

# Deciphering spatial genomic heterogeneity at a single cell resolution in multiple myeloma

**Maximilian Merz** (✉ [maximilian.merz@roswellpark.org](mailto:maximilian.merz@roswellpark.org))

Roswell Park Comprehensive Cancer Center

**Almuth Merz**

Roswell Park Comprehensive Cancer Center

**Jie Wang**

Roswell Park Cancer Institute

**Lei Wei**

Roswell Park Cancer Institute <https://orcid.org/0000-0002-1881-8666>

**Qiang Hu**

Roswell Park Cancer Institute <https://orcid.org/0000-0002-4090-5539>

**Nicholas Hutson**

Roswell Park Comprehensive Cancer Center

**Cherie Rondeau**

Roswell Park Comprehensive Cancer Center

**Kimberly Celotto**

Roswell Park Comprehensive Cancer Center

**Ahmed Belal**

Roswell Park Comprehensive Cancer Center

**Ronald Alberico**

Roswell Park Comprehensive Cancer Center

**AnneMarie Block**

Roswell Park Comprehensive Cancer Center

**Hemn Mohammadpour**

Roswell Park Cancer Institute

**Paul Wallace**

Roswell Park Cancer Institute

**Joseph Tario, Jr.**

Roswell Park Cancer Institute

**Jesse Luce**

Roswell Park Comprehensive Cancer Center

**Sean Glenn**

Roswell Park Comprehensive Cancer Center

**Prashant K. Singh**

Roswell Park Cancer Institute

**Megan Herr**

Roswell Park Comprehensive Cancer Center

**Theresa Hahn**

Roswell Park Comprehensive Cancer Center <https://orcid.org/0000-0002-3835-8855>

**Mehmet Samur**

Dana-Farber Cancer Institute <https://orcid.org/0000-0002-9978-5682>

**Nikhil Munshi**

Dana-Farber Cancer Institute <https://orcid.org/0000-0002-7344-9795>

**Song Liu**

Roswell Park Cancer Institute

**Philip McCarthy**

Roswell Park Cancer Institute

**Jens Hillengass**

Roswell Park Comprehensive Cancer Center, Buffalo, NY

---

**Article**

**Keywords:** Multiple myeloma, osteolytic lesion, bone disease, single-cell RNA sequencing, heterogeneity

**Posted Date:** January 16th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-146899/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Nature Communications on February 10th, 2022. See the published version at <https://doi.org/10.1038/s41467-022-28266-z>.

# Abstract

Osteolytic lesions (OL) characterize symptomatic multiple myeloma. The mechanisms of how malignant plasma cells (PC) cause OL in one region while others show no signs of bone destruction despite subtotal infiltration remain unknown. We report the first single-cell RNA sequencing (scRNA-seq) study of PC obtained prospectively from random bone marrow aspirates (BM) and paired imaging-guided biopsies of OL. We analyzed 148,630 PC from 24 different locations in 10 patients and observed vast inter- and intra-patient heterogeneity based on scRNA-seq analyses. Beyond the limited evidence for spatial heterogeneity from whole-exome sequencing, we found an additional layer of complexity by integrated analysis of anchored scRNA-seq datasets from the BM and OL. PC from OL were characterized by differentially expressed genes compared to PC from BM, including upregulation of genes associated with myeloma bone disease like *DKK1*, *HGF* and *TIMP-1* as well as recurrent downregulation of *JUN/FOS*, *DUSP1* and *HBB*. Assessment of PC from longitudinally collected samples revealed transcriptional changes after induction therapy. Our study, based on the largest number of PC analyzed by scRNA-seq, contributes to the understanding of destructive myeloma bone disease.

## Introduction

Multiple Myeloma (MM) is a heterogeneous disease with survival ranging from months to decades (1). Malignant plasma cells (PC) for histopathology and genetic assessment are isolated from iliac crest bone marrow aspirates in routine practice. However, PC are not homogeneously distributed within the bone marrow. Osteolytic lesions (OL) are areas of circumscribed bone loss caused by malignant PC infiltration. While OL can be visualized by positron emission computed tomography (PET/CT) in up to 80% of patients, their underlying biology remains to be clarified. Some patients show subtotal PC infiltration of the bone marrow in the iliac crest without signs of bone destruction while in the same patients, PC cause bone disease in distant locations such as the vertebral bodies. Therefore, OL might represent regions of increased infiltration as well as areas containing biologically different PC.

Efforts have been made to classify MM patients based on copy number changes (2), mutational burden (2) and gene expression profiling (GEP) (3–6). Recently, a retrospective multi-region whole exome sequencing (WES) study showed for the first time spatial genomic heterogeneity of paired samples from random bone marrow aspirates and distant lesions (7). Since focal lesions before and after therapy are associated with adverse outcome (8), site-specific high-risk PC populations might be responsible for treatment resistance and relapse. Therefore, sampling PC solely from the iliac crest might not be representative due to intra-patient spatial heterogeneity.

Furthermore, cancers are not composed by an aggregation of genetically identical cells, and bulk tissue sequencing might obscure biologically relevant differences between cells. Among several emerging

technologies to interrogate tumors at a single-cell resolution, single-cell RNA sequencing (scRNA-seq) can identify treatment resistant clones and subpopulations responsible for metastatic spread in several human cancers (9–12). The first scRNA-seq study in MM examined 20,568 PC from bone marrow samples from 29 patients with a variety of plasma cell disorders. This analysis demonstrated PC heterogeneity in MM and identified circulating tumor cells as well as minimal residual disease (MRD) after therapy (10). More recently, scRNA-seq was used to study the microenvironment in asymptomatic patients for the first time (13).

In this study, we conducted the first scRNA-seq analysis of 148,630 freshly purified PC obtained prospectively from random bone marrow aspirates and paired imaging-guided biopsies of OL in 10 patients with symptomatic MM. We demonstrate that scRNA-seq of PC from OL is feasible in a prospective clinical trial. Based on single cell transcriptomics, we observed inter- as well as intra-patient heterogeneity. While we showed limited spatial heterogeneity based on WES, scRNA-seq identified significant differences between both locations. Assessment of PC sampled after induction therapy showed transcriptional changes compared to baseline findings. Our study revealed a new layer of complexity to spatial heterogeneity in MM and contributes to the understanding of myeloma bone disease.

## Results

### Single cell RNA sequencing of plasma cells from guided biopsies of osteolytic lesions and corresponding bone marrow identifies inter-patient heterogeneity

We implemented a translational workflow (**Figure 1A**) to obtain and purify viable PC from OL and corresponding BM samples from patients with newly diagnosed or relapsed MM. The bone marrow biopsies from the iliac crest (bone marrow sample, BM) were obtained and processed at the same time as OL biopsies since differences in sample processing times might cause changes in MM gene expression (14). We performed scRNA-seq on paired samples from 10 patients (7 with newly diagnosed and 3 relapsed/refractory MM) with one patient having 2 OL biopsied. In three patients with NDMM, we obtained additional samples after induction therapy. Patient characteristics are summarized in **Table 1**.

Clustering of 148,746 single cells from BM and OL (median 7712 cells/sample) created a map of distinct populations based on transcriptomes from individual patients (**Figure 1B**). There were no significant differences between both locations regarding the purity of isolated and sequenced PC, underlining the feasibility of our protocol and comparability of paired samples. In total, 94.8% of cells from OL (n = 70,036) and 95.7% of cells from BM (n = 71,580) were PC (**Figure 1C**). Since the cell cycle can introduce heterogeneity in gene expression in very homogeneous cell populations like in our study, we assigned cell cycle scores to PC and found no significant differences between OL and BM (**Figure 1D**).

Next, we identified marker genes for malignant PC clusters from each individual patient (**Figure 1E**). As expected, the genes for cyclin D1-3 (*CCND1-3*) were preferentially expressed in patients with IgH translocations (NDMM02, NDMM04 and NDMM06) as detected by FISH (**Table 1**). Furthermore, genes associated with myeloma bone disease (*DKK1* and *FRZB*), cytokine signaling (*IFI27* and *IL6R*) and *EDNRB* were identified as marker genes for PC clusters from individual patients. All of the aforementioned genes were previously described as characteristic genes elevated in the molecular subtypes of MM that were derived from bulk GEP (15, 16). Besides these known marker genes, we identified *STMN1* to be preferentially expressed in small subsets of malignant PC (**Figure 1E**). No significant differences for the identified marker genes were found between BM and OL.

Following single gene analysis, we performed a gene set enrichment analysis (GSEA) using the curated MM subtype gene sets from the Molecular Signature Database (MSigDB). Consistent with the single gene analysis, gene-set analysis showed that patients could be grouped according to the molecular classification of MM. For example, patient RRMM02 in the low bone disease (LB) group showed higher expression of *IL6R* and *EDNRB* and NDMM02 in the MAF group (MF) showed higher levels of *CCND2* and lower levels of *DKK1* (**Figure 1F**). This demonstrates that scRNA-seq reveals inter-patient heterogeneity and identifies MM subtypes for individual patients.

### **Intra-patient heterogeneity based on scRNA-seq and inferred CNVs**

After investigating inter-patient heterogeneity using scRNA-seq data, we aimed at deciphering intra-patient heterogeneity on the transcriptional and chromosomal level. We performed clustering and differential expression analysis of PC for each individual patient. Specifically, we looked for genes that were recurrently differentially expressed in PC clusters and that have been associated to play a role in the pathogenesis and prognostication of MM. A summary of the differential expression analysis for each individual patient is available at **Supplemental Table 2**.

In 9 out of 10 patients, we identified clusters of malignant PC characterized by the overexpression of genes encoding the microtubule-associated proteins *STMN1* and *TUBA1B* (**Figure 2A** and **Supplemental Figure 1**). *STMN1* is among the 15 genes associated with high-risk disease in the GEP score identified by the Integroupe Francophone du Myelome (4). Furthermore, higher expression levels of *STMN1* and *TUBA1B* are associated with shorter progression-free and overall survival in the CoMMpass dataset of the Multiple Myeloma Research Foundation (**Supplemental Figure 2**). This underlines the biological and

prognostic relevance of the identified subclusters. Since the respective clusters accounted only for a small number of malignant PC in each individual patient (**Supplemental Figure 1**) they would have been missed by bulk sequencing.

CNVs are another cornerstone in the biology of MM and we have shown recently that cytogenetically defined subclones have major prognostic implications (17). However, in routine practice, chromosomal aberrations are usually detected by fluorescence in situ hybridization (FISH). The detection of subclones by FISH is limited by the number of analyzed cells and applied probes. We leveraged the scRNA-seq data to identify genome-wide CNVs and subclones for 141,616 PC using inferCNV (18). **Figure 2B** provides two examples for CNVs from a patient with newly diagnosed disease (NDMM01) and a patient with relapsed disease (RRMM01). UPhyloPlot2 (19) was used to draw phylogenetic trees from inferCNV outputs. In 6 of 7 with NDMM and 2 of 3 patients with RRMM, we were able to detect subclones based single cell transcriptomes (**Figure 2C** and **Supplemental Figure 3**). No significant differences in CNVs were found when comparing OL and BM.

### Limited evidence for spatial heterogeneity from whole-exome sequencing

Next, we aimed at characterizing spatial heterogeneity. Whole-exome sequencing (WES) has been utilized to investigate spatial genomic heterogeneity in MM based on banked, frozen PC from a retrospective study (7). To explore whether our study from scRNA-seq on freshly isolated PC could reveal another layer of complexity of spatial heterogeneity, we performed WES on all paired samples in our study. The comparison of PC with the matched germline normal identified a total of 1103 somatic mutations, including 1063 SNVs and 40 Indels (**Figure 3A**). Among these somatic mutations, 665 were predicted to cause an amino acid alteration, 72 were truncating and 366 were silent mutations. **Supplemental table 2** gives an overview of the individual variant calls from the WES analysis. For each patient, we quantified the similarity between the BM and OL by calculating the Jaccard score, defined as the ratio between shared mutations and all mutations. Patient NDMM03 was excluded from calculations since only limited numbers of malignant PCs were captured from the BM.

In 8 out of 10 patients (NDMM02-07 and RRMM02-03), the percentage of shared mutations between the bone marrow and OL were ~ 80% or higher, suggesting the BM and OL were highly consistent (**Figure 3 A and B**). In the patient with 2 OL (NDMM06), the Jaccard scores among all 3 locations (BM and two OL) were above 97% with a clonal *TP53* mutation present in all three locations.

Only in 2 out of 10 patients, WES revealed relevant spatial heterogeneity:

In patient NDMM01, 75% of all mutations were shared between the bone marrow and OL, and 25% of mutations were only present in the OL (**Figure 3A**) including a *BRAF V600E* mutation. In patient RRMM01 with an OL of the right clavicle with extramedullary spread, only 20% of all mutations were shared, with 24% of the mutations found only in the bone marrow, and 56% of the mutations unique to the OL (**Figure 3A**). Two distinct *BRAF* mutations were detected: *V600E* was only in BM, and another activating Class 2 *BRAF* mutation (*G469R*) in the OL. The latter mutation causes resistance to the BRAF inhibitor vemurafenib (20). Furthermore, we detected an additional *NRAS* mutation (*G12D*) in the OL. *NRAS* mutations have been associated to drive spatially divergent resistance to vemurafenib in BRAFmut MM (21). Both examples demonstrate that treatment with a BRAF inhibitor would most likely be ineffective against PC from different locations in individual patients.

### Single cell RNA sequencing revealed another layer of complexity and links site-specific gene expression to the development of osteolytic lesions

After analyzing WES from malignant PC from the OL and BM, we investigated whether the observed similarities between both locations are also reflected by scRNA-seq data. Therefore, average gene expression of PC from the OL and BM were correlated to each other (**Figure 3B**). In agreement with findings from WES, we found a high concordance of average gene expression between both locations. However, in every patient outliers in both directions were observed (**Figure 3B**).

To identify genes that are differentially expressed in PC from both conditions, we performed an integrated analysis after anchoring datasets from OL and BM for each individual patient as described before (22). After applying the integration procedure, malignant PC were robustly detected in all datasets and the same PC clusters were identified in BM and OL (**Figure 3C**). In all patients, we were able to find marker genes that were differentially expressed in malignant PC from OL compared to BM. Overall, 1140 genes were identified that were differentially expressed between OL and BM (**Figure 3D**). Genes that have been associated with the development of myeloma bone disease such as *DKK1*, *HGF* and *TIMP-1* (23, 24) were among the markers with higher expression levels in OL (**Figure 3D**). Furthermore, in agreement with the first scRNA-seq analysis in MM (10), we found *LAMP5* to be upregulated in PC from the OL in patient NDMM03 (**Figure 4B**).

Genes that were recurrently downregulated in PC from OL were *JUN* and *FOS* (6 of 10 patients), dual specificity phosphatase 1 (*DUSP1*, 5 of 10 patients) and hemoglobin beta chain (*HBB*, 3 of 10 patients)

(**Figure 3D and E**). Importantly, no somatic mutations were detected in the respective genes in PC from the BM and OL. Downregulation of *JUN/FOS*, *DUSP1* and *HBB* has been connected to extramedullary spread of MM in the past (25). Furthermore, *JUN/FOS* are linked to the malignant transformation of B-cells (26) and dissemination of clonal PC in a preclinical model (27). GSEA confirmed that downregulation of genes in pathways associated with normal B-cell was common in PC from the OL (**Figure 3F**).

Additionally, lower expression levels of genes encoding for the non-restricted light (e.g. *IGLC2/3* in NDMM01, **Figure 4A**) and heavy chain (e.g. *IGHM* in NDMM03, **Figure 4B**) were observed in PC from OL. Downregulation of the affected heavy chain was observed in PC from the OL in RRMM01 and RRMM03 providing a first link between evolving disease and the rare phenomenon of light chain escape that can be observed in heavily pretreated RRMM and EMD.

In patient RRMM01 with a PC tumor with extramedullary spread, we identified the gene for zinc-alpha2-glycoprotein (*AZGP1*) to be downregulated in the OL (**Figure 4D**). *AZGP1* is a known tumor suppressor gene causing mesenchymal-to-epithelial transition and inhibiting invasion and metastasis (28). While its expression was virtually absent in PC from the OL, it was homogeneously expressed in the magnitude of PC from the BM. However, a small cluster of PC with lower expression levels compared to the majority of PC was identified in the BM (**Figure 4D**). Trajectory inference revealed that PC from the respective cluster underwent the largest transcriptional change, might have occurred latest in the developmental process and might have given rise to the OL in the right clavicle (**Figure 4D**).

Our results show that scRNA-seq adds another layer of complexity compared to WES to spatial heterogeneity in MM. The scRNA-seq data allow us to link site-specific gene expression to the development of myeloma bone disease and identify subclusters that might be the origin for OL.

### **Single cell RNA sequencing identifies minimal residual disease and transcriptional changes after therapy**

In three patients (NDMM01, NDMM03 and NDMM06) we collected samples after 4 cycles of induction therapy. While in NDMM01 (**Figure 4A**) and NDMM06 (**Figure 4B**) we performed a regular bone marrow biopsy to assess residual disease, an imaging-guided biopsy of a residual focal MRI lesion in T8 was biopsied in NDMM03 (**Figure 4C**). NDMM01 and NDMM03 were in MRD-positive complete remission (CR) after 4 cycles RVD. NDMM06 had achieved a partial response (PR) after 4 cycles daratumumab-RVD. Correspondingly, we captured less malignant PC in NDMM01 (**Figure 4A**) and NDMM03 (**Figure 4B**), while

almost the same amount of malignant PC after therapy were isolated and sequenced in NDMM06 compared to the BM and the two OL at primary diagnosis (**Figure 4C**).

To identify transcriptional changes after therapy, we performed a differential expression analysis after integrating datasets before and after therapy. In both patients in MRD-positive CR, higher expression levels of *HLA-DRA*, *HLA-DRB1* and *HLA-DPB1* were observed in malignant PC after 4 cycles RVD.

In patient NDMM06 in PR, *ISG15* was upregulated in residual PC after 4 cycles of daratumumab-RVD. *ISG15* encodes an ubiquitin-like protein and its expression has been linked to Carfilzomib-resistance in a recent preclinical study (29). Additionally, *TXNIP* and *DDIT4* were downregulated upon therapy. Both genes are induced by dexamethasone and involved in glucocorticoid-induced apoptosis providing a link to steroid-refractoriness in the detected residual disease (30, 31).

This demonstrates that scRNA-seq identifies MRD in patients in remission and reveals transcriptional changes consistent with immunomodulatory effects of lenalidomide and drug resistance to proteasome inhibitors.

## Discussion

In this prospective study, we established a reproducible workflow to obtain and isolate viable PC of comparable quality from BM and paired OL. Our study links the accumulation of malignant PC with distinct transcriptomes to the development of myeloma bone disease and identifies transcriptional changes in residual disease.

scRNA-seq has been successfully utilized to study PC heterogeneity by Ledergor *et al.* who sequenced 24,126 single cells from random bone marrow samples of 40 individuals (11 normal controls and 29 patients with PC disorders of which 12 were MM patients) (10). After removal of contaminating non-PC (n=3,179, 15.4%), 20,568 single PC were analyzed with the focus on inter-patient and inter-diagnostic differences. In the current study, we analyzed 148,630 PC from 24 different locations (BM and paired distant OL) in 10 individuals with MM. We not only confirmed that subsets of single malignant PC in individual patients can be delineated based on single cell transcriptomics and inferred CNVs, but also identified subclusters of malignant PC that are characterized by the overexpression of genes associated with adverse outcome (*STMN1* and *TUBA1B*). Given the limited number of cells compared to the entire population of malignant PC in the respective patients, its significance would have been missed by bulk sequencing.

Beyond the identification inter- and intra-patient heterogeneity, we aimed at deciphering spatial heterogeneity and the underlying biology of OL. So far, only one retrospective study on frozen samples used WES and GEP, demonstrating spatial genomic heterogeneity by multi-regional sequencing in more than 75% of their analyzed patients (n=51) (7). In contrast to the study by Rasche *et al.* and in line with a recent study of WES on circulating PC and corresponding bone marrow PC (32), we demonstrated a high concordance between PC from BM and OL based on WES. Possible explanations for these differences might be the retrospective character of the study by Rasche *et al.* and the inclusion of more patients with advanced disease including extramedullary MM. Both might have introduced sampling bias compared to our prospective acquisition of fresh PC.

By analyzing scRNA-seq data from paired OL and BM, we revealed another layer of complexity and additional detail to the understanding of PC heterogeneity. Although single PC transcriptomes from both locations were highly correlated to each other, we identified differentially expressed genes in every individual patient through integrated analysis of anchored datasets. Beyond genes that have been associated with myeloma bone disease (*TIMP1*, *HGF*(24)) and impaired Wnt-signaling (*DKK1*(23)), we identified *LAMP5* to be overexpressed in OL. Our findings connecting *LAMP5* to the development of OL is supported by a recent retrospective trial in which bulk GEP was performed on PC from patients with smoldering myeloma with and without progression during follow-up. *LAMP5* was significantly overexpressed in patients with disease progression and 8 of 10 patients progressed with new OL in the respective study (33).

Additionally, *JUN/FOS*, *DUSP1* and *HBB* were recurrently downregulated in PC from OL. The respective genes have been connected to the extramedullary spread of MM in the past (25). *JUN/FOS* have been associated with malignant B-cell transformation (26) and progression of MM in a recent preclinical study (27). GSEA confirmed that downregulation of genes connected to regular B-cell function were a common feature of PC from OL.

The anti-metastatic tumor suppressor gene *AZGP1* was also downregulated in the EMD from patient RRMM01, while a small subset of PC with lower expression levels of *AZGP1* could also be identified in BM. Our findings and recent preclinical studies (27) support the hypothesis that OL are derived from a common PC ancestor developing molecular features to cause myeloma bone disease. Thus, MM would behave like a solid tumor with PC metastasizing to distant locations, inducing OL.

Beyond transcriptional changes in PC from OL, we were able to detect changes upon therapy in patients with residual disease. In both MRD-positive patients after 4 cycles RVD, we found an upregulation of HLA class II genes consistent with immunomodulatory effects on PC described for lenalidomide in preclinical studies (34). In the high-risk patient with del(17p) and a *TP53* mutation (NDMM06), genes associated with carfilzomib-resistance and steroid-induced apoptosis were dysregulated in residual PC after 4 cycles of daratumumab-RVD. These results indicate that scRNA-seq helps to identify residual disease and changes in single PC transcriptomes might help to identify new strategies to eradicate MRD.

Taken together, we have shown that scRNA-seq is feasible to analyze transcriptional heterogeneity in fresh clinical PC samples from a prospective trial. We provide for the first time a link between site-specific accumulation of malignant PC with a transcriptomic profile and the development of myeloma bone disease. Our findings contribute to the current understanding of MM heterogeneity and have implications follow-up and therapeutic decision-making.

## Patients And Methods

### Prospective trial of imaging-guided biopsies

In April 2019 we initiated a prospective clinical trial to analyze spatial and temporal evolution in newly diagnosed and relapsed MM. Within the ongoing trial, patients consented to an imaging-guided biopsy of OL identified by PET/CT in addition to the regular random bone marrow aspirate from the iliac crest (bone marrow sample, BM). Biopsies were performed at study entry and before the initiation of local or systemic therapy. Eligible patients with a confirmed diagnosis of MM according to International Myeloma Working Group (IMWG) criteria were at least 18 years of age with an Eastern Cooperative Oncology Group (ECOG) performance status of 0-2 and no contraindications against general anesthesia. Key exclusion criterion was a history of other malignancy except if the patient had been symptom-free and without active therapy for at least 5 years. The ongoing study was approved by the Roswell Park Comprehensive Cancer Center (Roswell Park) Institutional Review Board and was conducted in accordance with the Declaration of Helsinki.

### Medical imaging

Every patient with newly diagnosed MM before the initiation of therapy received a PET/CT according to the current IMWG guidelines (35). Imaging was performed on a GE Discovery ST PET/CT (GE Healthcare, Chicago, IL). Approximately 60 minutes after the intravenous administration of 10 mCi of <sup>18</sup>F-fluorodeoxyglucose (FDG) a low-dose, non-contrast, diagnostic quality whole-body CT was performed for assessment of bone disease, attenuation correction and anatomic orientation. Axial images were reconstructed at 3.75 mm. Afterwards a PET covering the same anatomical regions was obtained and

reconstructed with and without attenuation correction. OL were characterized by the presence of circumscribed areas with bone loss and increased tracer uptake compared to the surrounding bone marrow background.

### **Imaging guided biopsies of osteolytic lesions**

Every potential patient with new OL on PET/CT was discussed in a weekly multi-disciplinary tumor board. Patients were offered to participate in the study if interventional radiologists identified at least one potentially accessible lesion. Upon obtaining written informed consent, patients were scheduled for a biopsy either under CT or fluoroscopic guidance. Patients were placed in a prone position on the CT or fluoroscopy table. Local anesthesia of the skin and soft tissue surrounding the OL was performed using a combination of 1% lidocaine and 0.25% preservative free bupivacaine. Under intermittent CT/fluoroscopy guidance, a 13-gauge trocar needle was advanced to the bone surface. A powered bone access system (Arrow® OnControl®, Teleflex, NC) was used to advance the needle through the outer cortex, since it has been shown to reduce pain, procedure time and increases quality of specimen compared to manually obtained samples(36). After CT-/fluoroscopy-guided confirmation of the correct placement of the needle in the OL, 10 ml bone marrow aspirate and a core biopsy were obtained of the OL. The similar process was performed on the iliac crest for the RS.

### **Obtaining viable plasma cells from random aspirates and osteolytic lesions**

**Figure 1** summarizes the translational workflow in our prospective trial. Bone marrow aspirates from the RS and OL were collected in tubes containing Ethylenediaminetetraacetic acid (EDTA, BD Vacutainer®, BD, NJ). Plasma cells were immediately isolated using a CD138 positive selection kit according to manufacturer's instructions (EasySep™, STEMCELL Technologies, Vancouver, Canada). Cell numbers and viability of the positive and negative fraction after plasma cell separation were checked with Trypan Blue using an automated counter (Countess™ II, Thermo Fischer, MA). Plasma cell purity was assessed by fluorescence-activated cell sorting (FACS). Only in patient RRMM01 with an osteolytic lesion of the right clavicle and adjacent extramedullary disease (EMD), we obtained a soft tissue tumor that was mechanically disintegrated to bring cells into suspension.

On the same day, plasma cells were resuspended in RPMI1640 containing 10% fetal bovine serum (FBS) and subjected to scRNA-seq. Plasma cells that were not transferred for scRNA-seq were resuspended in FBS containing 10% Dimethylsulfoxide (DMSO) and frozen at -80°C.

In total, 167,453 of the initially sequenced 220,654 single cells from BM, OL and after therapy (median 7712 cells/sample) were captured after filtering out cells characterized by multiplets and high mitochondrial RNA expression that would be reflective of a high fraction of apoptotic or necrotic cells.

Quality assessment revealed that 77.2% (n = 74,922) of cells from OL and 71.9% (n = 73,814) of cells from BM passed the filtering process (**Table 1**).

### **Fluorescence-activated cell sorting (FACS)**

Immunophenotyping for plasma cells was performed on purified bone marrow samples from the bone marrow RS of the iliac crest as well as the aspirate from the OL. The following cell surface markers were used on fresh samples to identify plasma cells: CD38, CD138 and CD45 (37).

### **Fluorescence in situ hybridization (FISH)**

FISH analyses were performed on CD138-purified plasma cells counting at least 100 nuclei per sample and using probes for: 1q, 1p, 5q, 9 satellite III, del13q, 15 alpha satellite, del17p, t(4;14), t(11;14), t(14;16) and breakapart probes for IgH as well as MYC.

### **Single cell RNA sequencing (scRNA-seq)**

Single cell gene expression libraries were generated using the 10X Genomics platform, as described previously (9). In brief, cells were loaded into the Chromium Controller (10X Genomics, CA) where they were partitioned into nanoliter-scale Gel Beads-in-emulsion with a single barcode per cell. After reverse transcription, the resulting cDNA was amplified and used to generate libraries by enzymatic fragmentation, end-repair, a-tailing, adapter ligation, and PCR to add Illumina compatible sequencing adapters. The resulting libraries were evaluated on D1000 screentape using a TapeStation 4200 (Agilent Technologies, CA), and quantitated using Kapa Biosystems qPCR quantitation kit for Illumina (Illumina Inc., Ca). They were then pooled, denatured, and diluted to 350pM with 1% PhiX control library added. The resulting pool was then loaded into the appropriate NovaSeq Reagent cartridge and sequenced on a NovaSeq6000 following the manufacturer's recommended protocol (Illumina Inc., CA).

Cell Ranger (v3.1.0) was used to read alignment, filter, barcode and UMI counting. Analyses of scRNA-seq data were performed using the Seurat R toolkit (v3.2.2) for single cell genomics(22). The matched BM and OL samples from the same patient were merged together in the analysis. We filtered out the low quality or dying cells with more than 10% counts originating from the mitochondrial genes. The cells detected with less than 500 or more than 7,500 unique genes were also discarded to avoid empty droplet or multiplets. Data were normalized using the LogNormalization method from Seurat using a scale factor of 10,000. After feature selection and scaling of the normalized data, we performed PCA linear dimensional reduction. The first 30 PCs were used to construct the KNN graph and the Louvain algorithm

was performed for clustering the cells with a resolution parameter set to 0.2. We ran the UMAP method for the non-linear dimensional reduction to visualize the dataset. Gene expression profiles were annotated with publicly available datasets (Blueprint and ENCODE) using the R package SingleR (38). Cell cycle phases were scored using the list of cell cycle markers from Tirosh et al. that are preloaded in Seurat (18). Trajectory inference and pseudotime calculations with Monocle 3 (39) on Seurat objects were performed using Seurat Wrappers.

### **Copy number variations and phylogenetic trees from scRNA-seq**

CNVs were calculated using the inferCNV R package (18). Count matrices and cell annotations were extracted from the Seurat S4 object for every individual patient. Reference cell clusters were determined by the annotations from SingleR. Since SingleR cannot differentiate between malignant and non-malignant PC, the latter were determined by restricted expression of heavy and light chains. CNV prediction via hidden markov model (HMM) was performed at the level of subclusters instead of the entire sample by using the random trees method and setting the p-value to 0.05. Phylogenetic trees were generated using UPhyloPlot2 with arm length being proportional to the number of cells (19). Only subclusters accounting for at least 15% of total cells were visualized.

### **Integrated analysis of paired and longitudinal samples**

To identify clusters that are present in the OL and BM and to investigate differentially expressed genes between both conditions, we performed an integrated analysis as described previously (22). After identifying anchors between paired samples and integrating datasets for every individual patient, the standard Seurat workflow for clustering and visualization was performed. To get a broad overview on differences in gene expression between malignant OC from OL and BM, scatterplots for average gene expression were generated. Next, we identified differentially expressed genes between the two groups by using the Wilcoxon Rank Sum test. Results were visualized with the EnhancedVolcano R package (40). The same integrated analysis of anchored datasets was performed in the 3 patients with longitudinal samples collected after induction therapy.

### **Gene set enrichment analysis**

Gene set enrichment analysis (GSEA) was performed using the FGSEA R package (41) to investigate whether scRNA-seq data can be used to classify individual patients according to the molecular classification of MM from bulk GEP (15). The AddModuleScore() command from Seurat was used to calculate the average expression for each molecular subgroup and visualize results in a FeaturePlot.

## Whole exome sequencing (WES)

DNA was extracted from frozen plasma cells using kits according to manufacturer's instructions (DNeasy kit, Qiagen, Hilden, Germany) for bulk WES. Oral swabs (oragen×DNA, DNA genotek, Ontario, Canada) were collected for germline controls. After quality check (Qubit Fluorometric Quantification DNA and RNA Assay kits, Thermo Fisher, MA), samples were subjected to WES. SureSelect XT Low Input Target Enrichment System (Agilent Inc, CA) was used for individual exome capturing of each DNA sample. DNA was sheared using a Covaris S220 (Covaris Inc., MA) followed by end repair, P5 adaptor ligation, and 10 cycles of PCR to complete the P7 adapter. Unique dual-indexed libraries were purified with AMPureXP beads (Beckman Coulter, CA) and validated for appropriate size on a TapeStation 4200 DNA1000 screentape (Agilent Inc., CA). The purified library was then hybridized to the SureSelectXT Human All Exon V7 Capture library (Agilent Inc., CA). Afterwards, the hybridized regions were bound to streptavidin magnetic beads and washed to remove any non-specific bound products. Eluted library underwent a second 10 cycle PCR amplification to generate enough material for sequencing. Final libraries were purified, measured by TapeStation 4200 DNA1000 screentape, and quantitated using KAPA qPCR (KAPA Biosystems, Basel, Switzerland). Individual libraries were pooled in equimolar fashion at 2nM final concentration. Each pool was denatured and diluted to 350pM with 1% PhiX control library added. The resulting pool was then loaded into the appropriate NovaSeq Reagent cartridge and sequenced on a NovaSeq6000 following the manufacturer's recommended protocol (Illumina Inc., CA).

High quality paired-end reads passing Illumina RTA filter were initially processed against the NCBI human reference genome (GRCh37) using publicly available bioinformatic tools (42, 43) and Picard (<http://picard.sourceforge.net/>). Putative single nucleotide variants (SNVs) and insertions and deletions (indels) were identified by running variation detection module of Bambino (44). All putative SNVs were further filtered based on a standard set of criteria to remove the following common types of false calls: (1) the alternative allele was present in the matching normal sample and the contingency between the tumor and normal samples was not statistically significant; (2) the mutant alleles were only present in one strand and the strand bias was statistically significant; (3) the putative mutation occurred at a site with systematically dropped base quality scores; (4) the reads harboring the mutant allele were associated with poor mapping quality. Ambiguous cases were manually inspected to ensure accuracy. Putative indels were evaluated by a re-alignment process to filter out potential false calls introduced by unapparent germline events, mapping artifacts and homopolymer. The identified somatic mutations were compared to the public human germline databases including dbSNP (45), 1000 Genomes Project (46), National Heart, Lung, and Blood Institute's Exome Sequencing Project to further exclude remaining germline polymorphisms. All mutations were annotated using ANNOVAR (47) with NCBI RefSeq database.

## References

1. C. Röhlig, S. Knop, M. Bornhäuser, Multiple myeloma, *Lancet Lond. Engl.* **385**, 2197–2208 (2015).
2. B. A. Walker, E. M. Boyle, C. P. Wardell, A. Murison, D. B. Begum, N. M. Dahir, P. Z. Proszek, D. C. Johnson, M. F. Kaiser, L. Melchor, L. I. Aronson, M. Scales, C. Pawlyn, F. Mirabella, J. R. Jones, A. Brioli, A. Mikulasova, D. A. Cairns, W. M. Gregory, A. Quartilho, M. T. Drayson, N. Russell, G. Cook, G. H. Jackson, X. Leleu, F. E. Davies, G. J. Morgan, Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma, *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **33**, 3911–3920 (2015).
3. R. Kuiper, M. van Duin, M. H. van Vliet, A. Broijl, B. van der Holt, L. el Jarari, E. H. van Beers, G. Mulligan, H. Avet-Loiseau, W. M. Gregory, G. Morgan, H. Goldschmidt, H. M. Lokhorst, P. Sonneveld, Prediction of high- and low-risk multiple myeloma based on gene expression and the International Staging System, *Blood* **126**, 1996–2004 (2015).
4. O. Decaux, L. Lodé, F. Magrangeas, C. Charbonnel, W. Gouraud, P. Jézéquel, M. Attal, J.-L. Harousseau, P. Moreau, R. Bataille, L. Champion, H. Avet-Loiseau, S. Minvielle, Intergroupe Francophone du Myélome, Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myélome, *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **26**, 4798–4805 (2008).
5. D. Hose, T. Reme, T. Hielscher, J. Moreaux, T. Messner, A. Seckinger, A. Benner, J. D. Shaughnessy, B. Barlogie, Y. Zhou, J. Hillengass, U. Bertsch, K. Neben, T. Mohler, J. F. Rossi, A. Jauch, B. Klein, H. Goldschmidt, Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma, *Haematologica* **96**, 87–95 (2011).
6. J. D. Shaughnessy, F. Zhan, B. E. Burington, Y. Huang, S. Colla, I. Hanamura, J. P. Stewart, B. Kordsmeier, C. Randolph, D. R. Williams, Y. Xiao, H. Xu, J. Epstein, E. Anaissie, S. G. Krishna, M. Cottler-Fox, K. Hollmig, A. Mohiuddin, M. Pineda-Roman, G. Tricot, F. van Rhee, J. Sawyer, Y. Alsayed, R. Walker, M. Zangari, J. Crowley, B. Barlogie, A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1, *Blood* **109**, 2276–2284 (2007).
7. L. Rasche, S. S. Chavan, O. W. Stephens, P. H. Patel, R. Tytarenko, C. Ashby, M. Bauer, C. Stein, S. Deshpande, C. Wardell, T. Buzder, G. Molnar, M. Zangari, F. van Rhee, S. Thanendrarajan, C. Schinke, J. Epstein, F. E. Davies, B. A. Walker, T. Meissner, B. Barlogie, G. J. Morgan, N. Weinhold, Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing, *Nat. Commun.* **8**, 268 (2017).
8. P. Moreau, M. Attal, D. Caillot, M. Macro, L. Karlin, L. Garderet, T. Facon, L. Benboubker, M. Escoffre-Barbe, A.-M. Stoppa, K. Laribi, C. Hulin, A. Perrot, G. Marit, J.-R. Eveillard, F. Caillon, C. Bodet-Milin, B. Pegourie, V. Dorvaux, C. Chaletteix, K. Anderson, P. Richardson, N. C. Munshi, H. Avet-Loiseau, A. Gaultier, J.-M. Nguyen, B. Dupas, E. Frampas, F. Kraeber-Bodere, Prospective Evaluation of Magnetic Resonance Imaging and [18F]Fluorodeoxyglucose Positron Emission Tomography-Computed

- Tomography at Diagnosis and Before Maintenance Therapy in Symptomatic Patients With Multiple Myeloma Included in the IFM/DFCI 2009 Trial: Results of the IMAJEM Study, *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **35**, 2911–2918 (2017).
9. G. X. Y. Zheng, J. M. Terry, P. Belgrader, P. Ryzkin, Z. W. Bent, R. Wilson, S. B. Ziraldo, T. D. Wheeler, G. P. McDermott, J. Zhu, M. T. Gregory, J. Shuga, L. Montesclaros, J. G. Underwood, D. A. Masquelier, S. Y. Nishimura, M. Schnall-Levin, P. W. Wyatt, C. M. Hindson, R. Bharadwaj, A. Wong, K. D. Ness, L. W. Beppu, H. J. Deeg, C. McFarland, K. R. Loeb, W. J. Valente, N. G. Ericson, E. A. Stevens, J. P. Radich, T. S. Mikkelsen, B. J. Hindson, J. H. Bielas, Massively parallel digital transcriptional profiling of single cells, *Nat. Commun.* **8**, 1–12 (2017).
  10. G. Ledergor, A. Weiner, M. Zada, S.-Y. Wang, Y. C. Cohen, M. E. Gatt, N. Snir, H. Magen, M. Koren-Michowitz, K. Herzog-Tzarfati, H. Keren-Shaul, C. Bornstein, R. Rotkopf, I. Yofe, E. David, V. Yellapantula, S. Kay, M. Salai, D. Ben Yehuda, A. Nagler, L. Shvidel, A. Orr-Urtreger, K. B. Halpern, S. Itzkovitz, O. Landgren, J. San-Miguel, B. Paiva, J. J. Keats, E. Papaemmanuil, I. Avivi, G. I. Barbash, A. Tanay, I. Amit, Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma, *Nat. Med.* **24**, 1867–1876 (2018).
  11. S. V. Puram, I. Tirosh, A. S. Parikh, A. P. Patel, K. Yizhak, S. Gillespie, C. Rodman, C. L. Luo, E. A. Mroz, K. S. Emerick, D. G. Deschler, M. A. Varvares, R. Mylvaganam, O. Rozenblatt-Rosen, J. W. Rocco, W. C. Faquin, D. T. Lin, A. Regev, B. E. Bernstein, Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer, *Cell* **171**, 1611-1624.e24 (2017).
  12. J. G. Lohr, S. Kim, J. Gould, B. Knoechel, Y. Drier, M. J. Cotton, D. Gray, N. Birrer, B. Wong, G. Ha, C.-Z. Zhang, G. Guo, M. Meyerson, A. J. Yee, J. S. Boehm, N. Raje, T. R. Golub, Genetic interrogation of circulating multiple myeloma cells at single-cell resolution, *Sci. Transl. Med.* **8**, 363ra147 (2016).
  13. O. Zavidij, N. J. Haradhvala, T. H. Mouhieddine, R. Sklavenitis-Pistofidis, S. Cai, M. Reidy, M. Rahmat, A. Flaifel, B. Ferland, N. K. Su, M. P. Agius, J. Park, S. Manier, M. Bustoros, D. Huynh, M. Capelletti, B. Berrios, C.-J. Liu, M. X. He, E. Braggio, R. Fonseca, Y. E. Maruvka, J. L. Guerriero, M. Goldman, E. M. Van Allen, S. A. McCarroll, J. Azzi, G. Getz, I. M. Ghobrial, Single-cell RNA sequencing reveals compromised immune microenvironment in precursor stages of multiple myeloma, *Nat. Cancer* **1**, 493–506 (2020).
  14. T. Meißner, A. Seckinger, K. Hemminki, U. Bertsch, A. Foersti, M. Haenel, J. Duering, H. Salwender, H. Goldschmidt, G. J. Morgan, D. Hose, N. Weinhold, Profound impact of sample processing delay on gene expression of multiple myeloma plasma cells, *BMC Med. Genomics* **8** (2015), doi:10.1186/s12920-015-0161-6.
  15. F. Zhan, Y. Huang, S. Colla, J. P. Stewart, I. Hanamura, S. Gupta, J. Epstein, S. Yaccoby, J. Sawyer, B. Burington, E. Anaissie, K. Hollmig, M. Pineda-Roman, G. Tricot, F. van Rhee, R. Walker, M. Zangari, J. Crowley, B. Barlogie, J. D. Shaughnessy, The molecular classification of multiple myeloma, *Blood* **108**, 2020–2028 (2006).
  16. Y. Zhou, B. Barlogie, J. Shaughnessy, The molecular characterization and clinical management of multiple myeloma in the post-genome era, *Leukemia* **23**, 1941–1956 (2009).

17. M. Merz, A. Jauch, T. Hielscher, T. Bochtler, S. O. Schönland, A. Seckinger, D. Hose, U. Bertsch, K. Neben, M. S. Raab, J. Hillengass, H. Salwender, I. W. Blau, H.-W. Lindemann, I. G. H. Schmidt-Wolf, C. Scheid, M. Haenel, K. C. Weisel, H. Goldschmidt, Prognostic significance of cytogenetic heterogeneity in patients with newly diagnosed multiple myeloma, *Blood Adv.* **2**, 1–9 (2018).
18. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J.-R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. K. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A.-C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. V. Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jané-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq, *Science* **352**, 189–196 (2016).
19. M. A. Durante, D. A. Rodriguez, S. Kurtenbach, J. N. Kuznetsov, M. I. Sanchez, C. L. Decatur, H. Snyder, L. G. Feun, A. S. Livingstone, J. W. Harbour, Single-cell analysis reveals new evolutionary complexity in uveal melanoma, *Nat. Commun.* **11**, 496 (2020).
20. Z. Yao, R. Yaeger, V. S. Rodrik-Outmezguine, A. Tao, N. M. Torres, M. T. Chang, M. Drosten, H. Zhao, F. Cecchi, T. Hembrough, J. Michels, H. Baumert, L. Miles, N. M. Campbell, E. de Stanchina, D. B. Solit, M. Barbacid, B. S. Taylor, N. Rosen, Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS, *Nature* **548**, 234–238 (2017).
21. M. S. Raab, N. Lehnert, J. Xu, A. D. Ho, P. Schirmacher, H. Goldschmidt, M. Andrulis, Spatially divergent clonal evolution in multiple myeloma: overcoming resistance to BRAF inhibition, *Blood* **127**, 2155–2157 (2016).
22. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species, *Nat. Biotechnol.* **36**, 411–420 (2018).
23. H. van Andel, K. A. Kocemba, M. Spaargaren, S. T. Pals, Aberrant Wnt signaling in multiple myeloma: molecular mechanisms and targeting options, *Leukemia* **33**, 1063–1075 (2019).
24. E. Terpos, I. Ntanasis-Stathopoulos, M. Gavriatopoulou, M. A. Dimopoulos, Pathogenesis of bone disease in multiple myeloma: from bench to bedside, *Blood Cancer J.* **8**, 1–12 (2018).
25. C. V. Hedvat, R. L. Comenzo, J. Teruya-Feldstein, A. B. Olshen, S. A. Ely, K. Osman, Y. Zhang, N. Kalakonda, S. D. Nimer, Insights into extramedullary tumour cell growth revealed by expression profiling of human plasmacytomas and multiple myeloma, *Br. J. Haematol.* **122**, 728–744 (2003).
26. D. R. Carrasco, K. Sukhdeo, M. Protopopova, R. Sinha, M. Enos, D. E. Carrasco, M. Zheng, M. Mani, J. Henderson, G. S. Pinkus, N. Munshi, J. Horner, E. V. Ivanova, A. Protopopov, K. C. Anderson, G. Tonon, R. A. DePinho, The Differentiation and Stress Response Factor XBP-1 Drives Multiple Myeloma Pathogenesis, *Cancer Cell* **11**, 349–360 (2007).
27. Y. J. Shen, Y. Mishima, J. Shi, R. Sklavenitis-Pistofidis, R. A. Redd, M. Moschetta, S. Manier, A. M. Roccaro, A. Sacco, Y.-T. Tai, F. E. Mercier, Y. Kawano, N. K. Su, B. Berrios, J. Doench, D. E. Root, F. Michor, D. T. Scadden, I. M. Ghobrial, Progression signature underlies clonal evolution and dissemination of multiple myeloma, *Blood*, doi:10.1182/blood.2020005885.

28. B. Kong, C. W. Michalski, X. Hong, N. Valkovskaya, S. Rieder, I. Abiatari, S. Streit, M. Erkan, I. Esposito, H. Friess, J. Kleeff, AZGP1 is a tumor suppressor in pancreatic cancer inducing mesenchymal-to-epithelial transdifferentiation by inhibiting TGF- $\beta$ -mediated ERK signaling, *Oncogene* **29**, 5146–5158 (2010).
29. I. Riz, T. S. Hawley, R. G. Hawley, KLF4-SQSTM1/p62-associated prosurvival autophagy contributes to carfilzomib resistance in multiple myeloma models, *Oncotarget* **6**, 14814–14831 (2015).
30. Z. Wang, Y. P. Rong, M. H. Malone, M. C. Davis, F. Zhong, C. W. Distelhorst, Thioredoxin-interacting protein (txnip) is a glucocorticoid-regulated primary response gene involved in mediating glucocorticoid-induced apoptosis, *Oncogene* **25**, 1903–1913 (2006).
31. O. Decaux, M. Clément, F. Magrangeas, W. Gouraud, C. Charbonnel, L. Campion, H. A. Loiseau, S. Minvielle, Inhibition of mTORC1 activity by REDD1 induction in myeloma cells resistant to bortezomib cytotoxicity, *Cancer Sci.* **101**, 889–897 (2010).
32. S. Manier, J. Park, M. Capelletti, M. Bustoros, S. S. Freeman, G. Ha, J. Rhoades, C. J. Liu, D. Huynh, S. C. Reed, G. Gydush, K. Z. Salem, D. Rotem, C. Freymond, A. Yosef, A. Perilla-Glen, L. Garderet, E. M. Van Allen, S. Kumar, J. C. Love, G. Getz, V. A. Adalsteinsson, I. M. Ghobrial, Whole-exome sequencing of cell-free DNA and circulating tumor cells in multiple myeloma, *Nat. Commun.* **9**, 1691 (2018).
33. P. Storti, L. Agnelli, B. dalla Palma, K. Todoerti, V. Marchica, F. Accardi, G. Sammarelli, F. Deluca, D. Toscani, F. Costa, E. Vicario, G. Todaro, E. Martella, A. Neri, N. Giuliani, The transcriptomic profile of CD138<sup>+</sup> cells from patients with early progression from smoldering to active multiple myeloma remains substantially unchanged, *Haematologica* **104**, e465–e469 (2019).
34. Y. X. Zhu, E. Braggio, C.-X. Shi, L. A. Bruins, J. E. Schmidt, S. Van Wier, X.-B. Chang, C. C. Bjorklund, R. Fonseca, P. L. Bergsagel, R. Z. Orlowski, A. K. Stewart, Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide, *Blood* **118**, 4771–4779 (2011).
35. J. Hillengass, S. Usmani, S. V. Rajkumar, B. G. M. Durie, M.-V. Mateos, S. Lonial, C. Joao, K. C. Anderson, R. García-Sanz, E. R. Serra, J. Du, N. van de Donk, J. G. Berdeja, E. Terpos, E. Zamagni, R. A. Kyle, J. San Miguel, H. Goldschmidt, S. Giralt, S. Kumar, N. Raje, H. Ludwig, E. Ocio, R. Schots, H. Einsele, F. Schjesvold, W.-M. Chen, N. Abildgaard, B. C. Lipe, D. Dytfeld, B. M. Wirk, M. Drake, M. Cavo, J. J. Lahuerta, S. Lentzsch, International myeloma working group consensus recommendations on imaging in monoclonal plasma cell disorders, *Lancet Oncol.* **20**, e302–e312 (2019).
36. R. T. Swords, J. Anguita, R. A. Higgins, A. C. Yunes, M. Naski, S. Padmanabhan, K. R. Kelly, D. Mahalingam, T. Philbeck, L. Miller, T. A. Puga, F. J. Giles, M. C. Kinney, A. J. Brenner, A prospective randomised study of a rotary powered device (OnControl) for bone marrow aspiration and biopsy, *J. Clin. Pathol.* **64**, 809–813 (2011).
37. K. T. Soh, J. D. Tario, T. E. Hahn, J. Hillengass, P. L. McCarthy, P. K. Wallace, Methodological considerations for the high sensitivity detection of multiple myeloma measurable residual disease, *Cytometry B Clin. Cytom.* (2019), doi:10.1002/cyto.b.21862.
38. D. Aran, A. P. Looney, L. Liu, E. Wu, V. Fong, A. Hsu, S. Chak, R. P. Naikawadi, P. J. Wolters, A. R. Abate, A. J. Butte, M. Bhattacharya, Reference-based analysis of lung single-cell sequencing reveals a

- transitional profibrotic macrophage, *Nat. Immunol.* **20**, 163–172 (2019).
39. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis, *Nature* **566**, 496–502 (2019).
  40. K. Blighe, S. Rana, E. Turkes, B. Ostendorf, M. Lewis, *EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling* (Bioconductor version: Release (3.12), 2020; <https://bioconductor.org/packages/EnhancedVolcano/>).
  41. G. Korotkevich, V. Sukhov, A. Sergushichev, Fast gene set enrichment analysis, *bioRxiv*, 060012 (2019).
  42. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform, *Bioinformatics* **25**, 1754–1760 (2009).
  43. Q. Liu, Q. Hu, S. Yao, M. L. Kwan, J. M. Roh, H. Zhao, C. B. Ambrosone, L. H. Kushi, S. Liu, Q. Zhu, SeqSQC: A Bioconductor Package for Evaluating the Sample Quality of Next-generation Sequencing Data, *Genomics Proteomics Bioinformatics* **17**, 211–218 (2019).
  44. M. N. Edmonson, J. Zhang, C. Yan, R. P. Finney, D. M. Meerzaman, K. H. Buetow, Bambino: a variant detector and alignment viewer for next-generation sequencing data in the SAM/BAM format, *Bioinformatics* **27**, 865–866 (2011).
  45. S. T. Sherry, M.-H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, K. Sirotkin, dbSNP: the NCBI database of genetic variation, *Nucleic Acids Res.* **29**, 308–311 (2001).
  46. The 1000 Genomes Project Consortium, An integrated map of genetic variation from 1,092 human genomes, *Nature* **491**, 56–65 (2012).
  47. K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, *Nucleic Acids Res.* **38**, e164–e164 (2010).

## Declarations

### Financial support

This work is supported by a grant from the German Cancer Aid covering the salary of Maximilian Merz and an Institutional Educational Research Grant from Celgene Corporation. The Genomics, Biostatistics & Bioinformatics, and Flow & Image Cytometry Shared Resources are supported by the NIH/NCI Cancer Center Support Grant P30CA16056. Hemn Mohammadpour is funded by F32 CA239356. Black Swan Research Initiative of the International Myeloma Foundation.

Parts of this work were awarded with a Conquer Cancer Award at the 2020 ASCO meeting and presented at the ASH meeting in 2020.

### Conflict of interests:

PLM: Advisory Board/Consulting: BlueBird Biotech, Bristol-Myers Squibb, Celgene, Fate Therapeutics, Janssen, Juno, Karyopharm, Magenta Therapeutics, Sanofi, Takeda; Honoraria: BlueBird Biotech, Bristol-Myers Squibb, Celgene, Fate Therapeutics, Janssen, Juno, Karyopharm, Magenta Therapeutics, Sanofi, Takeda

JH: Advisory Boards/ Honoraria: Adaptive, Amgen, Bristol-Myers Squibb, Celgene, GlaxoSmithKline, Janssen, Oncotracker, Oncopeptide, Skyline, Takeda

MM: Advisory Boards/ Honoraria: Amgen, BMS, Celgene and Takeda

Institutional Research Support: Celgene

All other authors declare no conflict of interest.

### **Acknowledgement:**

The authors would like to thank our healthcare assistants Amy McCracken and Mahfoudh Saleh for their support during acquisition of the samples, preparing and caring for our patients. Special thanks to L. Shawn Matott for his continuous support regarding any problems with the High Performance Computing environment at Roswell Park.

### **Authors' Contributions**

**Conception and design:** Maximilian Merz, Almuth Maria Anni Merz, Philip McCarthy and Jens Hillengass

**Acquisition of data (acquired and managed patients, provided facilities, FISH, Flow cytometry, etc.):**

Maximilian Merz, Almuth Maria Anni Merz, Ahmed Belal, Ronald Alberico, AnneMarie W. Block, Cherie Rondeau, Kimberly Celotto, Hemn Mohammadpour, Megan Herr, Theresa Hahn, Paul K. Wallace, Joseph Tario, Jesse Luce, Sean T. Glenn, Prashant Singh, Philip L. McCarthy and Jens Hillengass

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):**

Maximilian Merz, Almuth Maria Anni Merz, Jie Wang, Lei Wei, Qiang Hu, Nicholas Hutson, Mehmet Samur, Nikhil Munshi, Song Liu, Philip McCarthy and Jens Hillengass

**Writing, review, and/or revision of the manuscript:** All authors

## **Supplementary Materials**

Table 1 is available in the Supplementary Files

**Supplemental Figure 1:**

## Intra-patient heterogeneity based on single plasma cell transcriptomes

Uniform manifold approximation and projection (UMAP) plots for each individual patient are shown as well as blended FeaturePlots for *STMN1* (red) and *TUBA1B* (green) expression. Combined FeaturePlots (yellow) showed the overlapping expression of both markers in the respective clusters (encircled in black). Heatmaps displaying the top 5 differentially expressed genes in the respective clusters are provided below the UMAs for every individual patient. Each row represents one gene (*STMN1* indicated by red arrow; *TUBA1B* indicated by green arrow) and each column represents one cluster (*STMN1+*/*TUBA1B+* clusters indicated by black arrows). The total number of displayed cells per cluster was limited to 50.

## Supplemental Figure 2:

**Prognostic significance of *STMN1* and *TUBA1B* expression in the MMRF CoMMpass dataset. (A)** Violin plots for *TUBA1B* and *STMN1* expression for each individual patient. *XBP1* was included in the top row as positive control to show the extend of the entire PC population. Only a small subset of malignant PC co-expressed *STMN1* and *TUBA1B*. **(B)** Progression-free (PFS) and overall survival (OS) for *STMN1* expression. **(C)** PFS and OS for *TUBA1B* expression. Patients with lower expression in blue. Patients with higher expression in red.

## Supplemental Figure 3:

### Intra-patient heterogeneity based on inferred copy number variations

Inferred CNVs from scRNA-seq data were used to visualize phylogenetic trees with UPhyloPlot2 with the length of each arm representing the relative size of the subclone. Only clones accounting for at least 15% of cells were visualized. In 8 of 10 patients we were able to identify subclones.

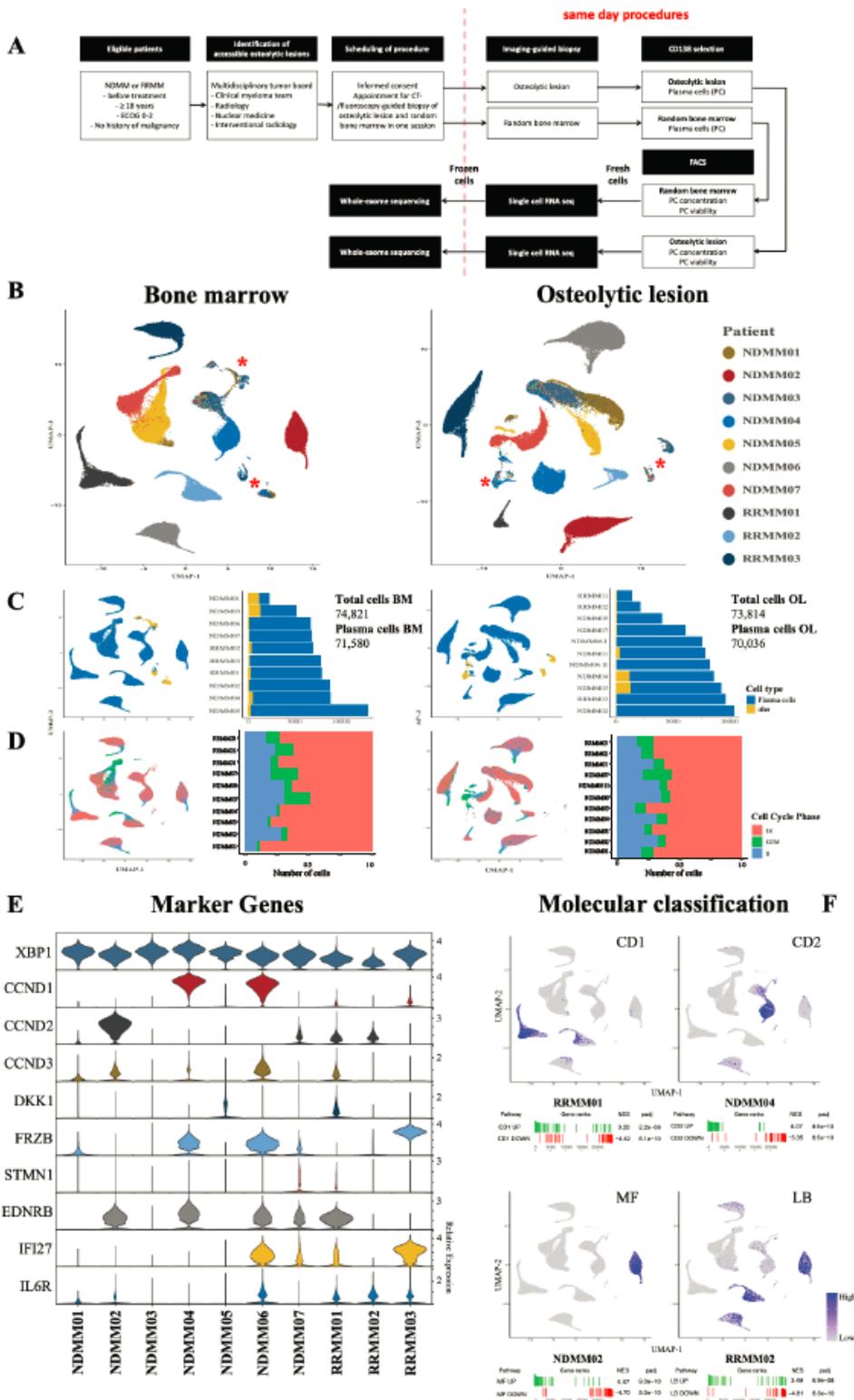
## Supplemental Table 1:

### Findings from whole-exome sequencing in the entire cohort

## Supplemental Table 2:

### Summary of differential expression analysis to identify subclusters of malignant plasma cells in individual patients

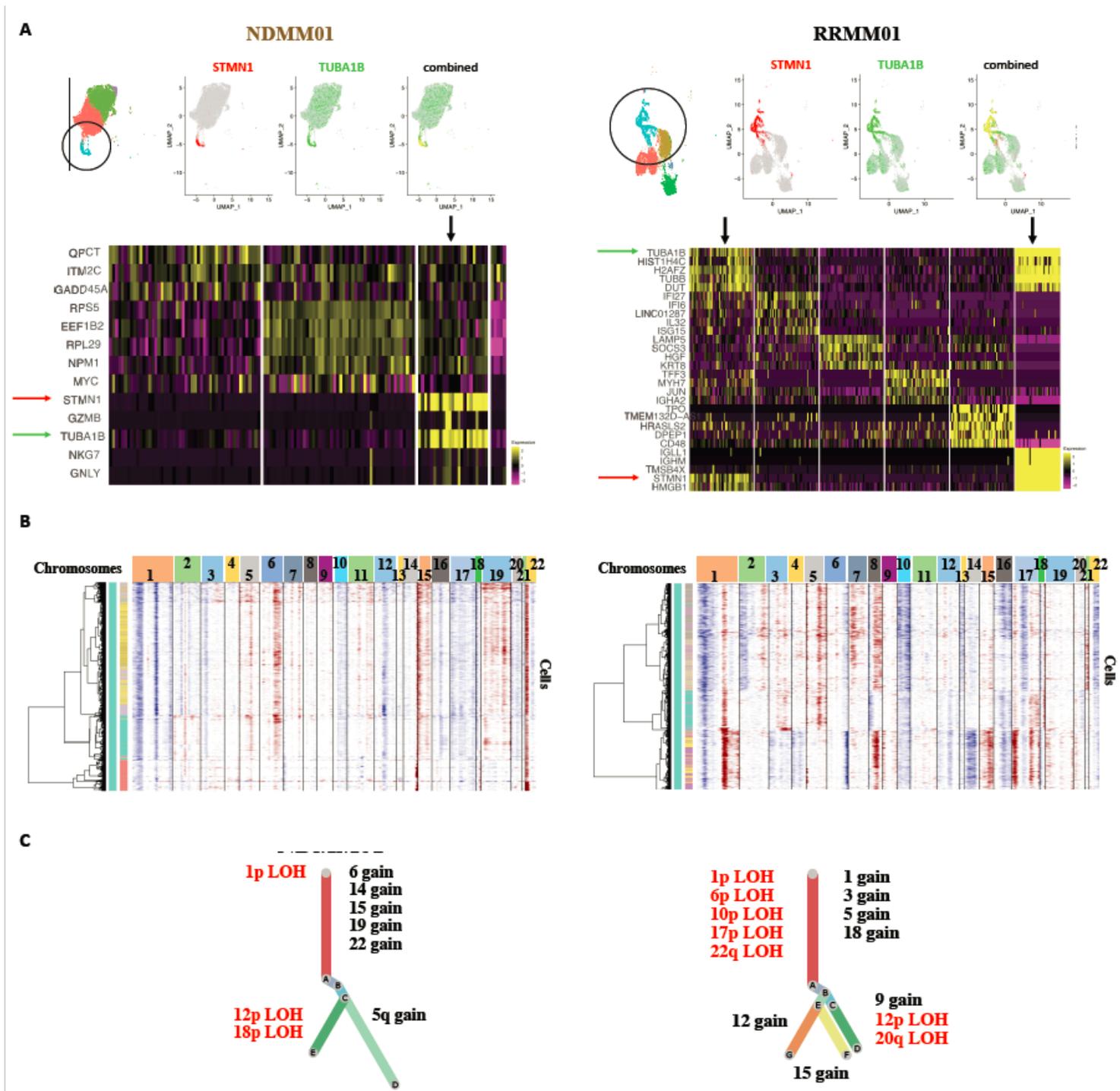
# Figures



**Figure 1**

Acquisition of viable plasma cells from imaging-guided biopsies of osteolytic lesions (OL) and bone marrow (BM) and comprehensive clustering of 148,746 cells shows inter-patient heterogeneity. (A) Summary of the workflow to obtain and isolate viable PC from BM and OL. Eligible patients with newly diagnosed (NDMM) or relapsed/refractory multiple myeloma (RRMM) were discussed in a multidisciplinary tumor board before the initiation of therapy. Patients consented to an imaging-guided

biopsy of an OL in addition to the routine random bone marrow aspirate from the iliac crest (random bone marrow samples, BM). To minimize changes in gene expression due to delayed sample processing, aspirates of the BM and OL were performed and analyzed on the same day as follows: 1) PC were isolated using a CD138 positive selection; 2) Fluorescence-activated cell sorting (FACS) was performed for quality control (QC); 3) cells were subjected to single cell RNA sequencing (scRNA-seq). The rest of the PCs were frozen at -80°C and later analyzed by whole exome sequencing (WES). (B) Uniform manifold approximation and projection (UMAP) plots for cells derived from the bone marrow (left side) and osteolytic lesion (right side) in 21 different location and 10 patients. Each patient is depicted by a specific color. PC from individual patients clustered together with overlap occurring only in regions of the few contaminating, non-malignant, non-PC as indicated by red asterisks (C) UMAPs after cell annotation with SingleR labelling PC (blue) and non-plasma cells (yellow). In total, 95% of captured and sequenced cells from BM and OL were PC. (D) UMAPs after removal of contaminating, non-PC and cell cycle scoring to rule out heterogeneity bias resulting from differences in cell cycle stage. No significant differences were observed in PC from BM and OL with the majority of malignant PC being captured in G1 phase. (E) Violin plots showing the identified marker genes for each individual patient. The width of the violin representing the number of cells expressing the respective gene with the relative expression being plotted on the y-axis. XBP1 was included in the top row as positive control to show the extend of the entire PC population. CCND1-3 were marker genes in patients with IgH-translocations (NDMM02/04/06) and homogeneously expressed in all PC. Other marker genes associated with bone disease (DKK1, FRZB), cytokine signaling (IFI27, IL6R) and EDNRB have been described in bulk gene expression profiling (GEP) studies. (F) Gene set enrichment analysis (GSEA) to classify patients according to the molecular subtypes published by the Arkansas group (14). Gene sets are available in the Molecular Signature Database (MSigDB). Examples are shown for the subgroups CD1 (RRMM01), CD2 (NDMM04), MF (NDMM02) and LB (RRMM02). GSEA for up- (green lines) and downregulated genes (red lines) were performed. The `AddModuleScore()` command from Seurat was used to calculate the average expression for the top 50 upregulated genes in the molecular subcategory and visualize findings in a FeaturePlot. Relative expression levels ranged from low (in grey) to high (in purple). Individual patients could be classified based on their single PC transcriptomes. NES = normalized enrichment score. Padj = adjusted p-value.



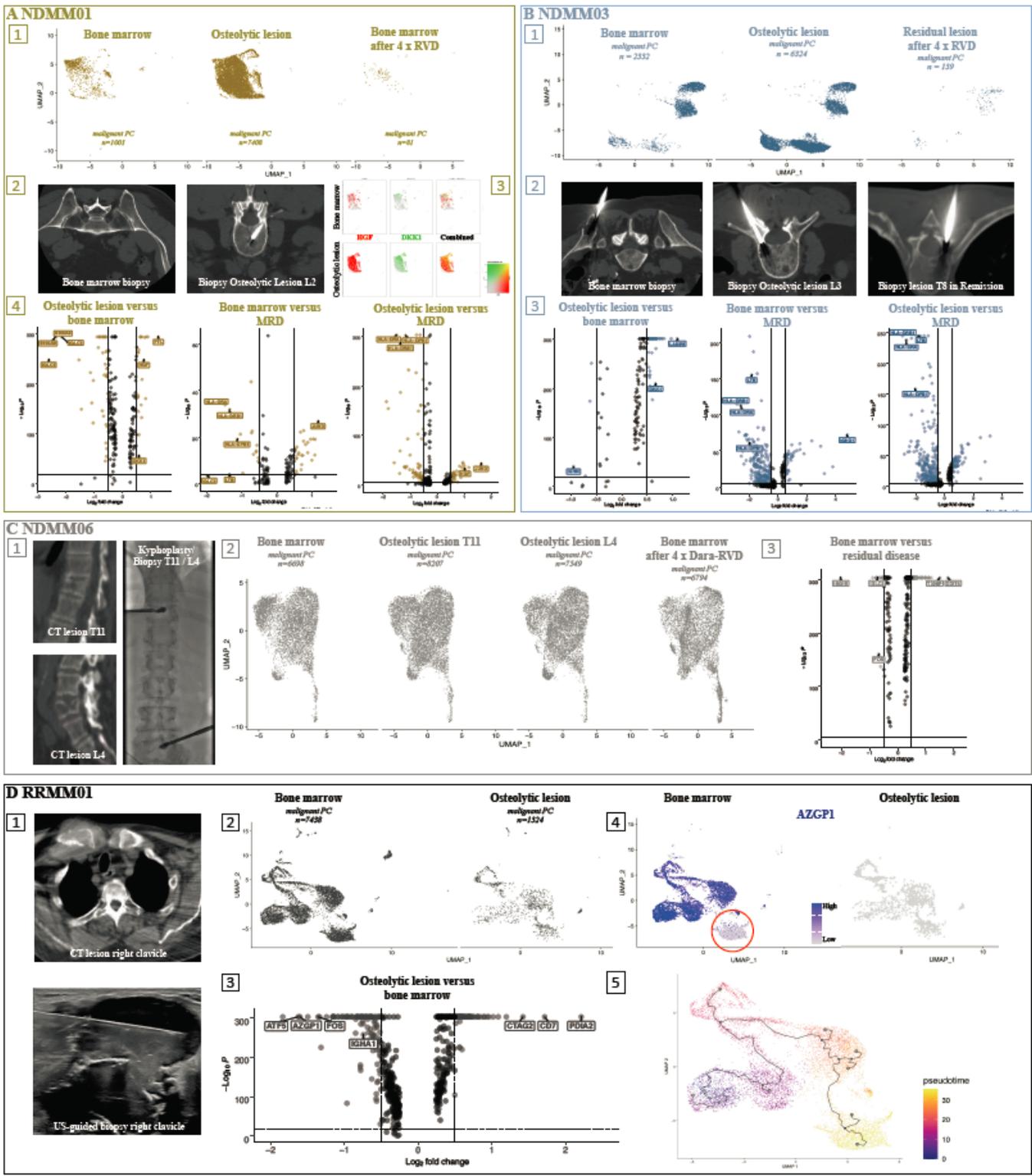
**Figure 2**

Intra-patient heterogeneity based on single plasma cell transcriptomes and inferred copy number variations. After the identification of inter-patient heterogeneity, we performed clustering and differential expression analysis for every individual patient to identify intra-patient heterogeneity. In 9 of 10 patients, we detected a subcluster characterized by the overexpression of genes encoding the microtubule-associated proteins STMN1 and TUBA1B. Both genes are associated with adverse outcome in the MMRF CoMMpass dataset (Supplemental Figure 2) and STMN1 is among the 15 genes that define high-risk disease in the IFM gene expression profiling score (4). (A) Uniform manifold approximation and

projection (UMAP) plots for each two patients (NDMM01 and RRMM01) are shown as well as blended FeaturePlots for STMN1 (red) and TUBA1B (green) expression. Combined FeaturePlots (yellow) showed the overlapping expression of both markers in the respective clusters (encircled in black). Heatmaps displaying the top 5 differentially expressed genes in the respective clusters are provided below the UMAPs for every individual patient. Each row represents one gene (STMN1 indicated by red arrow; TUBA1B indicated by green arrow) and each column represents one cluster (STMN1+/TUBA1B+-clusters indicated by black arrows). The total number of displayed cells per cluster was limited to 50. UMAPs and heatmaps for all patients are available in Supplemental Figure 1. To identify intra-patient heterogeneity based on copy number variations (CNV), we used the inferCNV R package. (B) shows two examples of CNVs calculated for patients NDMM01 and RRMM01. Each line represents one malignant PC and the columns reflect the respective chromosomal region. Gains are depicted in red, losses in blue. The predicted CNVs were used to visualize phylogenetic trees with UPhyloPlot2 as shown in (C) with the length of each arm representing the relative size of the subclone. Only clones accounting for at least 15% of cells were visualized. In 8 of 10 patients we were able to identify subclones. Phylogenetic trees created from inferCNV files from all individual patients are available in Supplemental Figure 3.



from single cell RNA sequencing. Second and fourth rows: Variant allele frequency (VAF) from WES. Results from plasma cells (PC) from the bone marrow (BM) are plotted to x-axis and from the osteolytic lesion (OL) to the y-axis for each individual patient. Jaccard indices were calculated to quantify the overlap between both samples based on WES. In WES scatter plots, effects of the mutations are delineated by different symbols (altering = circle, non-amino acid change = triangle, truncating = square). Only in patient RRMM01 with extramedullary disease (EMD), we found substantial differences as indicated by a Jaccard score of 0.2. Also average gene expression of PC from both location was highly correlated to each other. Nevertheless, outliers in both directions could be detected with scRNA-seq. (C) Integrating anchored datasets to identify genes that are differentially expressed in PC from OL and BM. First, samples from OL and BM were merged. Anchors between OL and BM samples were identified as described previously and datasets were integrated separately for each individual patient. Uniform manifold approximation and projection (UMAP) plots split by the origin of the PC are shown. The same clusters of malignant PC were identified in the OL and BM. (D) Differential expression analysis comparing OL and BM identified 1140 significantly up- and downregulated genes in PC from OL. While genes that have been associated with myeloma bone disease were upregulated (e.g. DKK1, HGF and TIMP1) in OL, JUN/FOS, DUSP1 and HBB were recurrently downregulated. (E) Top 15 up- and downregulated genes in PC from OL with regards to average log-fold change (avg\_logFC). (F) Gene set enrichment analysis demonstrated that genes connected to regular B-cell function were significantly downregulated in OL. NES = normalized enrichment score. Padj = adjusted p-value.



**Figure 4**

Single cell RNA sequencing links site-specific gene expression to the development of myeloma bone disease and identifies transcriptional changes in minimal residual disease after therapy. (A1) Uniform manifold approximation and projection (UMAP) plots split by the origin of plasma cells (PC) for patient NDMM01. After 4 cycles RVD, 81 malignant PC were isolated from the bone marrow and sequenced. (A2) CT images from the bone marrow (left side) and imaging-guided biopsy of L2 (right side). (A3) Blended

FeaturePlots showing expression levels of HGF (red) and DKK1 (green) or combined (yellow) in the osteolytic lesion (OL) and bone marrow (BM). It can be appreciated that the OL harbored more PC with higher expression levels compared to the BM. (A4) Volcano plots showing differentially expressed genes between OL and BM (left), BM and minimal residual disease (MRD, middle) as well as OL and MRD (right). (B2) UMAP plots split by the origin of PC for patient NDMM03. After 4 cycles RVD, 139 malignant PC were isolated and sequenced after imaging-guided biopsy of a residual MRI lesion in T8. (B2) CT images from the bone marrow (left side) imaging-guided biopsy of L3 (middle) and imaging-guided biopsy of T8 in remission (right side). (B3) Volcano plots showing differentially expressed genes between OL and BM (left), BM and MRD (middle) as well as OL and MRD (right). In both patients with MRD-positive disease we found human leucocyte antigen class II genes to be upregulated after therapy. (C1) CT images of fractured osteolytic lesions in T11 and L4 (left side) as well as biopsy and kyphoplasty under fluoroscopy of the respective lesions (right side) in patient NDMM06. (C2) UMAP plots split by the origin of PC. Almost same numbers of cells were captured from the bone marrow and both osteolytic lesions at primary diagnosis as well as the bone marrow after 4 cycles daratumumab-RVD. (C3) Volcano plot showing differentially expressed between initial bone marrow samples and in partial response (PR). ISG15 has been associated with carfilzomib-resistance and was upregulated in PC after therapy, while TXNIP and DDIT4 are linked to steroid-induced apoptosis. (D1) CT image of an OL in the right clavicle with an adjacent extramedullary tumor (top) and the respective ultrasound-guided biopsy (bottom) in patient RRMM01. (D2) UMAPs split by the origin of the cells. (D3) Volcano plot showing differentially expressed gene between OL and BM. (D4) The majority of malignant PC in the BM showed higher expression values auf AZGP1 compared to the OL. However, a small population with absent expression was detected in the BM (red circle). (D5) Trajectory analysis revealed that the respective cluster occurred later in the developmental process suggesting it gave rise to the distant OL. In Volcano plots, horizontal lines represent the p-value cutoff (10<sup>-10</sup>), vertical lines represent the log<sub>2</sub>-fold change cutoff (0.5).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.xlsx](#)
- [Supplementaltable1.xls](#)
- [Supplementaltable2.xlsx](#)
- [SupplementalFigures.pdf](#)