The Spatio-Temporal Evolution of Multiple Myeloma from Baseline to Relapse-Refractory States

Niels Weinhold (✉ nweinh80@gmail.com)

Department of Internal Medicine

Leo Rasche

University Hospital Wuerzburg

Carolina Schinke

University of Arkansas for Medical Sciences <https://orcid.org/0000-0002-2699-1741>

Francesco Maura

Sylvester Comprehensive Cancer Center, University of Miami <https://orcid.org/0000-0002-5017-1620>

Michael Bauer

Myeloma Center, University of Arkansas for Medical Sciences, Little Rock, AR <https://orcid.org/0000-0003-2498-7295>

Cody Ashby

Myeloma Center, University of Arkansas for Medical Sciences, Little Rock, AR <https://orcid.org/0000-0002-9361-0283>

Shayu Deshpande

University of Arkansas for Medical Sciences

Alexandra Poos

University Hospital of Heidelberg

Maurizio Zangari

Myeloma Center, University of Arkansas for Medical Sciences, Little Rock, AR

Sharmilan Thanendrarajan

University of Arkansas for Medical Sciences

Faith Davies

University of Arkansas for Medical Sciences

Brian Walker

Indiana University <https://orcid.org/0000-0002-8615-6254>

Bart Barlogie

Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY

Ola Landgren

Sylvester Comprehensive Cancer Center, University of Miami <https://orcid.org/0000-0001-6485-4839>

Gareth Morgan

NYU Perlmutter Cancer Center <https://orcid.org/0000-0002-4271-6360>

Frits van Rhee

Article

Keywords:

Posted Date: January 21st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1177985/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 August 3rd, 2022. See the published version at <https://doi.org/10.1038/s41467-022-32145-y>.

The Spatio-Temporal Evolution of Multiple Myeloma from Baseline to Relapse-Refractory States

Rasche, L^{1,2,3}, Schinke, C¹, Maura, F.⁴, Bauer, M¹, Ashby, C¹, Deshpande, S¹, Poos, A⁵, Zangari, M¹, Thanendrarajan, S¹, Davies, F.E.⁶, Walker, B.A.⁷, Barlogie B¹, Landgren O⁴, Morgan, GJ⁶, van Rhee, F.¹, Weinhold, N^{1,5}

¹ Myeloma Center, University of Arkansas for Medical Sciences, Little Rock, AR, USA

² Department of Internal Medicine 2, University Hospital of Würzburg, Würzburg, Germany

³ Mildred Scheel Early Career Center (MSNZ), University Hospital of Würzburg, Würzburg, Germany

⁴ Myeloma Program, Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL, USA

⁵ Department of Internal Medicine V, University Hospital of Heidelberg, Heidelberg, Germany

⁶ Perlmutter Cancer Center, New York University Langone Health, New York, NY, USA

⁷ Division of Hematology Oncology, Indiana University, Indianapolis, IN, USA

Corresponding Author:

Niels Weinhold, Department of Internal Medicine V, University Hospital of Heidelberg, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany

email: niels.weinhold@med.uni-heidelberg.de, phone: +496221564158

Running head: Spatio-temporal tumor evolution in myeloma

Competing Interests: The authors declare no conflict of interest.

Funding: This work was supported by P01 CA 55819 from the National Cancer Institute. LR was funded by the Deutsche Krebshilfe via the MSNZ program Würzburg, and the Interdisciplinary Center for Clinical Research Würzburg (IZKF). NW was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM125503, FM was supported by the American Society of Hematology and by Sylvester Comprehensive Cancer Center NCI Core Grant (P30 CA 240139). GJM was supported by the Leukemia and Lymphoma Society.

ABSTRACT

Deciphering the evolution of tumor cells in the whole bone marrow (BM) cavity is key to inform curative strategies in multiple myeloma. Here, we performed spatial-longitudinal sequencing, including 144 samples collected from 25 patients during up to 14 years. Applying imaging-guided sampling we observed three evolution patterns, including relapse driven by a single tumor cell, competing relapse sub-clones and a novel pattern characterized by unique treatment-resistant sub-clones at distinct BM locations. Our study highlights focal bone lesions (FL) as hotspots of tumor evolution and provides possible explanations for their negative prognostic impact. We show a close phylogenetic relationship between baseline FL and relapse sub-clones. In patients with ≥ 3 FDG-PET-positive FL at diagnosis, relapse was driven by multiple distinct sub-clones, whereas in other patients, a single-cell expansion was typically seen ($p < 0.01$). We observed resistant sub-clones that can be hidden over years, and epigenetic modifier mutations as a new class of relapse aberrations. In conclusion, imaging can enhance tumor evolution studies in myeloma. The sophisticated evolution mechanisms and presence of single cells that can drive relapse even after 10 years of remission suggest that a prerequisite for curative therapies in myeloma would be to overcome not only tumor heterogeneity but also dormancy.

INTRODUCTION

Deciphering the evolutionary trajectories shaping the clonal architecture of Multiple Myeloma (MM) from disease initiation to relapsed-refractory disease is the key to inform novel curative treatment strategies. Recent longitudinal studies in MM have identified a number of genetic aberrations, which are associated with increased clonal fitness and can lead to relapse. These aberrations frequently arise through branching pathways, especially in patients relapsing from deep responses ¹⁻⁷. However, the origin of relapse sub-clones as well as the extent of clonal diversity in treatment-resistant disease remain largely unknown.

Crucial hints to the origin of MM cells, which are the source of relapse, comes from functional imaging studies, which have identified nodular plasma cell (PC) accumulations in the remission bone marrow (BM) of some patients with undetectable minimal residual disease (MRD) at the iliac crest ⁸⁻¹⁰. Genomic sequencing of these nodules, which are called focal lesions, showed site-specific enrichment of aberrations associated with relapse ¹¹. While these observations support focal lesions as being the origin of resistant disease, and providing one possible explanation for their negative prognostic impact ¹²⁻¹⁴, the evolutionary pathways undergone by focal lesion sub-clones after treatment have not been defined.

Clonal diversity provides the fuel for drug resistance and treatment failure. Since current MM therapies induce deep responses, thereby potentially eradicating at least some sub-clones, it would be predicted that diversity decreases over time. According to the observations by Keats et al. ¹⁵, a number of relapse sub-clones compete for access to the BM niche. In contrast to this model, the recent description of melphalan-induced mutational signatures in tumor cells derived from multiple sites suggests that a single clonal initiating cell can drive MM relapse ^{16,17}. Consequently, the exact model of how clonal diversity is generated in MM during therapy and its impact on resistance remain controversial. This deficiency could be remedied by the addition of spatial and temporal data on clonal structure.

Here, we analyzed a unique set of longitudinal multi-region sequencing data together with matched clinical annotations, including a total of 144 tumor samples from 25 MM patients (**Fig. 1, Suppl. Table 1**). By tracking resistant clones in time and space we dissect the sophisticated spatio-temporal evolutionary pathways that are active in MM during treatment and demonstrate the power of combining functional imaging and tumor sequencing to delineate the most complete picture of sub-clonal evolution in MM available to date.

RESULTS

Sub-clonal spatial heterogeneity is linked to the composition of focal lesions

Recently, we characterized focal lesions as hotspots of regional tumor evolution in MM ¹¹ but the extent of spatial genomic heterogeneity outside of the focal lesions in the diffuse BM infiltrate remains unknown. From a clinical point of view, this information is important, since in daily clinical practice the left or right iliac crest are usually randomly chosen for diagnostic BM biopsies. To fill this gap in knowledge, we analyzed paired left and right iliac crest samples, which had been collected from 9 patients within a few days of each other at baseline.

We found a high degree of homogeneity in these paired samples with a median of 98% (range 96-100%) of mutations being shared, and in addition there were only minor sub-clonal differences between the two sites (**Fig. 2a and b**). The homogeneous sub-clonal mix at the iliac crest is an observation that is reassuring, supporting that a second independent random biopsy usually does not provide additional information. In contrast, we observed frequent major unshared mutations in paired baseline iliac crest and focal lesion samples with a significantly lower proportion of shared mutations (median 79% (62-96%), $p=0.001$, **Fig. 2a and c**). The picture seen at the iliac crest barely changed during therapy. Although we found shared-diff mutations (detectable in both samples but $>3\times$ CCF difference, see methods), there were no unshared major mutations and a median of 95% (92-98%) of mutations were shared between paired iliac crest samples (**Fig. 2b**). In paired iliac crest/focal lesion or focal lesion/focal lesion pairs, however, there was significant genomic heterogeneity at a similar level to that seen at baseline (**Fig. 2c**).

The largest level of genomic heterogeneity was seen, when we compared the first to the last available sample per patient, collected after a median of 6 treatment lines (range: 1-15) and 5.4 (1.5-13.2) years. On average, 41% (0-82%) of mutations were major in one sample but not detectable in the other (**Fig. 2d**). The spatio-temporal changes affecting known myeloma drivers are shown in **Fig. 2e**. Despite these huge changes in clonal composition, baseline and relapse were clonally related, even in patients with more than 10 years of follow-up, with the first and last dominating sub-clone having the same most recent common ancestor in all cases. Considering just the number of sub-clones per patient and sample, these changes would not be apparent. We observed on average 3 sub-clones both at baseline and after therapy (**Suppl. Fig. 1**). However, taking all time points into consideration, the total number of sub-clones per patient was significantly higher with an average of 5 (range 2-11, $p<0.01$). Mock oncogenetic trees for all patients are shown in **Suppl. Fig 2-27**.

Taken together, this data suggests that significant spatial genomic heterogeneity is associated with focal lesions, while the clonal composition of the diffuse MM infiltrate in the pelvis tends to be more homogeneous both at baseline and at subsequent distinct time points post treatment. Further, we show similar numbers of sub-clones at baseline and after therapy but major BM-wide changes in sub-clonal genetic composition in longitudinal analyses.

The relationship between focal lesion clones and resistant disease

A major implication of our initial findings is that pre-existing clones derived from focal lesions contribute to changes in clonal composition observed post treatment. To address the hypothesis, we analyzed longitudinal genetic data from patients, for whom paired iliac crest/focal lesion specimens at baseline and BM samples post treatment were available. In the majority of these cases ($n=7/12$), we observed a close relationship between the baseline clonal composition of the focal lesion and the relapse clone(s), with the major focal lesion sub-clone either being the precursor or sharing a phylogenetic branch with the relapse clone(s) (**Fig. 3a**). Except for one of these patients, the clonal branch leading to relapse was also detectable at the iliac crest but at a minor sub-clonal level, supporting the results of a recent baseline-relapse study¹⁸ (**Fig. 3b**).

Taken together, the data suggests that the precursors of resistant disease can frequently be found as major clones in focal lesions, supporting a link between the presence of focal lesion clones and the emergence of evolved clones at relapse.

Single clone expansion vs. multiple surviving sub-clones

Despite the close phylogenetic relationship between focal lesion clones and relapsed resistant disease states, we did not detect the unique relapse clone as a major sub-clone in the baseline focal lesion sample with only a few patients showing the selection of a pre-existing minor sub-clone at the iliac crest (**Fig. 3b**). One possible explanation for this lack of a tight association is that relapse is driven by a very small unique sub-clone or even a single cell in the majority of patients, as has recently been shown in the “single-cell expansion” model ¹⁶, making its detection at baseline highly challenging.

To test this hypothesis, we analyzed 21 patients for whom baseline and at least 2 follow-up samples were available, and performed a longitudinal analysis of the clonal architecture. All patients received melphalan (**Suppl. Table 1**), which, with some limitations (see methods), allowed us to use the mutational signature SBS-MM1 (melphalan) as a barcode for single-cell tracking. In one third (7/21) of patients, we detected only one expanded sub-clone at relapse, and in 5 of them the signature SBS-MM1 was detectable in mutations, which were common to all relapse cells (**Fig. 4a, Suppl. Table 2 & 3**). Since the underlying mutations are only detectable through bulk sequencing when one single cell expands, these observations support the existence of a single-cell expansion mechanism. However, in 10 patients we found evidence that more than one sub-clone (typically two) from the primary untreated tumor was involved in seeding of relapse cells. We also noted changes in sub-clonal proportions in 8 of these examples, reminiscent of the clonal competition model ¹⁵ (**Fig. 4b**).

In the remaining 4 patients we identified a special type of multi-clone resistant disease, where unique sub-clones from different branches are seen at distinct locations, in contrast to the classical clonal competition model, where multiple genetically distinct sub-clones co-exist at the same location. This “alternating spatial clonal dominance” pattern is illustrated in **Suppl. Fig. 28**. Thus, there are three distinct patterns of sub-clonal evolution in MM that occur over time in response to cycles of remission and relapse, which are 1) single-cell expansions and sweeps, 2) coexisting expanding (and competing) sub-clones, and 3) site-unique expansions of distinct sub-clones with the main difference between the first and the third pattern being the anatomical location of sub-clones.

The link between evolutionary pattern and clinical features

We investigated whether the three evolutionary patterns were associated with specific clinical features. Due to limited sample size, we restricted the analysis to the risk status defined by the GEP70 and the number of focal lesions on PET-CT at baseline, which are independent prognostic factors ^{12,19}. Furthermore, we included *best response during first-line* therapy as another parameter, as this was previously shown to be associated with branching evolution in MM ⁵. Risk status at baseline was not associated with the evolutionary patterns seen at relapse. In contrast, achieving CR was a significant determinant of the type of MM evolution. All patients in the single-cell expansion and alternating spatial clonal dominance categories had relapsed from CR but more than half of the patients with co-existing sub-clones (i.e. in the clonal competition model) had only achieved PR (n=3) or VGPR (n=3) as their best response (p<0.05, fisher’s exact test, **Fig. 5**). A strong association with the evolutionary patterns was seen for the number of focal lesions detected at baseline (**Fig. 5**). In all patients with ≥3 PET-positive focal lesions, an imaging pattern associated with poor outcome ^{12,14}, relapse was driven by multiple surviving sub-clones, whereas in patients with <3 PET-positive focal lesions a single-cell expansion was typically seen (7/10 patients, p<0.01, fisher’s exact test). These

observations support the relevance of the single sub-clone expansion model for patients with none or only a limited number of focal lesions at baseline. The data is also consistent with an increased number of focal lesions at baseline being associated with greater levels of spatial clonal heterogeneity which increase the likelihood of multiple sub-clones being able to survive treatment and lead to relapse.

Mixture of evolution patterns

The three patterns of evolution, which we describe herein, were not mutually exclusive with their components occurring either sequentially or in parallel in the majority of patients. On the one hand, all except one patient from the “single-cell (clone) expansion” group showed further clonal evolution and had evidence of sub-clonal competition during the course of the disease (**Fig. 4a**). On the other hand, 11 of the 14 patients with multiple surviving sub-clones showed the SBS-MM1 melphalan exposure signature in at least one sub-clone (**Suppl. Table 3**). Intriguingly, we observed a single-cell expansion in a focal lesion at the right sacrum and concomitant selection of a pre-existing clone without evidence of SBS-MM1 at the right iliac crest in patient #9 (**Suppl. Fig. 29**).

The best example of a mixed pattern of clonal evolution was seen in patient #12 (**Fig. 6**). This patient had a homogeneous clonal composition at the left and right iliac crest at baseline. During their follow-up over 14 years, three resistant clones were detected, which had the same pre-existing precursor but which were comprised of one of three unique clonal branches driven by three independent single-cell expansions. While the first clone dominated in focal lesions, the second was seen primarily at the iliac crest. The third clone, defined by a del(17p) among others, emerged 13 years after diagnosis following salvage ASCT and was not seen in any of the prior samples, suggesting that it had been in a dormant state for years. This clone dominated at the iliac crest at the last time point.

The frequency of clonal sweeps

Selective clonal sweeps, which refer to an expansion of a new clone and suppression of previously dominant clones, are a crucial component of MM evolution^{2,3,15}. Sweeps are clinically relevant, since the respective sub-clones could have new drug susceptibilities and resistance profiles. However, the frequency as well as the clinical context in which clonal sweeps occur remain poorly understood. To get deeper insights into this process, we analyzed 14 patients with at least two follow-up iliac crest specimens and observed no, 1, 2 and 3 sweeps in 1, 5, 6, and 2 patients, respectively, during the observation period. Clonal sweeps were not seen exclusively in patients relapsing from CR, but rather were seen in relapse from all response levels including the presence of major clonal changes in patients with progressive disease (**Suppl. Table 4**). We assume that a high proliferation rate and accelerated seeding of a more resistant sub-clone^{3,16} masked the extinction of previously dominant sub-clones in these patients. With each clonal sweep the dominant sub-clone(s) became more and more advanced as defined by the acquisition of novel mutations and/or CNAs. Thus, clonal sweeps are a common evolutionary event in MM during which the evolutionary fittest clones come to dominate in the BM cavity, raising the question as to the underlying drivers of fitness.

Spatial-temporal parallel evolution

The identification of parallel evolution by the analysis of spatial heterogeneity can provide a tool to identify the key drivers of relapse²⁰. Parallel evolution refers to a pattern in which sub-

clones evolve through independent mutations, targeting the same genes or the same molecular pathways ²⁰. Focusing on *NRAS*-, *KRAS*-, and/or *TP53*-mutated sub-clones, we observed this type of evolution in 5 patients in our data. Patient #13 presented with two unique *del(17p)/TP53^{mut}* double-hit events at different locations at baseline and another *TP53* double hit at relapse (**Suppl. Fig. 28**). Sub-clones with almost identical driver mutations were seen in patient #24, where the first branch, which dominated at baseline, had concomitant mutations in *NRAS*, *DIS3*, and *del(17p)*, while the second branch, which emerged at relapse, had mutations in *KRAS*, *TP53*, and *DIS3*. Both branches were associated with single-cell expansions, highlighting the commonality between the sub-clones. Stable evolution of two co-existing *KRAS*-mutated sub-clones was observed in patient #2 at the iliac crest. Yet, this patient developed an interval focal lesion at the left ischium that was related to one of the *KRAS*-branches but showed a site-unique *IKBKB* mutation and *del(2q37.2-q37.3)*, illustrating evolutionary changes in a patient with seemingly stable disease.

We extended the analysis to examine other genes recently implicated as MM drivers ²¹ and strikingly, we observed multiple mutations in epigenetic modifiers of the lysine demethylase (KDM)/methyltransferase (KMT) family in patient #12, who presented with three clonal branches, *KDM4B* and *KMT2D* were each affected in a branch (**Fig. 6**). In total, we found mutations in these two classes of epigenetic modifiers in 10/25 patients (**Suppl. Table 5**). While one patient showed *KDM* mutations at baseline, in all 10 patients mutations in *KDM* and/or *KMT* genes emerged during treatment. Since they appeared in resistant disease and evolved in parallel, mutations in the *KDM/KMT* gene family are promising candidates as novel relapse drivers in MM, further supporting the role of epigenetic modifiers in the evolution of high-risk disease ²².

Together, our analysis demonstrates for the first time spatially separated parallel evolution and supports mutations in epigenetic modifiers as promising candidates for drivers of MM relapse.

DISCUSSION

In this first longitudinal multi-region sequencing study of MM, we provide the most detailed picture of MM evolutionary pathways developing during therapy to date. Patients included in the study were extensively treated in Total Therapy trials or similar protocols and were salvaged with multi-agent therapies including a range of novel drugs available at that time. Due to the long follow-up time, we could include all molecular subgroups and even patients with long-lasting remissions of up to 10 years during frontline therapy.

The distinct genetic makeup of focal lesions¹¹, which was confirmed in this study, and their prognostic impact are consistent with them having a significant role in MM progression. As a novel observation we show a close relationship between baseline focal lesion and relapse clones. Yet, it was challenging to track down the site and nature of the unique pre-existing relapse clone at disease presentation, which could be due to the relatively low number of BM samples and the limited sensitivity of WES. However, we think that the mutagenic impact of treatment and recent advances in the analysis of mutational signatures provide better explanations. MM cells usually acquire new mutations when exposed to alkylating agents ²³ and in our set ~1/3 of new mutations could be linked to melphalan. As a result, clonal changes are generally inevitable during exposure to multiagent chemotherapy. Second, and even more important, single-cell expansions were often seen at relapse, in line with recent predictions

based on samples from post mortem cases¹⁶. As these single cells could be from anywhere in the BM, even the most sensitive technologies would not be able to detect them, especially if spatial heterogeneity is not considered.

Given the nature of the MM clonal architecture, which can include site-specific sub-clones, it is difficult to capture the full extent of clonal diversity in MM. It seems that MM evolution is driven by a limited number of major sub-clones that generate the majority of diversity at relapse. However, despite this speculation we reported previously on a patient with unique sub-clones present in four different focal lesions¹¹. Notably, this patient had >100 focal lesions, suggesting a tremendously higher amount of genomic heterogeneity than is seen in the current limited set of samples. The same sub-clonal diversity likely holds true for treated patients, where we observed the emergence of unique sub-clones after multiple relapses and >10 years of treatment, clearly supporting a clonal composition which is dominated by clones with the highest proliferation rate and potential for adaptation. As a result of this observation, it is likely that relapse is associated with hidden sub-clonal heterogeneity. This complex sub-clonal architecture is further supported by recent longitudinal studies which showed emergence of previously undetectable sub-clones in relapsed-refractory MM patients treated with IMiDs, BRAF inhibitors or anti-BCMA CAR T cells^{5,6,24,25}.

From a therapeutic perspective it is important to understand why some relapsed patients present with multiple detectable sub-clones whereas in others relapse is driven by just one single clonal expansion. Here we demonstrate that the depth of response is associated with the number of (detectable) resistant sub-clones. While single sub-clone expansions were only seen in patients achieving CR, all patients relapsing from less deep responses had multiple detectable surviving sub-clones. A further important variable we describe for the first time is that the likelihood of multiple sub-clones surviving treatment and leading to progression is significantly increased in patients with multiple focal lesions at baseline. This not only demonstrates the value of incorporating imaging data together with molecular analyses but also provides a biological explanation for the negative prognostic impact of focal lesion number and size^{12,13,26}. Apparently, multiple unique advanced sub-clones at different BM locations increase the chance of there being multiple resistant sub-clones that can mediate a short time to progression.

Our observation further underscores the need for functional imaging at diagnosis and for response assessment^{9,27,28}. For instance, in some patients multiple surviving sub-clones did not coexist at the same location, as described by Keats *et al.*¹⁵, but dominated at different locations in the BM; a pattern which we called “alternating spatial clonal dominance”. This pattern can also be seen in a subset of newly diagnosed patients¹¹. We speculate that the potential of tumor cells to circulate via the peripheral blood is restricted in these patients but appreciate that the underlying mechanism remains elusive. As a variation of this unique pattern, we observed patients with spatial clonal dominance in which there was some exchange between sites.

Critically examining the sequencing data, we only considered mutations, which were seen in $\geq 20\%$ of MM cells in at least one of the patient’s samples in order to avoid an overestimation of heterogeneity. As a result, the study was focused on partially expanded sub-clones and thus rather underestimated the full extent of diversity. Hence, our approach could be enhanced in the future by using single-cell methods^{29–31}. Yet, it is important to keep in mind that MM cells may be subject to Muller’s ratchet, a phenomenon characterized by continuous

accumulation of aberrations, including deleterious mutations ³². However, as clinicians and cancer biologists we have a major interest in the few mutations and ancestors that drive progression and treatment resistance ³³.

In conclusion, the study highlights the complexity of MM evolution, including spatial heterogeneity during treatment, a link between sub-clones and growth patterns (focal vs. diffuse), resistant clones that may be hidden/dormant over many years of follow-up, as well as the emergence of multiple selective clonal sweeps derived from single cells (**Fig. 7**). We show the value of medical imaging as a tool to identify new patterns of MM evolution, which allowed us to provide one possible explanation for the negative prognostic impact of baseline focal lesions. The massive clonal changes over time, even in relapsed MM patients with seemingly stable disease, will make it necessary to assess the current genomic profiles if targeted therapies are considered. Last but not least, presence of single cells that can drive relapse even after 10 years of remission suggest that a prerequisite for curative therapies would be to overcome not only tumor heterogeneity but also dormancy.

METHODS

Patients and samples

We performed whole-exome sequencing (WES) of CD138 purified bio-banked MM cells, which had been collected from 25 patients at the University of Arkansas for Medical Sciences between 2003-2017. The patients were enrolled in total therapy protocols ³⁴⁻³⁶ or treated with total therapy-like frontline regimen including multi-agent induction therapy, stem cell transplantation, and intensified maintenance therapy (**Suppl. Table 1**). For the comparison of paired samples in treated patients, we included specimens which were collected within 2 months of the same treatment line. We discriminated between randomly collected samples from the iliac crest and CT-guided specimens from focal lesions, representing diffuse tumor cell infiltration and nodular tumor cell accumulations, respectively. In total, our study included 144 tumor samples, including 96, 43, 3 and 2 iliac crest specimens, focal lesion biopsies, soft tissue specimens and peripheral blood plasma cell leukemia samples, respectively. CT-guided biopsies were taken from all types of bone and extramedullary lesions excluding lesions located to the long bones due to the risk of peri-interventional fracture (Suppl. Data 1). Not all patients and samples were included into each analysis, e.g. in order to validate our recent findings, samples which were included in our previous analysis of spatial heterogeneity ¹¹ were not considered for comparisons of paired samples. Patient #25 was only considered for the comparison of paired iliac crest samples, since the patient did not receive treatment during the observation period (**Suppl. Table 6**).

Informed consent for treatment and sample procurement was obtained in accordance with the Declaration of Helsinki for all cases included in this study, which was approved by the local institutional review board (#02815).

Whole-exome sequencing and mutation calling

Tumor and control DNA were isolated from CD138-positive PCs and peripheral blood leukapheresis products collected after induction therapy, respectively. WES libraries were prepared using the SureSelectQXT sample prep kit in combination with the SureSelect Clinical

Research Exome kit (Agilent) or the HyperPlus kit (Kapa Biosystems) combined with the SeqCap EZ MedExome kit (Roche Nimblegen). The exome kits contained additional baits covering the Ig and MYC loci. De-multiplexing of raw paired-end sequencing data was performed using bcl2fastq. Reads were aligned to the human genome reference GRCh37 release 75 using BWA-mem version 0.7.12³⁷. Sambamba v0.5.6³⁸ was used to sort and index bam files and to mark duplicates. Somatic single-nucleotide variants (SNVs) were called using MuTect v1.1.7³⁹ and Strelka v2.9.10⁴⁰ with default parameters. The intersection of SNVs identified by both variant callers was filtered using the ffilter.pl script (<https://github.com/ckandoth/variant-filter>) with default parameters. After exclusion of variants located in immunoglobulin loci, we determined read counts for all mutations and samples per patient using the Rsamtools R package v1.24.0 and the following inclusion criteria: unique reads, coverage exceeding 20× in all samples of the patient, a mapping quality of at least 20 and base quality of at least 20 at the site of the variant. To maintain a conservative approach and avoid an overestimation of heterogeneity, we only included SNVs with a cancer clonal fraction (CCF) of ≥0.20 in at least one sample (corresponding to a clonal proportion of 20%) and called this SNP in the paired sample(s) if at least two variant reads were detected (usually, MuTect only calls mutations if at least three variant reads are detected). Furthermore, for heterogeneous mutations, we performed manual somatic variant refinement using IGV⁴¹ according to a published standard operating procedure⁴². This filtered set of SNVs, which had a median coverage of 121x, was annotated using SNPEff v4.3⁴³ and used for downstream analyses. Missense, nonsense, splice-site, and frameshift SNVs were defined as non-silent. For the description of heterogeneity, we used the same terminology as described previously¹¹. Briefly, we called mutations *shared*, if they were present in both samples with the same or similar CCF. We classified mutations with at least a three-fold difference in CCF as shared-differential ("*shared-diff*"). We called mutations *unshared*, if they were detectable in only one of the paired samples and discriminated between minor (CCF<60%) and major (CCF≥60%) mutations.

Translocation calling and copy number profiling

Ig and MYC translocations were identified using Manta v1.5.0⁴⁴. Translocations with somatic variant quality scores <30 or "imprecise" calls were removed. All translocation calls were manually inspected in IGV⁴¹. For annotation of translocations we used ANNOVAR⁴⁵. CNAs were called using Sequenza. For each sample, the accuracy of copy number calls was verified by manual inspection of LogR and BAF values for each CNA. To avoid overcalling heterogeneity we used a threshold of 5 Mb for global CNA analyses as described¹¹. For the detection of deletions affecting MM driver genes we used a threshold of 1 Mb.

Evolutionary patterns and sub-clonal reconstruction

The CCF was calculated as described⁴⁶. Briefly, the mutation copy number was determined using the following equation:

$$n_{mut} = f_s \times \frac{1}{p} [pn_{locus}^t + 2(1 - p)]$$

where n_{mut} is the mutation copy number, f_s is the fraction of mutated reads (variant allele frequency), p is the tumor purity, and n_{locus}^t is the locus-specific copy number. For p and n_{locus}^t

we used the values predicted by Sequenza. We then compared the expected f_s value to values assuming the mutation was on 1,2,3, ..., C chromosomes and assigned n_{chr} the value of C with the maximum likelihood using a binomial distribution. Finally the CCF was determined by dividing n_{mut} by n_{chr} . Clonal substructures were inferred using SciClone⁴⁷ with default parameters and the filtered set of SNVs. For the manual design of mock phylogenetic trees, the output of SciClone was further interpreted after inclusion of copy number data. Sub-clones were defined based on SciClone clusters and presence of at least three mutations or at least one copy number aberration. To identify expansions of single tumor cells during treatment we applied a recently published strategy based on mutational footprints of chemotherapies^{16,48,49}. This strategy is based on the fact that the underlying mutations are only detectable through bulk sequencing when single cells expand. Briefly, we first fitted sub-clone defining mutations in each patient (cut-off of 25 mutations, median: 57, range: 25-405) with the mutational signature single-base substitution (SBS)-MM1 (melphalan exposure) and the latest COSMIC reference (<https://cancer.sanger.ac.uk/cosmic/signatures/SBS/>) for SBS1, SBS2, SBS5, SBS8, SBS9, SBS13, and SBS18 using *mmsig* according to the authors' recommendations^{6,17,50}. Presence of SBS-MM1 was further confirmed using the mSigAct signature presence test (function: `process.one.group()`) and a p value cut-off of 0.05 to account for the limited number of mutations in WES^{48,51}.

Derivation of the GEP70 risk signature

For risk stratification, we applied the GEP70 model, which is based on Affymetrix U133Plus2.0 microarray data (Santa Clara, CA) for CD138-enriched PCs¹⁹. Raw intensity values were MAS5 normalized, converted to log2 scale and corrected for batch effects using M-ComBat⁵². The GEP70 corresponds to the average log2 expression of 51 upregulated genes minus the average log2 expression of 19 downregulated genes, and scores ≥ 0.66 indicate high-risk.

Medical imaging

PET with CT attenuation correction (PET-CT) and diffusion-weighted magnetic resonance imaging with background suppression (DWIBS) were done as recently described¹³. Briefly, PET-CT was performed on a Biograph 6 PET/CT system (Siemens Medical Solutions, PA, USA), a GE Discovery IQ scanner (GE Healthcare, IL, USA) or a CTI-Reveal scanner (Siemens Medical Solutions). Images were acquired from the vertex to toes. After iterative reconstruction, images were reviewed using the PET volume computer assisted reading software (AW server, version 3.2, General Electric, WI, USA). DWIBS was performed on a 1.5 Tesla Achieva scanner (PHILIPS, MA, USA). Scanning was performed from vertex to toes in 7 to 9 slabs, depending on the patient's height. A coronal whole-body T1 turbo spin echo image was used as a localizer. Images were analyzed in an inverted grayscale with fused whole-body maximum-intensity projection reconstructions of the diffusion and exponential apparent-diffusion coefficient images. For PET-CT, a focal lesion was defined as a circumscribed focus with increased FDG uptake compared to its surroundings. For DWIBS, a focal lesion was defined as a well delineated focal intensity above the surrounding background BM ≥ 1 cm in size.

Statistical methods

Statistical analyses were carried out using the R software package 3.6.0. Group comparisons of continuous variables were done using the Mann-Whitney Wilcoxon test for independent groups. Differences in treatment responses and evolution patterns between groups was assessed using Fisher's exact test.

Data availability

The WES dataset is deposited in dbGAP (accession: phs2625.v1) under restricted access; access can be obtained via a material transfer agreement with the University of Arkansas for Medical Sciences. All other data are either presented in the manuscript or available on request.

Figure Legends

Fig. 1: Patient characteristics, sample origin, and collection time points. In (a) the origin and number of investigated samples are depicted. In (b) gender, molecular subtype, risk according to the GEP70¹⁹, as well as the type of samples and the time points of sample collection are shown. EMD: extramedullary disease, PCL: plasma cell leukemia, NHD: non-hyperdiploid karyotype, HY: hyperdiploid, NA: not available, Lines: number of treatment lines (for a detailed description of treatment see Suppl. Table 1).

Fig. 2: Spatial heterogeneity in time and space. We compared mutational profiles of paired samples collected from different bone marrow sites at baseline and during treatment, and discriminated between randomly collected samples from the iliac crest and CT-guided specimens from focal lesions. In (a) the functional imaging and the corresponding cancer clonal fraction plots for detected mutations for patient #5 with paired samples from the right and left iliac crest as well as a sample from a left sacrum focal lesion are shown. In (b) and (c) the proportion of shared and heterogeneous mutations in paired left/right iliac crest and paired iliac crest/focal lesion samples collected baseline and within two months of the same treatment line are presented, respectively. In (d) the mutational profiles of the first and the last available sample per patient are compared. For this comparison, patient #25 was not considered. For patients, where an iliac crest sample was compared to a focal lesion or a peripheral blood sample, the ID is printed in bold with an underscore. Mutations were called shared (grey color), if they were present with the same or similar cancer clonal fraction (CCF) in paired samples. For heterogeneous mutations we discriminated between sub-clonal (CCF<60%) and clonal (CCF≥60%) mutations. Mutations with at least a three-fold difference in CCF were classified as shared-differential (sub-clonal: green; clonal: blue). We called mutations unshared, if they were detectable in only one of the paired samples (sub-clonal: pink; clonal: red). In (e), the spatio-temporal changes affecting known myeloma driver aberrations are presented, including MYC translocations, gain(1q), del(17p), as well as myeloma driver genes ²¹, which showed non-silent mutations in at least two patients in our set. Grey indicates shared (clonal) events. Green, red, and blue denote spatial, temporal and spatio-temporal differences/changes, respectively.

Fig. 3: Phylogenetic relationship between baseline and relapse clones. We analyzed the phylogenetic relationship between baseline clones from different bone marrow regions and relapse clones to elucidate the origin of resistant disease. In (a) the mock phylogenetic tree for P6 is shown as an example for a pattern where the unique baseline focal lesion clone and the dominant relapse clone share the phylogenetic branch. The boxes show the proportion of detected sub-clones with each sub-clone having a distinct color. For instance, 10 blue dots indicate a proportion of 10% for the “blue” sub-clone. Dashed lines illustrate the origin/relationship of these sub-clones and the color code corresponds to the color of sub-clones. The grey box/card indicate that tumor data is not available for the respective site at relapse. (b) For 12 patients with paired baseline iliac crest/focal lesion samples the relationship is shown separately for the focal lesion (left bar), and the iliac crest (middle bar). The right bar shows the relationship between baseline iliac crest and relapse clones for all patients with an iliac crest sample at baseline. No close relationship (grey): baseline and

relapse clones represent different evolutionary branches as shown for the dominant iliac crest and the relapse clone in (a); selected: a pre-existing resistant clone is selected (example in Suppl. Fig. 1); branch: the two clones share the same branch, as shown for the baseline focal lesion and the relapse clone in (a); precursor: linear relationship between baseline and relapse clone, i.e. the clone acquires new aberrations; SC: sub-clone is related to relapse clone; MC: major/dominant clone is related to relapse clone.

Fig. 4: Evolutionary patterns during treatment. In (a) the mock phylogenetic tree for patient #22 is shown as an example for a pattern where relapse is initially driven by a single cell. Subsequently, the first relapse clone is replaced by a further evolved descendant clone with a new del(4q). Expansion of a single cell was demonstrated by a mutational signature analysis which showed significant contribution of the melphalan signature SBS-MM1 to relapse clone defining mutations. In (b) two independent branches drive relapse and longitudinally alternate in dominance in patient #14 which resembles the classical alternating clonal dominance model by Keats *et al*¹⁵. Yet, clonal dominance also alternates in space. The boxes show the proportion of detected sub-clones with each sub-clone having a distinct color. Dashed lines illustrate the origin/relationship of sub-clones/branches. Selected clone-defining mutations are shown for each branch, with known myeloma drivers being in bold. The length of branches does not indicate the extent of differences between clones but the time point of their first appearance. Grey boxes/cards indicate that tumor data is not available for the respective timepoints and sites.

Fig. 5: Evolutionary patterns and association with clinical features. In (a) the 3 evolutionary patterns, which we observed in this study, are illustrated. These include 1) expansion and sweep driven by single tumor cells, 2) coexisting expanding sub-clones, and 3) site-unique expansions of distinct sub-clones with the main difference between the first and the third pattern being the anatomical location of sub-clones. In (b) boxplots are shown for the association between these patterns and 1) the number of PET-positive focal lesions at baseline (upper panel) and 2) the response to first-line therapy (lower panel), respectively.

Fig. 6: Spatial-longitudinal alternating dominance of single-cell expansions. The figure shows spatial-longitudinal data for patient #12. At baseline, the patient presented with a t(4;14) and the same clonal composition was seen at the left and right iliac crest. The first follow-up sample (1st FU in upper plot) was collected from a focal lesion at the sacrum after the 6th treatment line 11 years after enrolment (prior best response: VGPR, detailed treatment information in Suppl. table 1). The dominant clone (blue branch in mock tree) was positive for the SBS-MM1-signature, in line with a single-cell expansion. The precursor of this clone (grey-sub-clone in tree) was already detectable at the sub-clonal level at baseline. Two months later, a sample was collected from the right iliac crest after treatment with carfilzomib (best response: SD). While the focal lesion clone was detectable at the sub-clonal level, another SBS-MM1-positive clone with a MYC translocation (green branch in tree) dominated at this location, suggesting clonal co-existence or competition. Despite extensive treatment with pomalidomide, polychemotherapy and MEK inhibitors (best response: SD), a very similar clonal composition was seen in two focal lesions and 6 iliac crest samples, consecutively collected over 10 months. However, after another ASCT, the patient achieved CR, and suddenly dramatic clonal changes appeared. The focal lesion clone became dominant at the iliac crest and a new clone emerged, which was not seen in prior samples (red branch in tree).

This del(17p) and signature SBS-MM1 positive clone shared the pre-existing precursor with the other two relapse clones but represented another branch. The clone dominated at the iliac crest at the last available time point. Single-cell expansions were predicted based on the presence of the SBS-MM1 melphalan signature. The boxes show the proportion of detected sub-clones with each sub-clone having a distinct color. Dashed lines illustrate the origin/relationship of sub-clones/branches. Selected clone-defining mutations are shown in the mock phylogenetic tree, with known myeloma drivers being in bold. The length of branches of the phylogenetic tree corresponds to the time point of first appearance of respective clones.

Fig. 7: Summary of the spatial temporal evolution of multiple myeloma. The figure summarizes our findings and the complexity of myeloma evolution seen in this study. It includes spatial heterogeneity at baseline and after treatment, a link between sub-clones and growth patterns, resistant clones that may be hidden/quiescent over many years of follow-up, as well as the emergence of multiple selective clonal sweeps derived from single cells, diversification of expanding sub-clones and competition between them.

FIGURES

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

a

b

Fig. 6

Fig. 7

REFERENCES

- 1 Magrangeas F, Avet-Loiseau H, Gouraud W, Lodé L, Decaux O, Godmer P *et al.* Minor clone provides a reservoir for relapse in multiple myeloma. *Leukemia* 2013; **27**: 473–481.
- 2 Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I *et al.* Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun* 2014; **5**: 2997.
- 3 Weinhold N, Ashby C, Rasche L, Chavan SS, Stein C, Stephens OW *et al.* Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. *Blood* 2016; **128**: 1735–1744.
- 4 Bolli N, Maura F, Minvielle S, Gloznik D, Szalat R, Fullam A *et al.* Genomic patterns of progression in smoldering multiple myeloma. *Nat Commun* 2018; **9**: 3363.
- 5 Jones JR, Weinhold N, Ashby C, Walker BA, Wardell C, Pawlyn C *et al.* Clonal evolution in myeloma: the impact of maintenance lenalidomide and depth of response on the genetics and sub-clonal structure of relapsed disease in uniformly treated newly diagnosed patients. *Haematologica* 2019; **104**: 1440–1450.
- 6 Maura F, Bolli N, Angelopoulos N, Dawson KJ, Leongamornlert D, Martincorena I *et al.* Genomic landscape and chronological reconstruction of driver events in multiple myeloma. *Nat Commun* 2019; **10**: 3835.
- 7 Boyle EM, Deshpande S, Tytarenko R, Ashby C, Wang Y, Bauer MA *et al.* The molecular make up of smoldering myeloma highlights the evolutionary pathways leading to multiple myeloma. *Nat Commun* 2021; **12**: 293.
- 8 Moreau P, Attal M, Caillot D, Macro M, Karlin L, Garderet L *et al.* 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. *Journal of Clinical Oncology*. 2017; **35**: 2911–2918.
- 9 Rasche L, Alapat D, Kumar M, Gershner G, McDonald J, Wardell CP *et al.* Combination of flow cytometry and functional imaging for monitoring of residual disease in myeloma. *Leukemia* 2019; **33**: 1713–1722.
- 10 Alonso R, Cedena MT, Gómez-Grande A, Ríos R, Moraleda JM, Cabañas V *et al.* Imaging and bone marrow assessments improve minimal residual disease prediction in multiple myeloma. *Am J Hematol* 2019; **94**: 853–861.
- 11 Rasche L, Chavan SS, Stephens OW, Patel PH, Tytarenko R, Ashby C *et al.* Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. *Nat Commun* 2017; **8**: 268.
- 12 Zamagni E, Patriarca F, Nanni C, Zannetti B, Englaro E, Pezzi A *et al.* Prognostic relevance of 18-F FDG PET/CT in newly diagnosed multiple myeloma patients treated with up-front autologous transplantation. *Blood* 2011; **118**: 5989–5995.
- 13 Rasche L, Angtuaco EJ, Alpe TL, Gershner GH, McDonald JE, Samant RS *et al.* The presence of large focal lesions is a strong independent prognostic factor in multiple myeloma. *Blood* 2018; **132**: 59–66.

- 14 Davies FE, Rosenthal A, Rasche L, Petty NM, McDonald JE, Ntambi JA *et al.* Treatment to suppression of focal lesions on positron emission tomography-computed tomography is a therapeutic goal in newly diagnosed multiple myeloma. *Haematologica* 2018; **103**: 1047–1053.
- 15 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**: 1067–1076.
- 16 Landau HJ, Yellapantula V, Diamond BT, Rustad EH, Maclachlan KH, Gundem G *et al.* Author Correction: Accelerated single cell seeding in relapsed multiple myeloma. *Nat Commun* 2021; **12**: 591.
- 17 Rustad EH, Yellapantula V, Leongamornlert D, Bolli N, Ledergor G, Nadeu F *et al.* Timing the initiation of multiple myeloma. *Nat Commun* 2020; **11**: 1917.
- 18 Corre J, Cleyne A, Robiou du Pont S, Buisson L, Bolli N, Attal M *et al.* Multiple myeloma clonal evolution in homogeneously treated patients. *Leukemia* 2018; **32**: 2636–2647.
- 19 Shaughnessy JD Jr, Zhan F, Burington BE, Huang Y, Colla S, Hanamura I *et al.* A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* 2007; **109**: 2276–2284.
- 20 Rübber A, Araujo A. Cancer heterogeneity: converting a limitation into a source of biological information. *J Transl Med* 2017; **15**: 190.
- 21 Walker BA, Mavrommatis K, Wardell CP, Ashby TC, Bauer M, Davies FE *et al.* Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. *Blood* 2018; **132**: 587–597.
- 22 Pawlyn C, Bright MD, Buros AF, Stein CK, Walters Z, Aronson LI *et al.* Overexpression of EZH2 in multiple myeloma is associated with poor prognosis and dysregulation of cell cycle control. *Blood Cancer J* 2017; **7**: e549.
- 23 Maura F, Weinhold N, Diamond B, Kazandjian D, Rasche L, Morgan G *et al.* The mutagenic impact of melphalan in multiple myeloma. *Leukemia* 2021. doi:10.1038/s41375-021-01293-3.
- 24 Da Vià MC, Dietrich O, Truger M, Arampatzi P, Duell J, Heidemeier A *et al.* Homozygous BCMA gene deletion in response to anti-BCMA CAR T cells in a patient with multiple myeloma. *Nat Med* 2021; **27**: 616–619.
- 25 Raab MS, Lehnert N, Xu J, Ho AD, Schirmacher P, Goldschmidt H *et al.* Spatially divergent clonal evolution in multiple myeloma: overcoming resistance to BRAF inhibition. *Blood*. 2016; **127**: 2155–2157.
- 26 McDonald JE, Kessler MM, Gardner MW, Buros AF, Ntambi JA, Waheed S *et al.* Assessment of Total Lesion Glycolysis by F FDG PET/CT Significantly Improves Prognostic Value of GEP and ISS in Myeloma. *Clin Cancer Res* 2017; **23**: 1981–1987.
- 27 Ulaner GA, Sobol NB, O'Donoghue JA, Kirov AS, Riedl CC, Min R *et al.* CD38-targeted Immuno-PET of Multiple Myeloma: From Xenograft Models to First-in-Human Imaging. *Radiology* 2020; **295**: 606–615.
- 28 Zamagni E, Nanni C, Dozza L, Carlier T, Bailly C, Tacchetti P *et al.* Standardization of

- F-FDG-PET/CT According to Deauville Criteria for Metabolic Complete Response Definition in Newly Diagnosed Multiple Myeloma. *J Clin Oncol* 2021; **39**: 116–125.
- 29 de Jong MME, Kellermayer Z, Papazian N, Tahri S, Hofste Op Bruinink D, Hoogenboezem R *et al.* The multiple myeloma microenvironment is defined by an inflammatory stromal cell landscape. *Nat Immunol* 2021; **22**: 769–780.
 - 30 Zavidij O, Haradhvala NJ, Mouhieddine TH, Sklavenitis-Pistofidis R, Cai S, Reidy M *et al.* Single-cell RNA sequencing reveals compromised immune microenvironment in precursor stages of multiple myeloma. *Nat Cancer* 2020; **1**: 493–506.
 - 31 Cohen YC, Zada M, Wang S-Y, Bornstein C, David E, Moshe A *et al.* Identification of resistance pathways and therapeutic targets in relapsed multiple myeloma patients through single-cell sequencing. *Nat Med* 2021; **27**: 491–503.
 - 32 Muller HJ. THE RELATION OF RECOMBINATION TO MUTATIONAL ADVANCE. *Mutat Res* 1964; **106**: 2–9.
 - 33 Caravagna G, Heide T, Williams MJ, Zapata L, Nichol D, Chkhaidze K *et al.* Subclonal reconstruction of tumors by using machine learning and population genetics. *Nat Genet* 2020; **52**: 898–907.
 - 34 Jethava YS, Mitchell A, Epstein J, Zangari M, Yaccoby S, Tian E *et al.* Adverse Metaphase Cytogenetics Can Be Overcome by Adding Bortezomib and Thalidomide to Fractionated Melphalan Transplants. *Clin Cancer Res* 2017; **23**: 2665–2672.
 - 35 Jethava Y, Mitchell A, Zangari M, Waheed S, Schinke C, Thanendrarajan S *et al.* Dose-dense and less dose-intensive total therapy 5 for gene expression profiling-defined high-risk multiple myeloma. *Blood Cancer J* 2016; **6**: e471.
 - 36 van Rhee F, Szymonifka J, Anaissie E, Nair B, Waheed S, Alsayed Y *et al.* Total Therapy 3 for multiple myeloma: prognostic implications of cumulative dosing and premature discontinuation of VTD maintenance components, bortezomib, thalidomide, and dexamethasone, relevant to all phases of therapy. *Blood* 2010; **116**: 1220–1227.
 - 37 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
 - 38 Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 2015; **31**: 2032–2034.
 - 39 Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013; **31**: 213–219.
 - 40 Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* 2012; **28**: 1811–1817.
 - 41 Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G *et al.* Integrative genomics viewer. *Nature Biotechnology*. 2011; **29**: 24–26.
 - 42 Barnell EK, Ronning P, Campbell KM, Krysiak K, Ainscough BJ, Sheta LM *et al.* Standard operating procedure for somatic variant refinement of sequencing data with paired tumor and normal samples. *Genet Med* 2019; **21**: 972–981.
 - 43 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L *et al.* A program for

annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 2012; **6**: 80–92.

- 44 Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M *et al.* Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 2016; **32**: 1220–1222.
- 45 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; **38**: e164.
- 46 Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC *et al.* The landscape of cancer genes and mutational processes in breast cancer. *Nature* 2012; **486**: 400–404.
- 47 Miller CA, White BS, Dees ND, Griffith M, Welch JS, Griffith OL *et al.* SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput Biol* 2014; **10**: e1003665.
- 48 Boot A, Huang MN, Ng AWT, Ho S-C, Lim JQ, Kawakami Y *et al.* In-depth characterization of the cisplatin mutational signature in human cell lines and in esophageal and liver tumors. *Genome Res* 2018; **28**: 654–665.
- 49 Pich O, Muiños F, Lolkema MP, Steeghs N, Gonzalez-Perez A, Lopez-Bigas N. The mutational footprints of cancer therapies. *Nat Genet* 2019; **51**: 1732–1740.
- 50 Rustad EH, Nadeu F, Angelopoulos N, Ziccheddu B, Bolli N, Puente XS *et al.* mmsig: a fitting approach to accurately identify somatic mutational signatures in hematological malignancies. *Commun Biol* 2021; **4**: 424.
- 51 Ng AWT, Poon SL, Huang MN, Lim JQ, Boot A, Yu W *et al.* Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia. *Sci Transl Med* 2017; **9**. doi:10.1126/scitranslmed.aan6446.
- 52 Stein CK, Qu P, Epstein J, Burows A, Rosenthal A, Crowley J *et al.* Removing batch effects from purified plasma cell gene expression microarrays with modified ComBat. *BMC Bioinformatics*. 2015; **16**. doi:10.1186/s12859-015-0478-3.

Supplementary Files

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

- [Supplmaterial.pdf](#)
- [NCOMMS2150104RS.pdf](#)