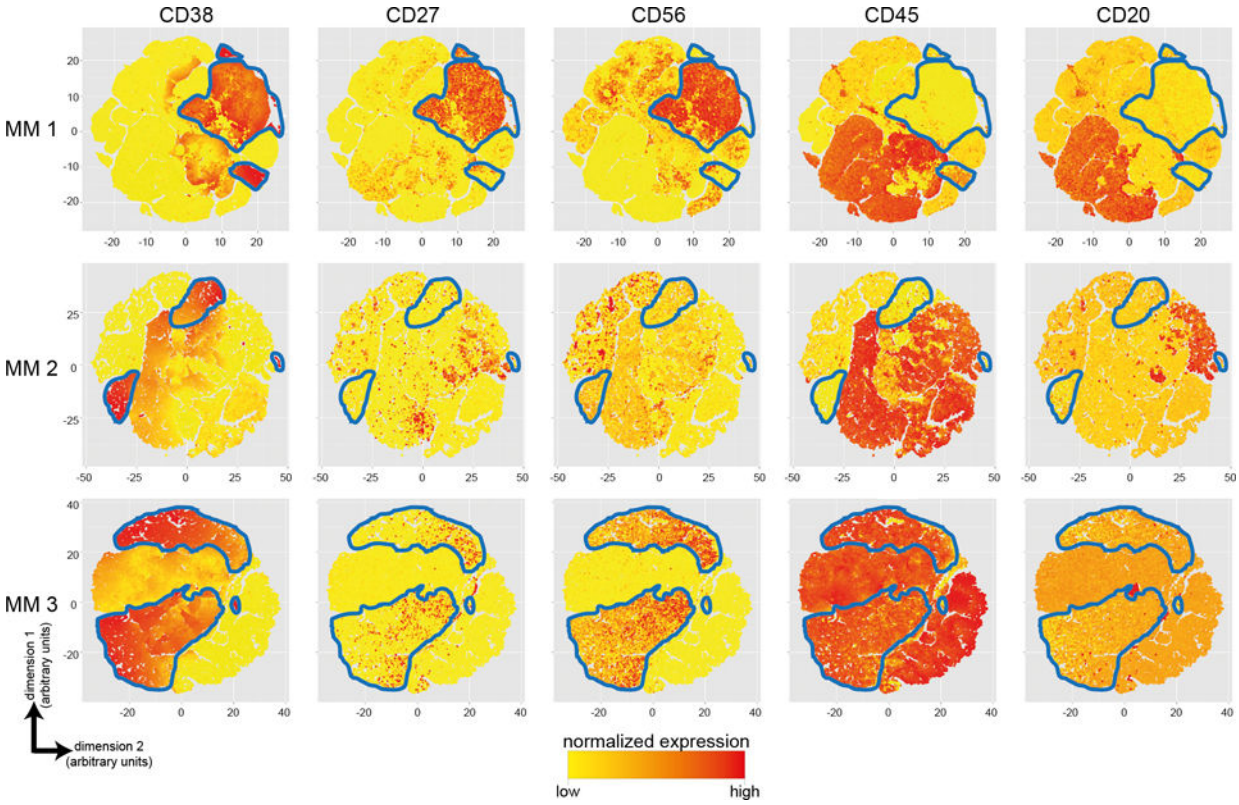

Figure 1. Phenotypic diversity of plasma cell populations in multiple myeloma bone marrow

t-SNE plots were generated from 13-parameter bone marrow FACS data of three multiple myeloma patients. Each data point represents a single cell and the location of each single cell is determined taking into account the distances of all markers in the FACS panel between individual cells. The expression levels of the markers indicated on top were normalized and color-coded (yellow to red). Plasma cells (in all samples $CD38^{hi}CD20^{-}$) occupy heterogeneous phenotypic locations and show variable expression of CD27, CD56, and CD45 underlining the intra- and inter-sample phenotypic diversity. The areas potentially including polyclonal physiologic and/or monoclonal malignant plasma cells based on high CD38 expression are circled in blue. Cells of each patient were FACS-sorted once. MM: multiple myeloma

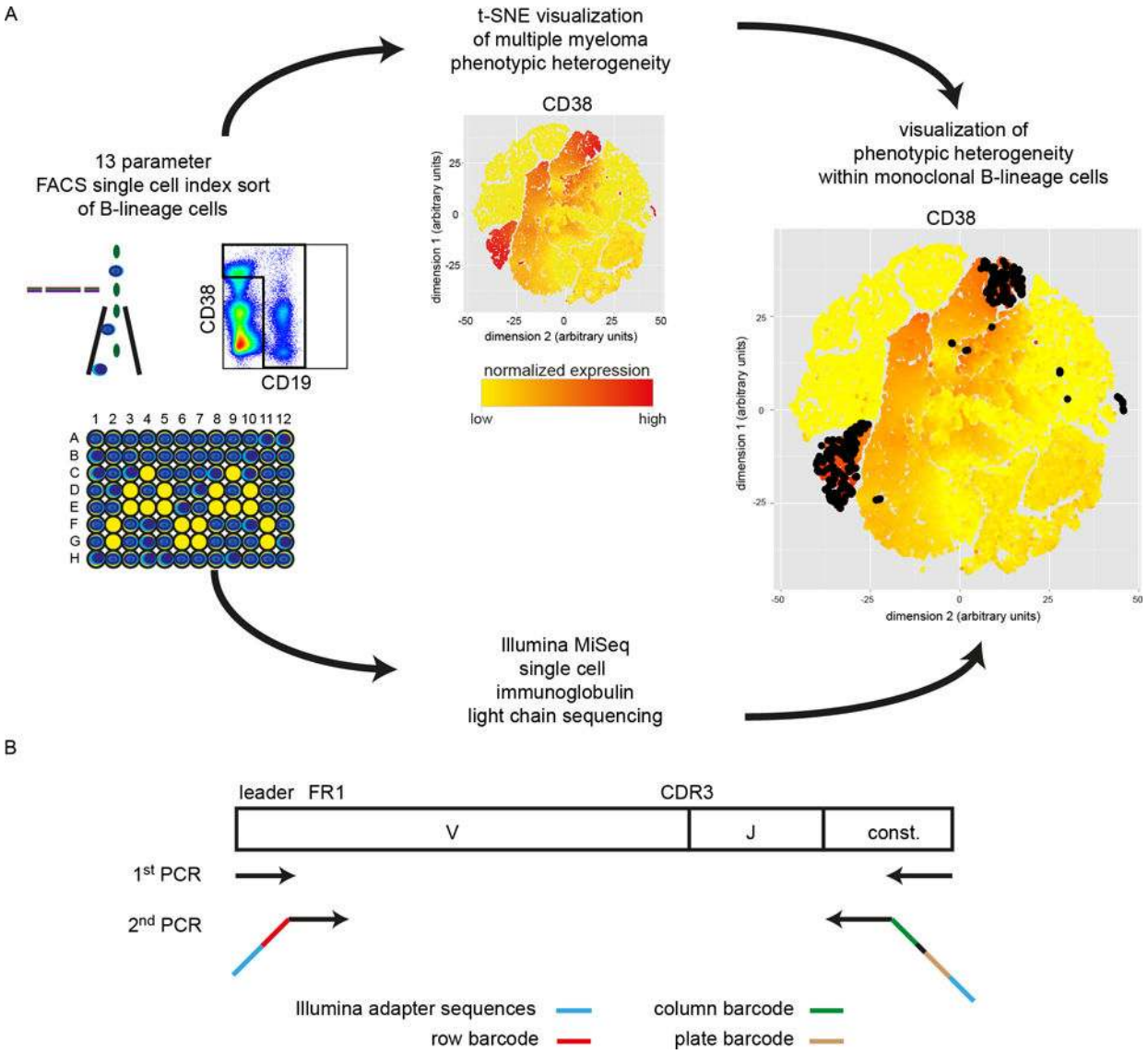

Figure 2. Technology schematic for phenotypic tracking of single molecularly defined B-lineage clones

(A) Single B-lineage cells identified by CD19 and CD38 characteristics were index sorted into 96-well plates using a 13-parameter FACS panel (see Supplementary Fig. S1 for detailed gating strategy). Single-cell light chain mRNA was reverse transcribed, amplified, barcoded, and sequenced. In parallel, FACS data were visualized as t-SNE maps and single monoclonal B-lineage cells of the predominant clone were mapped onto the t-SNE plots (black) to visualize their phenotypic distribution. (B) Illustration of the amplification and barcoding strategy. For each single cell a first amplification is followed by a nested second amplification that introduces plate-, column-, and well-specific barcodes and the first part of the Illumina adapter sequences. The remaining parts of the adapter sequences were attached in a third PCR reaction (not shown in the figure, see Suppl. Tab. S4 for primer sequences).

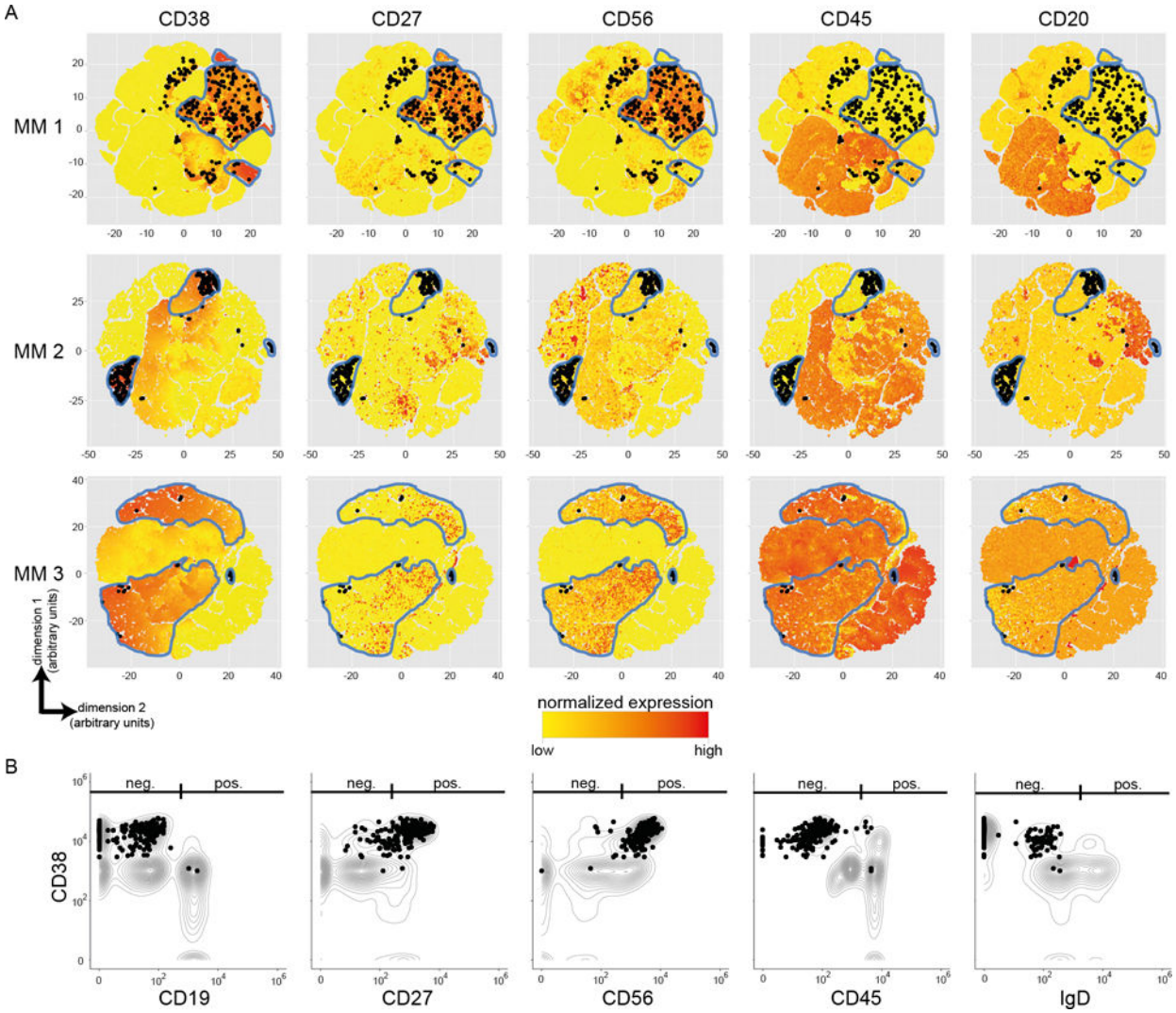

Figure 3. Phenotypic diversity of individual multiple myeloma clones

(A) Cells of the predominant clones were overlaid in black over t-SNE maps of three multiple myeloma samples. The areas potentially including physiologic and/or malignant plasma cells based on high CD38 expression are circled in blue. (B) For a more detailed visualization of individual phenotypes, single cells of the predominant B-lineage clone in MM1 were plotted as filled black circles in two-dimensional space. Positive and negative gates were set based on the distribution of all cells (clonal and non-clonal) in the individual dataset visualized as contour in the background. The absolute counts of monoclonal cells of the predominant clones are 136, 218, and 64 for MM1, MM2, and MM3 respectively. Note that there can be positional overlap due to similar phenotypes. MM: multiple myeloma

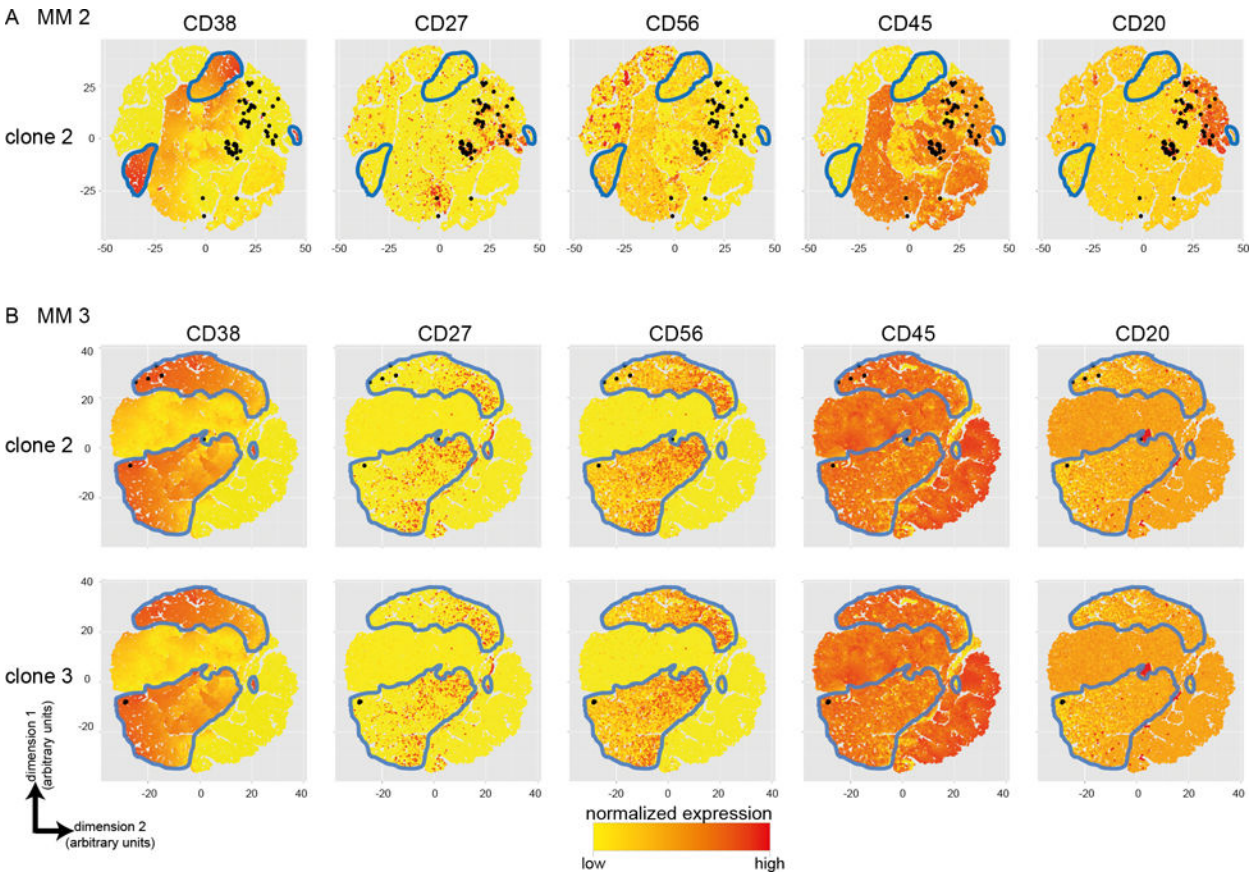

Figure 4. Phenotypic diversity of less predominant B-lineage clones

Immune phenotypes of cells of less predominant clones (in black) were visualized. The areas presumably representing physiologic and/or malignant plasma cell phenotypes are circled in blue. While cells of the second predominant clone of multiple myeloma 2 (A) co-localize with normal-phenotype B lymphocytes, cells of the second and third predominant clones in multiple myeloma 3 (B) map to areas that have been shown to be occupied with multiple myeloma cells in Fig. 3. See Supplementary Fig. S6 for detailed immune phenotypes.

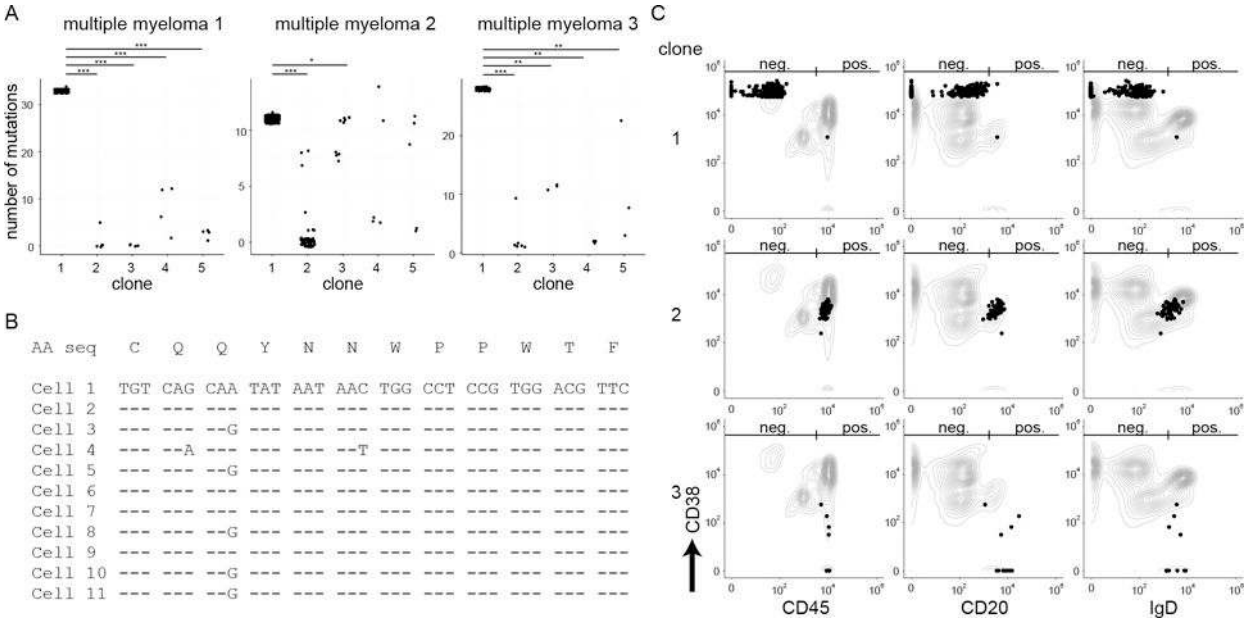

Figure 5. Convergent expansion in less predominant B-lineage clones

(A) Numbers of somatic mutations in the V genes of the five most predominant clones (1–5) in three multiple myeloma samples were determined. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. P values were calculated using the Wilcoxon Rank Sum test and corrected for multiple testing applying Bonferroni correction. (B) Shows the alignment of the CDR3 nucleotide sequences of the third predominant clone of multiple myeloma 2 as an example. The CDR3 shows silent nucleotide exchanges when comparing sequences from different cells suggesting an antigen-driven convergent expansion process. (C) Shows phenotypic characteristics for selected markers in the three predominant clones (filled black circles) of multiple myeloma 2 as an example. Positive and negative gates were defined based on the distribution of cells in the whole dataset (contour). For a detailed visualization of all markers in all investigated clones see Supplementary Fig. S6. AA seq: amino acid sequence.

Table 1

Patient characteristics

Patient 3 received chemotherapy for a period of six months completing the last cycle two months before the study-related bone marrow aspiration. Additional clinical and laboratory data can be found in Supplementary Table S1. MM: multiple myeloma, F: female, M: male, BM: bone marrow

MM	Age (years)	Sex	Ig-/light chain restriction	Stage (Salmon and Durie)	BM plasma cell infiltration (aspirate)	Last chemotherapeutic regimen
1	48	M	IgG λ	III	37 %	none
2	58	M	λ light chain	III	47 %	none
3	70	F	λ light chain	III	12 %	bortezomib, dexamethasone, cyclophosphamide

Table 2

clone frequencies and CDR3 sequences

Numbers indicate absolute cell numbers, percentages indicate percentages of all sequenced cells in the individual patients. Sequences were aligned using MUSCLE.

Clone	multiple myeloma 1			multiple myeloma 2			multiple myeloma 3		
	# cells	CDR3 sequence	light chain	# cells	CDR3 sequence	light chain	# cells	CDR3 sequence	light chain
1st	136 (28 %)	Q—VWDSDSDSYV	λ	218 (36 %)	—QAWDSSTVL	λ	64 (23 %)	QAWDRST–VV	λ
2nd	5 (1 %)	A—AWDDSLSGPV	λ	54 (9 %)	QQYNNWPPRYT-	κ	5 (2 %)	QQYNNWPPRYT	κ
3rd	4 (1 %)	A—AWDDSLNGPV	λ	11 (2 %)	QQYNNWPP-WT-	κ	3 (1 %)	QAWDSST–VL	λ
4th	4 (1 %)	A—AWDDSLNGVV	λ	5 (1 %)	—QAWDSSTVV	λ	3 (1 %)	QQYDNLP–LT	κ
5th	4 (1 %)	QQRSNWP–LT—	κ	5 (1 %)	QQANSFPR–T-	κ	3 (1 %)	QQYGSSP–FT	κ
# cells		483			603			283	