

# Dynamic super-enhancer core regulatory circuits and epigenetic landscapes drive malignant progression and refractory disease in multiple myeloma

**Ariosto Siqueira Silva** (✉ [ariosto.silva@moffitt.org](mailto:ariosto.silva@moffitt.org))

H Lee Moffitt Cancer Center and Research Institute

**Rafael Canevarolo**

H Lee Moffitt Cancer Center and Research Institute <https://orcid.org/0000-0002-8722-8512>

**Mark Meads**

H Lee Moffitt Cancer Center and Research Institute

**Maria Silva**

H Lee Moffitt Cancer Center and Research Institute

**Praneeth Sudalagunta**

Moffitt Cancer Center <https://orcid.org/0000-0003-1283-9332>

**Gabriel De Avila**

H Lee Moffitt Cancer Center and Research Institute

**Raghunandan Alugubelli**

H Lee Moffitt Cancer Center and Research Institute

**Alexandre Tungesvik**

University of South Florida

**Elissa Bell**

Florida State University

**Karen Burger**

H Lee Moffitt Cancer Center and Research Institute

**Amit Kulkarni**

M2Gen

**Oliver Hampton**

M2Gen <https://orcid.org/0000-0001-9164-1070>

**Zhijie Jiang**

M2Gen

**Hongyue Dai**

M2Gen

**Christopher Cubitt**

Moffitt Cancer Center <https://orcid.org/0000-0001-6343-7451>

**Jamie Teer**

H. Lee Moffitt Cancer Center and Research Institute <https://orcid.org/0000-0003-4513-0282>

**Eric Welsh**

Moffitt Cancer Center

**Sean Yoder**

H. Lee Moffitt Cancer Center

**Bijal Shah**

Moffitt Cancer Center

**Jianguo Tao**

Moffitt Cancer Center

**Lori Hazlehurst**

West Virginia University

**Robert Gatenby**

Moffitt Cancer Center

**Daniel Sullivan**

H Lee Moffitt Cancer Center and Research Institute

**Melissa Alsina**

H. Lee Moffitt Cancer Center and Research Institute

**Taiga Nishihori**

H Lee Moffitt Cancer Center and Research Institute

**Jason Brayer**

H Lee Moffitt Cancer Center and Research Institute

**John L. Cleveland**

Department of Tumor Biology, The Moffitt Cancer Center and Research Institute <https://orcid.org/0000-0002-7561-9065>

**Dalton William**

Moffitt Cancer Center

**Robert Gillies**

Moffitt Cancer Center <https://orcid.org/0000-0002-8888-7747>

**Rachid Baz**

H Lee Moffitt Cancer Center and Research Institute

**Kenneth Shain**

Moffitt Cancer Center

---

**Article**

**Keywords:** Plasma Cell Malignancy Multiple Myeloma, Epigenetic States, Tumor Microenvironment, Cell Cycle, Metabolic Control Genes

**Posted Date:** January 4th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-125536/v1>

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

[Read Full License](#)

---

# Abstract

The plasma cell malignancy multiple myeloma (MM) evolves from a pre-malignant state and remains all but incurable due to emergence of therapy resistance. Despite intensive analyses, mechanisms driving MM progression and refractory disease are poorly understood. Integrating topologic, expression and epigenetic analyses of 1,016 patient specimens, we report super-enhancer core regulatory circuits (SECRCs) that drive and sustain MM epigenetic states. Reprogramming of cell identity and tumor microenvironment genes drive malignant conversion, while alterations in cell cycle and metabolic control genes cause refractory disease. Thus, select epigenetic states drive progression and therapy resistance, providing strategies to prevent and effectively treat MM.

## Introduction

Multiple myeloma (MM) is a cancer of plasma cells that evolves from a pre-malignant state (MGUS)<sup>1,2</sup>. Unfortunately, MM remains incurable for the vast majority of patients, as a consequence of the emergence of multi-drug resistance after repeated courses of therapeutic regimens<sup>2-4</sup> (**Figure 1a**). Extensive studies have linked MM progression and refractory disease with the accumulation of cytogenetic abnormalities, mutations, and dysregulated DNA methylation<sup>5-7</sup>. However, the mechanisms driving MM progression and multi-drug resistance remain unclear, in part, due to the significant interpatient heterogeneity<sup>5,8-10</sup>.

We hypothesize that the genetic diversity observed in MM converge into a common set of biological traits, subject to evolutionary dynamics imposed by tumor microenvironment (TME) and therapy. To test this hypothesis, we developed an unsupervised systems biology approach, integrating genetic, transcriptional, epigenetic and longitudinal clinical data, from a new cohort of MM patients, through Moffitt Cancer Center's Total Cancer Care protocol<sup>11</sup>. Analysis of matched RNAseq and whole exome sequencing (WES), from CD138+ (plasma) cells and tissue samples from 1,016 bone marrow aspirates across MM spectrum (**Figure 1b**), demonstrated that despite a large number of cytogenetic abnormalities, no consistent increases in driver mutation frequencies were associated with progression from MGUS to newly diagnosed MM (NDMM), or with the development of refractory disease (late relapse, LRMM). Rather, MM progression correlated with decreased expression of genes that control cell identity and TME interactions, whereas LRMM demonstrated increased expression of cell cycle and metabolism-associated genes. Furthermore, these genes were enriched for transcriptional regulation based on specific chromatin accessibility regulating histone alterations. Single-cell chromatin accessibility analysis (scATAC-seq) of primary specimens revealed that changes in MM chromatin accessibility involved transcription factors associated with induced pluripotent stem cells (iPSC)<sup>12</sup> and formation of super-enhancers (SE). The essential role of self-regulatory SE-regulated genes, which are differentially expressed during MM progression and in LRMM, was independently confirmed in the analyses of publicly available databases reporting silencing viability assays in MM cell lines. Additionally, common MM mutations and cytogenetic abnormalities correlated with differential expression observed in the transitions from pre-malignant to

NDMM (NRAS, KRAS, del13q), or LRMM (TP53, del17p, amp1q21). Thus, transitions from MGUS to NDMM, and later to LRMM, are driven by epigenetic dysregulation of super-enhancers that occur in concert with other independent genomic events.

## Results

**Transcriptional topology reveals regulatory pathways in MM biology.** To identify the mechanisms driving transcriptional alterations across MM spectrum, we recreated a transcriptional topology of MM using dimensionality reduction analysis (t-SNE<sup>26</sup>, *Extended Data Figure 2a*). In this topology, 16,738 genes were spatially distributed according to co-expression across the cohort of 844 RNAseq samples. An unsupervised approach (fuzzy c-Means<sup>27</sup>) was used to segment clusters of co-expressing genes, resulting in 500 gene clusters (*Extended Data Figure 2b*), each of which contained a number of genes controlled by a shared transcriptional program. A knowledge-based approach, consisting in highlighting genes in manually curated Cancer Hallmarks and KEGG pathways (*Extended Data Figure 2c-j*), revealed that some gene sets are transcriptionally coordinated (e.g. Ribosome, *Extended Data Figure 2j*), while others are under more diffuse control (e.g. Focal Adhesion, *Extended Data Figure 2h*). Unsupervised hierarchical clustering, based on Pearson correlation between ssGSEA NES, calculated across all samples, for aforementioned gene clusters and Cancer Hallmarks, identified two superclusters manifest in MM, **a** and **b** (*Extended Data Figure 3a*). Supercluster **a**'s expression correlated with 14 hallmarks, including *protein secretion*, *oxidative phosphorylation*, and *DNA repair*, while the expression of supercluster **b** correlated with 17 hallmarks involved in *EMT*, *inflammation*, *immunogenicity* and *hypoxia-response*. These observations suggested that these biological pathways are coordinated by a shared mechanism of transcriptional regulation. Importantly, the integrity of these superclusters was conserved in the t-SNE (*Figure 3a*), and their differential expression, as estimated by ssGSEA NES, significantly changed, with **a** increasing and **b** decreasing (*Figure 3b*), along transitions to NDMM or LRMM.

Independently, genes differentially expressed during progression and refractory disease also co-localized in the transcriptional topology, suggesting shared transcriptional mechanisms (*Figure 3c-d*). Specifically, the under-expressed genes identified in the MGUS-to-NDMM transition (*Figure 3c*), overlapped with supercluster **b**, while those differentially expressed in the NDMM-to-LRMM transition (*Figure 3d*) overlapped with supercluster **a**. These differentially expressed genes were not enriched for any particular chromosomal region (*Extended Data Figure 3b*), suggesting the involvement of global, rather than local transcriptional dysregulation mechanisms. Accordingly, topologic analysis of the MM transcriptome indicates at least two conserved gene superclusters associated with biology driving malignant transformation and treatment refractory disease in MM.

**Epigenetic modifiers are dysregulated in MM.** Using the pathway analysis aggregator *Enrichr*<sup>28</sup>, based on Chip-seq data (ENCODE<sup>29</sup> and ChEA consensus), we identified DNA-binding proteins (DBP) with enriched binding to the genes differentially expressed between MM states. 13 DBPs were enriched for the genes

differentially expressed between MGUS/SMOL and NDMM, whereas 75 DBPs were enriched for the NDMM-to-LRMM transition (**Extended Data Figure 3d-e**). While the former included genes associated with PRC2 complex<sup>30</sup> and plasma cell differentiation<sup>31</sup> (*SUZ12*, *EZH2* and *SOX2*), the latter were notably associated with telomere formation/TERT activation in cancer (*RFX5*, *CEBPB*, *MAX*, *SP1*, *MYC*, *STAT3*, *E2F1*, *KLF4*, *ETS1*, *USF2*, *USF1*)<sup>32</sup> as well as the formation of topologically associated domains (TADs), essential to the formation of super-enhancers and establishment of cell identity<sup>29,33,34</sup> (*YY1*, *CTCF*, *GABPA* and *NRF1*). Given the role of these DBPs, and the contribution of epigenetics to cancer<sup>1,7</sup> and hematopoietic cell identity<sup>31</sup>, we have investigated the putative role of methylation or acetylation modifications of histone 3 lysine 27 in transcriptional dysregulation across MM spectrum. **Figure 3e** depicts gene clusters *individually* enriched for histone modifications responsible for increased (H3K27ac) or decreased (H3K27me3) chromatin accessibility<sup>35-37</sup>, based on enrichment score calculated from Chip-seq data repositories. H3K27ac- and H3K27me3-enriched gene clusters coalesced in two large regions of the t-SNE, suggesting the overall transcriptional topology of MM is defined by two epigenetic regulatory mechanisms that control cell differentiation or super-enhancers. Differential expression of these two putatively epigenetically-regulated gene sets, as calculated by ssGSEA NES, revealed significant suppression of H3K27me3-enriched genes between MGUS/SMOL and NDMM samples, while the expression of H3K27ac-enriched genes increased from NDMM to LRMM (**Figure 3f**). These findings were also observed in 8 individual sequential samples of MM patients in this cohort (**Extended Data Figure 4**). Collectively, these data suggest that the transition from pre-malignant to active disease is characterized by increased H3K27me3 histone modification, and consequent reduced chromatin accessibility and decreased expression, of gene sets associated with TME-dependency and cytokine-mediated signaling cascades. In turn, refractory MM correlated with increased H3K27ac histone modification, chromatin accessibility and expression of gene sets involved in cell cycle and metabolism. Additionally, we have observed a strong correlation between expression of epigenetically regulating genes with opposing roles, such as HAT1 and HDAC1 as well as SUZ12 and KDM1A (**Extended Data Figure 5**), suggesting that increased activity of acetylating/de-acetylating, or methylating/demethylating enzymes must be coordinated, to maintain distinct epigenetics states required for transitions to NDMM or LRMM.

**Genome-wide alterations in chromatin accessibility are a hallmark of MM progression.** Epigenetic control of chromatin accessibility has been linked to cell state changes in malignant progression, tumor metastasis and survival<sup>38-42</sup>. To determine whether alterations in transcriptomic programming were linked to chromosomal accessibility in this MM cohort, we conducted scATAC-seq in 10 primary MM samples, across 4 disease states: SMOL (3 samples/5,016 cells), NDMM (4 samples/21,005 cells), ERMM (1 sample/1,987 cells), and LRMM (2 samples/6,953 cells). Consistent with RNAseq-based predictions, we observed significant changes in chromatin accessibility of DBP motifs during MM progression (**Figure 4a-b**). Among the most significantly differentially accessible motifs, there was an inverse correlation of accessibility between POU5F1 (OCT4) and CTCF and YY1 (**Figure 4c-h**, **Extended Data Figure 6a-d**). Together with SOX2, POU5F1 is part of the Yamanaka factors, which regulate pluripotency and cell

differentiation<sup>12</sup> and, with NANOG, these pluripotency-inducer transcription factors (“OSN”) modulate gene repression by recruitment of Polycomb group chromatin regulators, which induce reduction in chromatin accessibility through spread of H3K27me3<sup>43-46</sup>, disrupting cohesin and CTCF and YY1 binding to chromatin, ultimately disassembling TADs in non-OSN SEs. These findings suggest a coordinated mechanism behind the opposing accessibility pattern of OSN and CTCF and YY1 binding sites<sup>47</sup>. Six DBPs were both enriched for binding to differentially expressed genes (**Extended Data Figure 3d**) and showed differential motif accessibility in the MGUS-to-NDMM transition: RUNX1, TCF3, GATA2, IRF8, ESR1 and TP63. Twenty-nine others matched these requirements for the NDMM-to-LRMM transition, among these were RFX5, USF2, YY1, NRF1 and POU5F1. We have identified 1,883 genes whose differential expression across disease states was associated with changes in chromatin accessibility at promoter and distal levels. These genes were used in a pseudotime analysis<sup>48</sup> to estimate the trajectory from SMOL-to-LRMM samples, based on overall (distal + promoter) single cell chromatin accessibility (**Figure 5a-b**). This topology suggested increased inter-patient tumor heterogeneity with disease progression, with cells from different SMOL samples more closely located, while NDMM and LRMM samples were markedly further away in pseudotime.

**Super-enhancer-regulated self-regulatory gene networks maintain MM transcriptional dysregulation.** SE-based gene regulation is central to maintaining cell identity, cancer progression and drug response<sup>41,42,49</sup>. SE-regulated transcription factors (TF) form self-reinforcing transcriptional networks through binding to their own SE regions, stabilizing chromatin in an accessible state (H3K27ac), dramatically increasing their transcription and that of thousands of downstream genes, as a means of epigenetic memory<sup>50</sup>. These networks are termed SE-regulated core regulatory circuits (SECRC)<sup>50</sup>. We generated a transcriptional network, based on a previously established method<sup>50,51</sup>, to evaluate the role of SECRCs in transcriptional reprogramming in MM<sup>37,50,51</sup>. This analysis identified 60 SECRC TFs and 1,865 downstream genes regulated by enhancers with transcription factor binding sites identified for at least one the 60 SECRCs, as determined by dbSuper/GeneCards, in a consensus of B cell lineages. We have investigated changes in chromatin accessibility of these 60 SECRCs across a pseudotime trajectory connecting 3 SMOL, 2 NDMM and one LRMM sample (**Figure 5b-c**). EBF1 and ZEB2 were examples of genes with no significant changes, while RFX5, YY1 and CTCF showed marked increase in accessibility.

In addition, RFX5 was notable as a transcription factor whose changes in expression, motif and combined enhancer/promoter accessibility significantly increased towards refractory disease (**Extended Data Figure 6e-g**), suggesting a driver role. This concept was reinforced by Pearson correlation between *RFX5* expression (RNAseq) and chromatin accessibility (scATAC-Seq) of YY1 and POU5F1 groups in seven MM samples with both datasets available (**Extended Data Figure 7**). Additionally, Pearson correlation between motif accessibility and ssGSEA of gene clusters, cancer hallmarks and KEGG pathways in these seven samples provided biological pathways and gene sets positively- or negatively-correlated with YY1 and POU5F1 motifs (**Extended Data Figure 8**). These observations affirm the

correlation between chromatin accessibility of OSN and TAD motifs and differential expression of pathways involved in MM progression and refractory disease, as well as expression of SECRCs (*e.g.*, *RFX5*) and the accessibility of these motifs.

We thus investigated SECRCs that could putatively regulate OSN or TAD-associated chromatin accessibility motifs, and thus drive MM progression: 8 of the 60 SECRCs were under-expressed and 3 over-expressed in NDMM when compared to MGUS/SMOL samples. Notably, among the under-expressed SECRC were *RB1*, *ELF1*, *IKZF1*, *FOXP1*, *KLF6*, *SPI1*, *ZEB2*, and *EBF1*. The most differentially under-expressed SECRC was *EBF1* (**Figure 5d**), which regulates B cell maturation and with *SPI1*<sup>52</sup> stabilize chromatin accessibility<sup>53,54</sup>. *RB1*<sup>55</sup> and *KLF6*<sup>56</sup> are both tumor suppressor genes. The 3 overexpressed SECRCs have previously been identified as critical to MM biology: *IRF4*, *PRDM1* and *MZF1* have been linked to MM progression via aberrant regulation of *HGF*<sup>51,57,58</sup>. We identified 5 SECRCs under-expressed and 24 over-expressed in LRMM compared to NDMM samples (**Figure 5e**). Among these, 7 were also identified as TFs enriched for the ND→LRMM progression, based on Chip-seq data (**Extended Data Figure 3e**): *YY1*, *CTCF*, *RELA*, *RFX5*, *ELF1*, *NFE2L2* and *CREB1*. The most differentially over-expressed SECRC in LRMM was *RFX5*. The under-expressed SECRCs were *JUN*, *MZF1*, *ZHX2*, *CHD2* and *KLF6*. Note that *MZF1* demonstrated a reversal of differential expression from MGUS to NDMM. The oncogenic SECRC *MYC* was over-expressed in LRMM vs. NDMM samples (**Figure 5f**), but was not listed in this analysis due to the FDR cutoff used ( $q < 0.01$ ): *MYC*'s  $q = 0.015$ , mean z-score 0.2144 (LR) vs. -0.0416 (ND). Collectively, these findings highlight the putative role of specific SECRC in MM progression and drug resistance and the cooperativity between these SEs in altering large gene expression programs.

**SECRCs are essential for MM biology.** We have demonstrated that the expression of certain SECRCs are associated with MM progression and refractory disease in primary MM samples. Using cell lines as surrogates for active MM disease, their dependence on SECRC expression was investigated through CRISPR knock-down (KD) of each of the 60 SECRC genes in the fitness of a panel of 20 MM human cell lines from DepMap portal<sup>59</sup> (**Figure 5g**). CRISPR KD-mediated suppression of genes such as *EBF1*, *SPI1*, *ZEB2*, *FOXP1*, *RB1* and *KLF6*, had minimum effect in MM cell lines' fitness. In contrast, *MZF1*, *IRF4*, *SMARCA5*, *PRDM1*, *YY1*, *ZNF217*, *RFX5*, *CTCF* and *IRF4* had higher relative expression in MM cell lines, and significantly reduced fitness following KD. Additionally, despite missing the FDR cut-off of statistical differential expression in LRMM, *MYC* was highly expressed in MM cell lines, and its KD was among the most negatively impactful in cell line fitness (**Figure 5g**). Thus, KD of SECRCs under-expressed in NDMM (as compared to MGUS/SMOL) did not affect cell fitness. In contrast, KD of SECRCs over-expressed in NDMM or LRMM (as compared to MGUS/SMOL) significantly reduced fitness in MM cell lines. Thus, despite the limitations in representing primary MM through cell lines<sup>60</sup>, these results indicate that the activation of these SECRCs is conserved across MM cell lines and primary samples alike. To this end, these SECRCs are mechanisms of epigenetic regulation of the survival and a proliferative state<sup>61</sup> and potentially serve as therapeutic targets.

**Multifactorial influences of the MM epigenetic landscape.** Our data indicated that transcriptional reprogramming follows an evolving epigenetic landscape. Epigenetics has been linked to the high-risk translocation of IgH enhancer-driven expression of the methyltransferase *NSD2/MMSET* in t(4;14)-positive tumors<sup>38,62</sup>. Further, MM progression is associated with accumulation of cytogenetic abnormalities and mutations in putative driver genes<sup>5</sup>. Therefore, we also examined changes in mutational load of individual samples and mutational frequency of individual genes associated with progression (MGUS, n=64; SMOL, n=65; and NDMM, n=199) and refractory (ERMM, n=342; and LRMM, n=146) disease states (**Extended Data Figure 9**). In agreement with previous studies<sup>5</sup>, we observed increased mutational load per sample, and mutational frequency of putative “driver genes”, across disease state (**Figure 6a-b**). Only two of these genes demonstrated statistically different frequency between two disease states, KRAS and NRAS, between MGUS and SMOL (**Extended Data Figure 10a-c**)<sup>63</sup>, although neither could solely account for disease progression and refractory disease. These observations led us to investigate whether well-established genetic aberrations in MM could converge within a common transcriptional dysregulation mechanism, thus defining a common “driver biology”. **Figure 6c-d** and **Extended Data Figure 10d-k** illustrate the differentially expressed genes and gene clusters in samples positive or negative for frequently mutated genes and cytogenetic abnormalities. Notably, mutations in *NRAS* and *KRAS*, as well as del13q, correlated with differential expression associated with transition from pre-malignant to active disease, while *TP53* mutation or loss in del17p was associated with refractory disease. In contrast, amp1q21 is associated with differential expression of both transitions, suggesting it contributes to both pre-malignant and refractory disease. We have confirmed that amp1q21, a marker of high-risk MM and disease progression<sup>64</sup>, correlated with over-expression of *RFX5* and *MCL1*, both genes being located in 1q21 region, although samples without amp1q21 also show comparable expression of these genes, suggesting amp1q21 as a sufficient but not necessary condition for *RFX5* over-expression (**Extended Data Figure 10l**). Additionally, scATAC-seq analysis of LRMM samples demonstrates dramatic increase in chromatin accessibility of *RFX5*, but only moderate for *MCL1* (**Figure 5c**), indicating the former’s association with SE activation, but not the latter’s. t(11;14) was associated with a unique pattern of differential expression, but which did not match either pattern identified for progression or refractory disease, confirming its neutral prognostic impact<sup>65</sup>. Moreover, mutations in large genes such as *TTN*, *MUC4* and *MUC16*, despite increased frequency across MM spectrum (**Extended Data Figure 9b**), did not significantly correlate with transcriptional reprogramming associated with disease progression or refractory disease, serving as a negative control for this analysis (**Figure 6c**, **Extended Data Figure 10 j-k**).

## Discussion

Integrating WES, RNA-seq and scATAC-seq of a new cohort of MM patient specimens, we have shown that a unifying SECRC control of distinct epigenetic transcriptional programs drives MM progression and refractory disease. Despite being associated with increased mutational load per sample, as well as

increased mutational frequency of putative driver genes (e.g. KRAS, NRAS, TP53)<sup>66</sup>, adaptations in MM biology between MGUS-to-NDMM and NDMM-to-LRMM states are fundamentally implemented by transcriptional reprogramming. The transition from pre-malignant to active MM is dominated by reduced expression of genes linked to cell identity, cell adhesion and inflammation, whereas hallmarks of refractory MM are associated with cell cycle, DNA repair, metabolism, protein and RNA synthesis and degradation. Underscoring the importance of these transcriptional programs is the fact that they are prognostic in defining time to progression of SMOL patients and overall survival of patients with active disease.

Notably, integrating the topology of MM transcriptional control and pathway analysis, revealed that epigenetic-mediated changes in chromatin accessibility are the driving mechanism of transcriptional changes observed in MM progression and refractory disease. Specifically, there is an enrichment for H3K27me3 and PRC2-involved DBPs (SUZ12 and EZH2) in under-expressed genes that characterize the MGUS-to-NDMM transition, while H3K27ac and TAD-forming DBPs (YY1, CTCF, GABPA and NRF1) are enriched in the progression to refractory disease. ScATAC-seq analyses of primary MM samples across MM spectrum confirmed these conclusions and revealed that the two families of motifs are inversely correlated with chromatin accessibility patterns, where (1) those of the OSN factors POU5F1 and SOX2 (among others) have increased accessibility from SMOL to NDMM and a decrease in therapy refractory disease, and (2) those of CTCF, YY1, NRF1 (among others) that direct the formation of topologically associating domains (TADs) in chromatin, had the inverse behavior. Given the roles of OSN factors in inducing iPSCs, and those of TADs in establishing SEs and cell identity, we propose that the progression from pre-malignant towards refractory disease should be described as “dysdifferentiation”, a term previously proposed to explain loss of cell identity in cancer cells<sup>67</sup>, where pre-existing epigenetic mechanisms of transcriptional control are hijacked through oncogenic events (e.g., mutations, translocations, SECRC dysregulation, etc.).

Mechanistically, transcriptional modeling and CRISPR fitness screens revealed that select SECRCs drive and sustain MM states. Such stable self-regulating transcriptional circuits provide a form of epigenetic memory<sup>52</sup>, and MM progression and refractory disease were specifically associated with SECRCs that are either under-expressed (e.g. EBF1, ZEB2, during MM progression) or overexpressed (e.g., IRF4, PRDM1 during MM progression and RFX5, YY1, CTCF in the transition to LRMM). Notably, publicly available CRISPR screens of 20 human MM cell lines revealed that NDMM-to-LRMM associated SECRCs were necessary for cell fitness, suggesting targeting such SECRCs as a therapeutic strategy to overcome the emergence of refractory disease.

Pre-malignant MM cells are limited to the perivascular niche<sup>68</sup>, but as they acquire independence from the TME, they expand into other niches of the marrow (e.g. endosteum, interstitium), leading to increased tumor burden<sup>69</sup> and active MM. TME-independence is acquired by a process akin to EMT in epithelial cells<sup>70</sup>, involving loss of hematopoietic cell markers and adhesion molecules. Following the initiation of

treatment, MM cells are selected for two traits: resistance to therapy and accelerated growth, where the fastest growing clones among the therapy-resistant cells re-populate the tumor during relapse.

The role of MM well-established mutations, and cytogenetic abnormalities, in regulation of epigenetic modifications in histones has been previously characterized in different cancers and iPSC<sup>71,72</sup>. More specifically, the OSN core transcriptional network is modulated by TGF- $\beta$ , MAPK, Akt-PI3K and Wnt signaling pathways, among others<sup>73-76</sup>. Here, we have shown the correlation between the occurrence of these genetic events and transcriptional dysregulation associated with MM progression and refractory disease. Collectively, our findings support a unified model of MM progression and emergence of refractory disease, where niche and therapy adaptations are result of genome-wide transcriptional reprogramming via hijacking of OSN, TAD and SECRC networks, which can be provoked by combinations of mutations, cytogenetic abnormalities, or environmental stress-induced epigenetic dysregulation<sup>77</sup> (**Figure 6d-f**). Among the new putative targets identified herein was RFX5, a TF that was originally identified in *MHCII* transcriptional regulation<sup>78,79</sup> yet that also controls telomere maintenance<sup>21,22</sup>. Critically, our data indicate that clinical management of MM patients should include therapeutic targeting of altered transcriptional programming, to both mitigate malignant transformation and/or prevent development of drug resistance, and that the significant inter-patient transcriptional heterogeneity manifest within clinical classifications, such as SMOL and therapy naïve MM, will require precision therapeutic approaches.

## Declarations

### Acknowledgements

The authors sincerely thank our multiple myeloma patients and their families for donating their samples for research purposes. We also thank the members of the Shain, Silva and Cleveland labs, and the Multiple Myeloma Working Group for scientific discussions. We are also grateful for the expert assistance and service of the Proteomics & Metabolomics, Molecular Genomics, Translational Research, Comparative Medicine, Flow Cytometry, and Biostatistics and Bioinformatics Cores of Moffitt Cancer Center, and Cores of M2Gen. This work was supported in parts by the H. Lee Moffitt Cancer Center Physical Sciences in Oncology (PSOC) Grant 1U54CA193489-01A1 (to K.H.S., A.S.S., R.J.G, J.L.C., R.A.G), H. Lee Moffitt Cancer Center's Team Science Grant (A.S.S., K.H.S.), Miles for Moffitt Foundation (A.S.S.), the Cortner-Couch Endowed Chair for Cancer Research from the University of South Florida School of Medicine (J.L.C.), 1R01CA195727 Targeting CD44-mediated calcium signaling for the treatment of relapsed myeloma (L.H. and A.S.S.), the Cancer Center Support Grant P30-CA076292 to the Moffitt Cancer Center, and personal donations from the Pentecost Family Foundation (to R.B., K.H.S. A.S.S.) and Lisa France Kennedy (J.L.C.). Access to primary cells was made possible through the Total Cancer Care Protocol at the Moffitt Cancer Center. The funding agencies have not played any role in study design, data collection, data analysis, interpretation, writing of the report, and the decision to submit it for publication.

## Author Contributions

A.S. and K.H.S. coordinated the study. A.S., K.H.S., R.B., R.J.G., W.S.D., D.M.S., J.C., J.B., T.N., M.A., R.A.G., L.H., J.T. and B.D.S. have developed the concept of the study and generation of the cohort used in this study. G.D.A., R.R.A., A.T., P.R.S. have conducted Electronic Health Records collection, abstraction, structuring and statistical analysis of clinical variables. R.R.C., E.T.B. and A.S. have conducted enrichment analysis and SECRC network modeling. R.R.C., M.B.M., M.C.S., P.R.S., K.L.B. have designed/conducted experiments. P.R.S., J.T., E.W. have conducted quality control analysis of RNAseq data. A.K., O.H., Z.J., H.D. have conducted processing of RNA-seq/WES data pipeline. S.Y. conducted scATAC-seq data analysis in 10x Genomics pipeline. C.L.C. conducted bone marrow aspirate processing and isolation of MM cells. R.R.C., M.B.M., A.S., J.L.C., R.J.G. and K.H.S. wrote the manuscript. All authors have read, reviewed and approved the manuscript.

## Competing Interests

R.B. reports research funding to the institution from Janssen, Oncopeptides, Karyopharm, GlaxoSmithKline, Sanofi, Abbvie, Bristol Myers Squibb, outside the submitted work; Also, member of advisory board for Janssen, Karyopharm, Abbvie, Bristol Myers Squibb, Sanofi; J.B. reports research funding to the institution from Bristol Myers Squibb and WindMIL Therapeutics, outside the submitted work; Also, member of advisory board for Janssen, and speakers bureau for Amgen, Bristol Myers Squibb, Janssen; M.A. is a member of the advisory board for GlaxoSmithKline and Celgene, and speakers bureau for Janssen and GlaxoSmithKline; N.T. reports research funding to the institution from Novartis and Karyopharm; K.H.S. reports honoraria from Bristol Myers Squibb, Celgene, Janssen, Adaptive, Sanofi, and Takeda, and research funding to the institution from AbbVie and Karyopharm, outside the submitted work. Also, member of advisory board for Celgene, Sanofi, and Bristol Myers Squibb, and speaker's bureau for Celgene, GlaxoSmithKline, Bristol Myers Squibb, Sanofi, Karyopharm, Takeda, Janssen, and Amgen; A.K., O.H., Z.J., H.D., W.S.D. are or were employed by M2Gen at the time of writing of this manuscript, which is part of ORIEN/AVATAR consortium. W.S.D. receives royalties from patents owned by the Moffitt Cancer Center and licensed to M2Gen. The other authors have nothing to disclose.

## Additional Information

**Supplementary Information** is available for this paper online.

**Correspondence and requests for materials** should be addressed to K.H.S. and A.S.S.

**Biological materials availability:** All analysis was performed on primary samples from Moffitt Cancer Center patients enrolled in the ORIEN/AVATAR consortium study, and thus are not publicly available.

**Code availability:** Not custom software was developed for this study, except for scripts used for data formatting, filtering and sorting. All other software from third parties were cited with version details in the Materials and Methods section. Scripts used for data formatting, filtering and parsing, as well as calls to MATLAB and R libraries are available upon request.

**Data availability:** The following data is available as supplemental material in this publication:

1. Z-normalized log-transformed FPKM RNAseq data from 844 bone marrow aspirates, used in this study;
2. Maf file with mutational data and clinical classification of MM samples used in this study;
3. FISH/cytogenetics data for samples from this study;
4. A table with de-identified information, matching clinical status, RNAseq, mutation and FISH/cytogenetics data;
5. scATAC-seq data in the form of Loupe Browser software file is available in Moffitt Cancer Center's U54 PS-ON/CSBC portal (<https://www.doi.org/10.7303/syn23595070>).

## References

- 1 Furukawa, Y. & Kikuchi, J. Epigenetic mechanisms of cell adhesion-mediated drug resistance in multiple myeloma. *Int J Hematol* **104**, 281-292, doi:10.1007/s12185-016-2048-5 (2016).
- 2 Meads, M. B., Gatenby, R. A. & Dalton, W. S. Environment-mediated drug resistance: a major contributor to minimal residual disease. *Nature reviews. Cancer* **9**, 665-674, doi:10.1038/nrc2714 (2009).
- 3 Mayo, A. & Reeder, C. Risk Stratification in Multiple Myeloma: Putting the Pieces Together. *J Adv Pract Oncol* **7**, 318-321 (2016).
- 4 Rajkumar, S. V. Multiple myeloma: 2016 update on diagnosis, risk-stratification, and management. *Am J Hematol* **91**, 719-734, doi:10.1002/ajh.24402 (2016).
- 5 Kumar, S. K. *et al.* Multiple myeloma. *Nat Rev Dis Primers* **3**, 17046, doi:10.1038/nrdp.2017.46 (2017).
- 6 Bustoros, M. *et al.* Genomic Profiling of Smoldering Multiple Myeloma Identifies Patients at a High Risk of Disease Progression. *Journal of Clinical Oncology* **38**, 2380+, doi:10.1200/Jco.20.00437 (2020).
- 7 Agirre, X. *et al.* Whole-epigenome analysis in multiple myeloma reveals DNA hypermethylation of B cell-specific enhancers. *Genome Research* **25**, 478-487, doi:10.1101/gr.180240.114 (2015).
- 8 Landau, H. J. *et al.* Accelerated single cell seeding in relapsed multiple myeloma. *Nature Communications* **11**, doi:ARTN 3617 10.1038/s41467-020-17459-z (2020).

- 9 Maura, F. *et al.* Genomic landscape and chronological reconstruction of driver events in multiple myeloma. *Nature Communications* **10**, doi:ARTN 3835  
10.1038/s41467-019-11680-1 (2019).
- 10 Maura, F., Landgren, O. & Morgan, G. J. Designing Evolutionary Based Interception Strategies to Block the Transition from Precursor Phases to Multiple Myeloma. *Clin Cancer Res*, doi:10.1158/1078-0432.CCR-20-1395 (2020).
- 11 Dalton, W. S. The "total cancer care" concept: linking technology and health care. *Cancer Control* **12**, 140-141, doi:10.1177/107327480501200211 (2005).
- 12 Liu, X. *et al.* Yamanaka factors critically regulate the developmental signaling network in mouse embryonic stem cells. *Cell Res* **18**, 1177-1189, doi:10.1038/cr.2008.309 (2008).
- 13 Zhan, F. *et al.* Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood* **99**, 1745-1757, doi:10.1182/blood.v99.5.1745 (2002).
- 14 Nechanitzky, R. *et al.* Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. *Nat Immunol* **14**, 867+, doi:10.1038/ni.2641 (2013).
- 15 Buffa, S. *et al.* A novel B cell population revealed by a CD38/CD24 gating strategy: CD38(-)CD24 (-) B cells in centenarian offspring and elderly people. *Age (Dordr)* **35**, 2009-2024, doi:10.1007/s11357-012-9488-5 (2013).
- 16 Alaterre, E. *et al.* CD24, CD27, CD36 and CD302 gene expression for outcome prediction in patients with multiple myeloma. *Oncotarget* **8**, 98931-98944, doi:10.18632/oncotarget.22131 (2017).
- 17 Wischhof, L. *et al.* The SWI/SNF subunit Bcl7a contributes to motor coordination and Purkinje cell function. *Sci Rep* **7**, 17055, doi:10.1038/s41598-017-17284-3 (2017).
- 18 Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16**, 178-189, doi:10.1038/nrm3941 (2015).
- 19 Ma, C. & Staudt, L. M. LAF-4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF-4, the gene fused to MLL in t(4;11) leukemias. *Blood* **87**, 734-745, doi:DOI 10.1182/blood.V87.2.734.bloodjournal872734 (1996).
- 20 Lefevre, L. *et al.* Combined transcriptome studies identify AFF3 as a mediator of the oncogenic effects of beta-catenin in adrenocortical carcinoma. *Oncogenesis* **4**, doi:10.1038/oncsis.2015.20 (2015).

- 21 Chen, D. B. *et al.* Regulatory factor X5 promotes hepatocellular carcinoma progression by transactivating tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta and suppressing apoptosis. *Chinese Med J-Peking* **132**, 1572-1581, doi:10.1097/Cm9.0000000000000296 (2019).
- 22 de la Guardia, R. D. *et al.* Expression profile of telomere-associated genes in multiple myeloma. *Journal of Cellular and Molecular Medicine* **16**, 3009-3021, doi:10.1111/j.1582-4934.2012.01628.x (2012).
- 23 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).
- 24 Meads, M. B., Hazlehurst, L. A. & Dalton, W. S. The bone marrow microenvironment as a tumor sanctuary and contributor to drug resistance. *Clinical Cancer Research* **14**, 2519-2526, doi:10.1158/1078-0432.Ccr-07-2223 (2008).
- 25 Jimenez-Sanchez, A. *et al.* Unraveling tumor-immune heterogeneity in advanced ovarian cancer uncovers immunogenic effect of chemotherapy. *Nat Genet* **52**, 582-593, doi:10.1038/s41588-020-0630-5 (2020).
- 26 van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J Mach Learn Res* **9**, 2579-2605 (2008).
- 27 Wang, P. H. Pattern-Recognition with Fuzzy Objective Function Algorithms - Bezdek,Jc. *Siam Rev* **25**, 442-442 (1983).
- 28 Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128, doi:10.1186/1471-2105-14-128 (2013).
- 29 Nepf, S. *et al.* An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* **489**, 83-90, doi:10.1038/nature11212 (2012).
- 30 Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343-349, doi:10.1038/nature09784 (2011).
- 31 Scharer, C. D., Barwick, B. G., Guo, M. Y., Bally, A. P. R. & Boss, J. M. Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. *Nature Communications* **9**, doi:ARTN 1698 10.1038/s41467-018-04125-8 (2018).
- 32 Horikawa, I. & Barrett, J. C. Transcriptional regulation of the telomerase hTERT gene as a target for cellular and viral oncogenic mechanisms. *Carcinogenesis* **24**, 1167-1176, doi:10.1093/carcin/bgg085

(2003).

- 33 Corces, M. R. *et al.* Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature Genetics* **48**, 1193-1203, doi:10.1038/ng.3646 (2016).
- 34 Weintraub, A. S. *et al.* YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**, 1573+, doi:10.1016/j.cell.2017.11.008 (2017).
- 35 Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-395, doi:10.1038/cr.2011.22 (2011).
- 36 Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Mol Cell* **49**, 825-837, doi:10.1016/j.molcel.2013.01.038 (2013).
- 37 Jin, Y. *et al.* Active enhancer and chromatin accessibility landscapes chart the regulatory network of primary multiple myeloma. *Blood* **131**, 2138-2150, doi:10.1182/blood-2017-09-808063 (2018).
- 38 Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science* **357**, doi:ARTN eaal2380 10.1126/science.aal2380 (2017).
- 39 Corces, M. R. *et al.* The chromatin accessibility landscape of primary human cancers. *Science* **362**, doi:10.1126/science.aav1898 (2018).
- 40 LaFave, L. M. *et al.* Epigenomic State Transitions Characterize Tumor Progression in Mouse Lung Adenocarcinoma. *Cancer Cell* **38**, 212-228 e213, doi:10.1016/j.ccell.2020.06.006 (2020).
- 41 Zhao, X. *et al.* Unification of de novo and acquired ibrutinib resistance in mantle cell lymphoma. *Nat Commun* **8**, 14920, doi:10.1038/ncomms14920 (2017).
- 42 Zhao, X. *et al.* BCL2 Amplicon Loss and Transcriptional Remodeling Drives ABT-199 Resistance in B Cell Lymphoma Models. *Cancer Cell* **35**, 752-766 e759, doi:10.1016/j.ccell.2019.04.005 (2019).
- 43 Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev* **23**, 2484-2489, doi:10.1101/gad.1837309 (2009).
- 44 Yeap, L. S., Hayashi, K. & Surani, M. A. ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. *Epigenetics Chromatin* **2**, 12, doi:10.1186/1756-8935-2-12 (2009).
- 45 Yuan, P. *et al.* Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev* **23**, 2507-2520, doi:10.1101/gad.1831909 (2009).

- 46 Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762-767, doi:10.1038/nature08398 (2009).
- 47 Liu, Z., Scannell, D. R., Eisen, M. B. & Tjian, R. Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* **146**, 720-731, doi:10.1016/j.cell.2011.08.005 (2011).
- 48 Pliner, H. A. *et al.* Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. *Mol Cell* **71**, 858-871 e858, doi:10.1016/j.molcel.2018.06.044 (2018).
- 49 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).
- 50 Saint-Andre, V. *et al.* Models of human core transcriptional regulatory circuitries. *Genome Res* **26**, 385-396, doi:10.1101/gr.197590.115 (2016).
- 51 Sengupta, S. & George, R. E. Super-Enhancer-Driven Transcriptional Dependencies in Cancer. *Trends Cancer* **3**, 269-281, doi:10.1016/j.trecan.2017.03.006 (2017).
- 52 Mayran, A. & Drouin, J. Pioneer transcription factors shape the epigenetic landscape. *Journal of Biological Chemistry* **293**, 13795-13804, doi:10.1074/jbc.R117.001232 (2018).
- 53 Boller, S. *et al.* Pioneering Activity of the C-Terminal Domain of EBF1 Shapes the Chromatin Landscape for B Cell Programming. *Immunity* **44**, 527-541, doi:10.1016/j.immuni.2016.02.021 (2016).
- 54 Treiber, T. *et al.* Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription- independent poising of chromatin. *Immunity* **32**, 714-725, doi:10.1016/j.immuni.2010.04.013 (2010).
- 55 Dyson, N. J. RB1: a prototype tumor suppressor and an enigma. *Gene Dev* **30**, 1492-1502, doi:10.1101/gad.282145.116 (2016).
- 56 Reeves, H. L. *et al.* Kruppel-Like factor 6 (KLF6) is a tumor-suppressor gene frequently inactivated in colorectal cancer. *Gastroenterology* **126**, 1090-1103, doi:10.1053/j.gastro.2004.01.005 (2004).
- 57 Tian, E. *et al.* Allelic mutations in noncoding genomic sequences construct novel transcription factor binding sites that promote gene overexpression. *Gene Chromosome Canc* **54**, 692-701, doi:10.1002/gcc.22280 (2015).
- 58 Chapuy, B. *et al.* Discovery and Characterization of Super-Enhancer-Associated Dependencies in Diffuse Large B Cell Lymphoma (vol 24, pg 777, 2013). *Cancer Cell* **25**, 545-546, doi:10.1016/j.ccr.2014.03.029 (2014).
- 59 Meyers, R. M. *et al.* Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature Genetics* **49**, 1779+, doi:10.1038/ng.3984

(2017).

60 Sarin, V. *et al.* Evaluating the efficacy of multiple myeloma cell lines as models for patient tumors via transcriptomic correlation analysis. *Leukemia* **34**, 2754-2765, doi:10.1038/s41375-020-0785-1 (2020).

61 Kuehl, W. M. & Bergsagel, P. L. Molecular pathogenesis of multiple myeloma and its premalignant precursor. *J Clin Invest* **122**, 3456-3463, doi:10.1172/JCI61188 (2012).

62 Kalff, A. & Spencer, A. The t(4;14) translocation and FGFR3 overexpression in multiple myeloma: prognostic implications and current clinical strategies. *Blood Cancer Journal* **2**, doi:ARTN e89 10.1038/bcj.2012.37 (2012).

63 Dutta, A. K. *et al.* Subclonal evolution in disease progression from MGUS/SMM to multiple myeloma is characterised by clonal stability. *Leukemia* **33**, 457-468, doi:10.1038/s41375-018-0206-x (2019).

64 Hanamura, I. *et al.* Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. *Blood* **108**, 1724-1732, doi:10.1182/blood-2006-03-009910 (2006).

65 Saxe, D., Seo, E. J., Bergeron, M. B. & Han, J. Y. Recent advances in cytogenetic characterization of multiple myeloma. *Int J Lab Hematol* **41**, 5-14, doi:10.1111/ijlh.12882 (2019).

66 Gatenby, R. A., Cunningham, J. J. & Brown, J. S. Evolutionary triage governs fitness in driver and passenger mutations and suggests targeting never mutations. *Nature Communications* **5**, doi:ARTN 5499 10.1038/ncomms6499 (2014).

67 Zs Nagy, I., Cutler, R. G. & Semsei, I. Dysdifferentiation Hypothesis of Aging and Cancer - a Comparison with the Membrane Hypothesis of Aging. *Ann Ny Acad Sci* **521**, 215-225, doi:DOI 10.1111/j.1749-6632.1988.tb35280.x (1988).

68 Sapoznikov, A. *et al.* Perivascular clusters of dendritic cells provide critical survival signals to B cells in bone marrow niches. *Nat Immunol* **9**, 388-395, doi:10.1038/ni1571 (2008).

69 Silva, A. S. & Gatenby, R. A. Adaptation to survival in germinal center is the initial step in onset of indolent stage of multiple myeloma. *Molecular pharmaceuticals* **8**, 2012-2020, doi:10.1021/mp200279p (2011).

70 Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J Clin Invest* **119**, 1420-1428, doi:10.1172/JCI39104 (2009).

- 71 Zhu, J. *et al.* Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. *Nature* **525**, 206-211, doi:10.1038/nature15251 (2015).
- 72 Hosogane, M., Funayama, R., Nishida, Y., Nagashima, T. & Nakayama, K. Ras-Induced Changes in H3K27me3 Occur after Those in Transcriptional Activity. *Plos Genetics* **9**, doi:ARTN e1003698 10.1371/journal.pgen.1003698 (2013).
- 73 Medvedev, S. P., Shevchenko, A. I. & Zakian, S. M. Molecular basis of Mammalian embryonic stem cell pluripotency and self-renewal. *Acta Naturae* **2**, 30-46 (2010).
- 74 James, D., Levine, A. J., Besser, D. & Hemmati-Brivanlou, A. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* **132**, 1273-1282, doi:10.1242/dev.01706 (2005).
- 75 Li, J., Li, J. & Chen, B. Oct4 was a novel target of Wnt signaling pathway. *Mol Cell Biochem* **362**, 233-240, doi:10.1007/s11010-011-1148-z (2012).
- 76 van Andel, H., Kocemba, K. A., Spaargaren, M. & Pals, S. T. Aberrant Wnt signaling in multiple myeloma: molecular mechanisms and targeting options. *Leukemia* **33**, 1063-1075, doi:10.1038/s41375-019-0404-1 (2019).
- 77 Prickaerts, P. *et al.* Hypoxia increases genome-wide bivalent epigenetic marking by specific gain of H3K27me3. *Epigenetics Chromatin* **9**, 46, doi:10.1186/s13072-016-0086-0 (2016).
- 78 Villard, J. *et al.* A functionally essential domain of RFX5 mediates activation of major histocompatibility complex class II promoters by promoting cooperative binding between RFX and NF-Y. *Molecular and Cellular Biology* **20**, 3364-3376, doi:Doi 10.1128/Mcb.20.10.3364-3376.2000 (2000).
- 79 Wright, K. L. & Ting, J. P. Epigenetic regulation of MHC-II and CIITA genes. *Trends Immunol* **27**, 405-412, doi:10.1016/j.it.2006.07.007 (2006).

## Figures

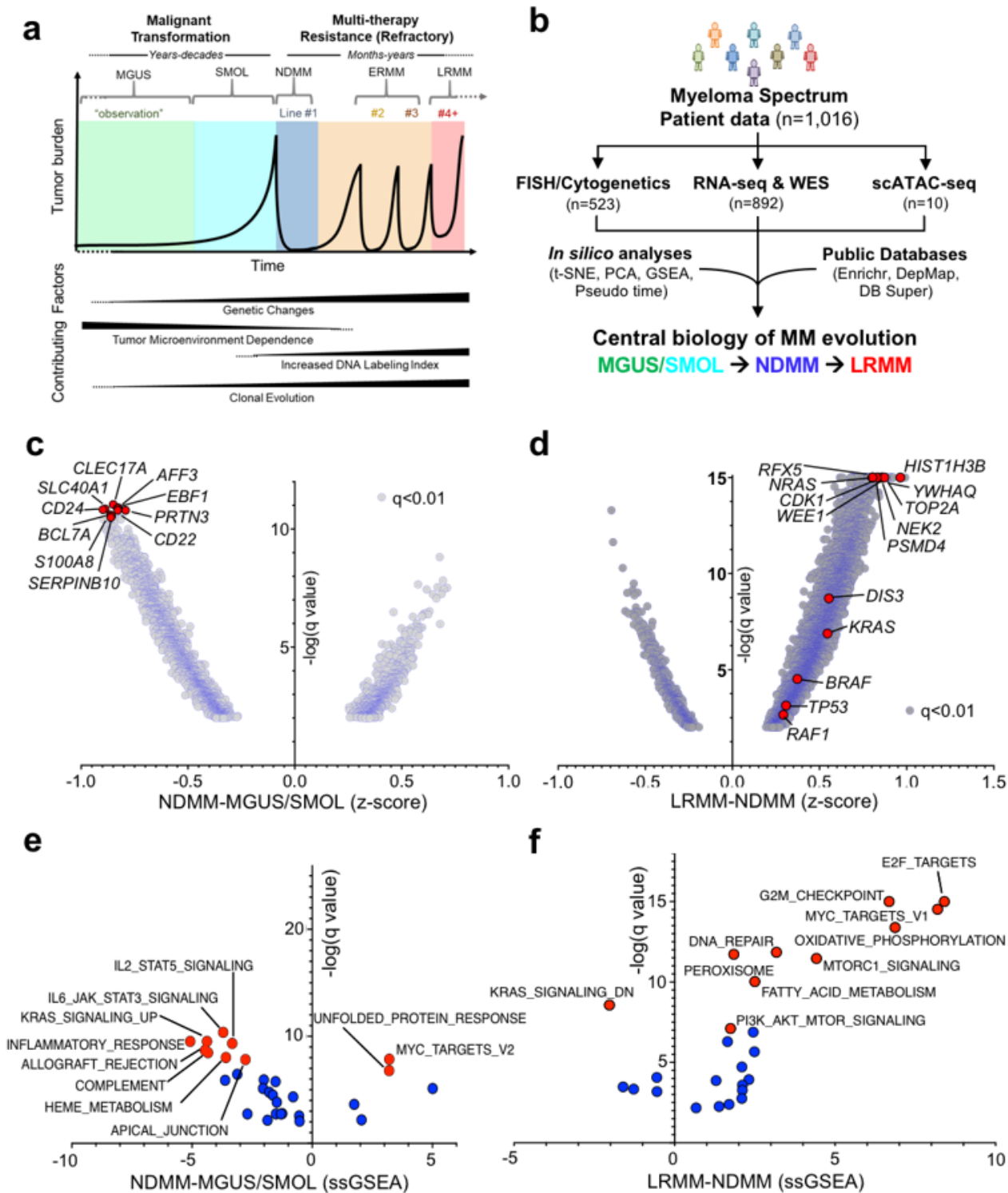


**Figure 1**

Differential transcriptional profile across MM spectrum. a, Model of MM clinical evolution suggests that transition from pre-malignant - monoclonal gammopathy of undetermined significance (MGUS) or smoldering MM (SMOL) - to active disease (newly diagnosed, NDMM) consists of loss of TME dependency and accumulation of genetic abnormalities. Successive lines of therapy response are followed by early relapses (ERMM, 1-3 lines of therapy) and eventually refractory disease (late relapse

MM, LRMM, >3 lines of therapy), characterized by high expression of cell cycle genes and increased genetic heterogeneity. While MM progression is selected by TME-imposed restrictions (i.e., limited availability of plasma cell-supporting niches in the marrow), refractory disease is the product of selection for the fastest growing therapy-refractory clones during relapse and tumor re-growth. b, We have explored the central biology involved in MM progression and refractory disease by leveraging a new database of combined clinical and molecular data, from a cohort of 1,016 bone marrow aspirates collected from patients across MM spectrum treated at Moffitt Cancer Center. c, Differential gene expression associated with MM progression or d, refractory disease. e, Differential relevance (based on enrichment score) of Cancer Hallmarks during MM progression or f, refractory disease. Differential gene expression was calculated on 16,738 expressed genes (z-normalized as  $\text{Log}_2(\text{FPKM}+10^{-3})$ ). Single sample gene set enrichment analysis (ssGSEA) was used to calculate the enrichment score of each gene set per sample. Unpaired two-sided t-tests with multiple test correction (q-value < 1%) were conducted in all comparisons above.

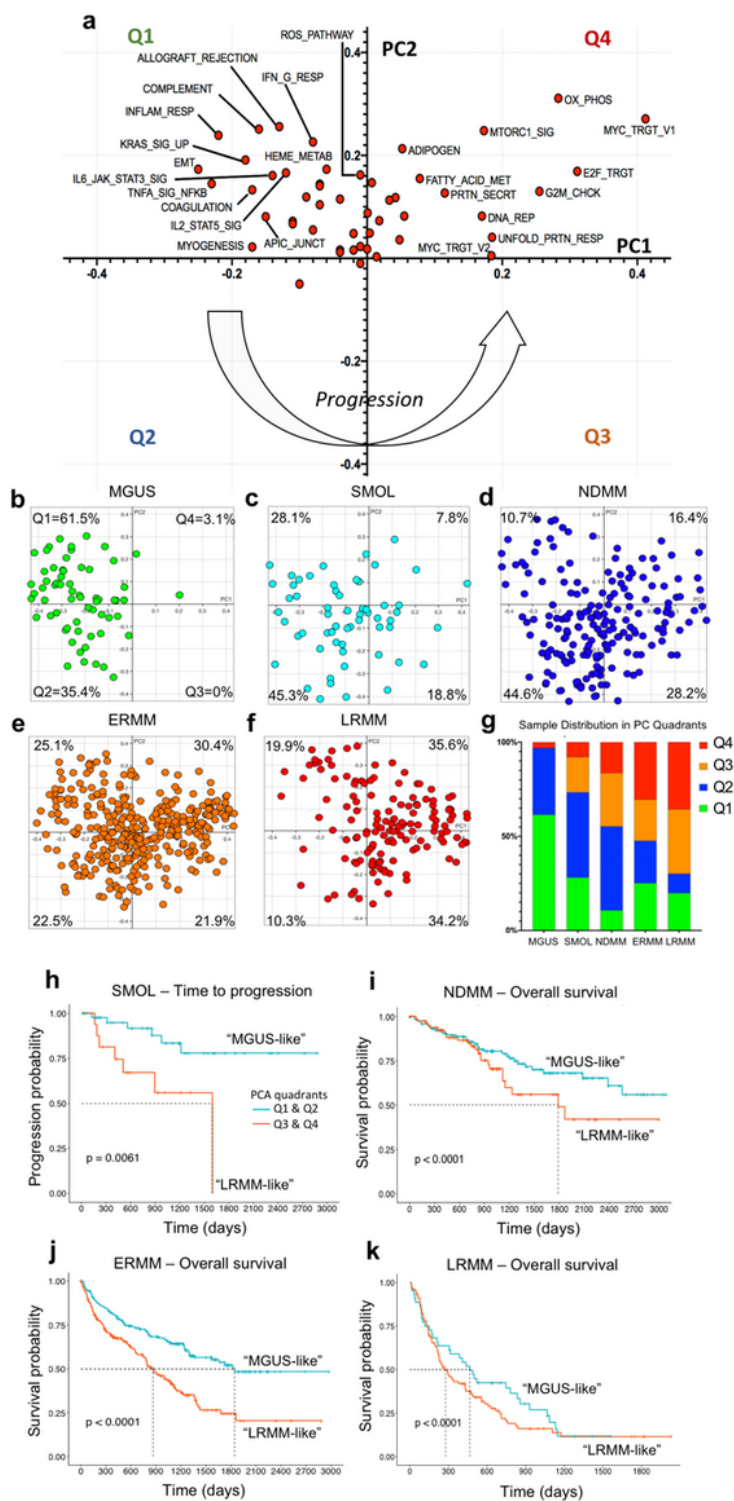


**Figure 2**

Unsupervised principal component analysis identifies biology driving MM and prognosis. Principal component analysis (PCA) was calculated on RNA-seq of 844 samples from patients across MM spectrum, using ssGSEA enrichment score of Cancer Hallmarks as variables. a, Projection of the Cancer Hallmarks' feature loadings indicate their relative contribution to the spatial distribution of the samples in the PCA. Curved arrow represents the proposed trend observed in the distribution of the samples across

disease states. b – f, Samples were displayed in 5 plots, according to disease state, to facilitate visualization. Percentages depict the proportion of samples in each quadrant. g, Distribution of samples from each disease state in the 4 PC quadrants annotated in a. MGUS samples are primarily located in quadrants Q1 and Q2, while LRMM samples are primarily located in quadrants Q3 and Q4. Samples in quadrants Q1 or Q2 were labeled as “MGUS-like” while those in quadrants Q3 and Q4 were labeled as “LRMM-like”. “MGUS-like” patients had a better prognosis than “LRMM-like”, as measured by (h) time to progression of SMOL patients, as well as overall survival of NDMM (i), ERMM (j) and LRMM (k) patients, indicating that the enrichment of specific cancer hallmarks has clinical implications.

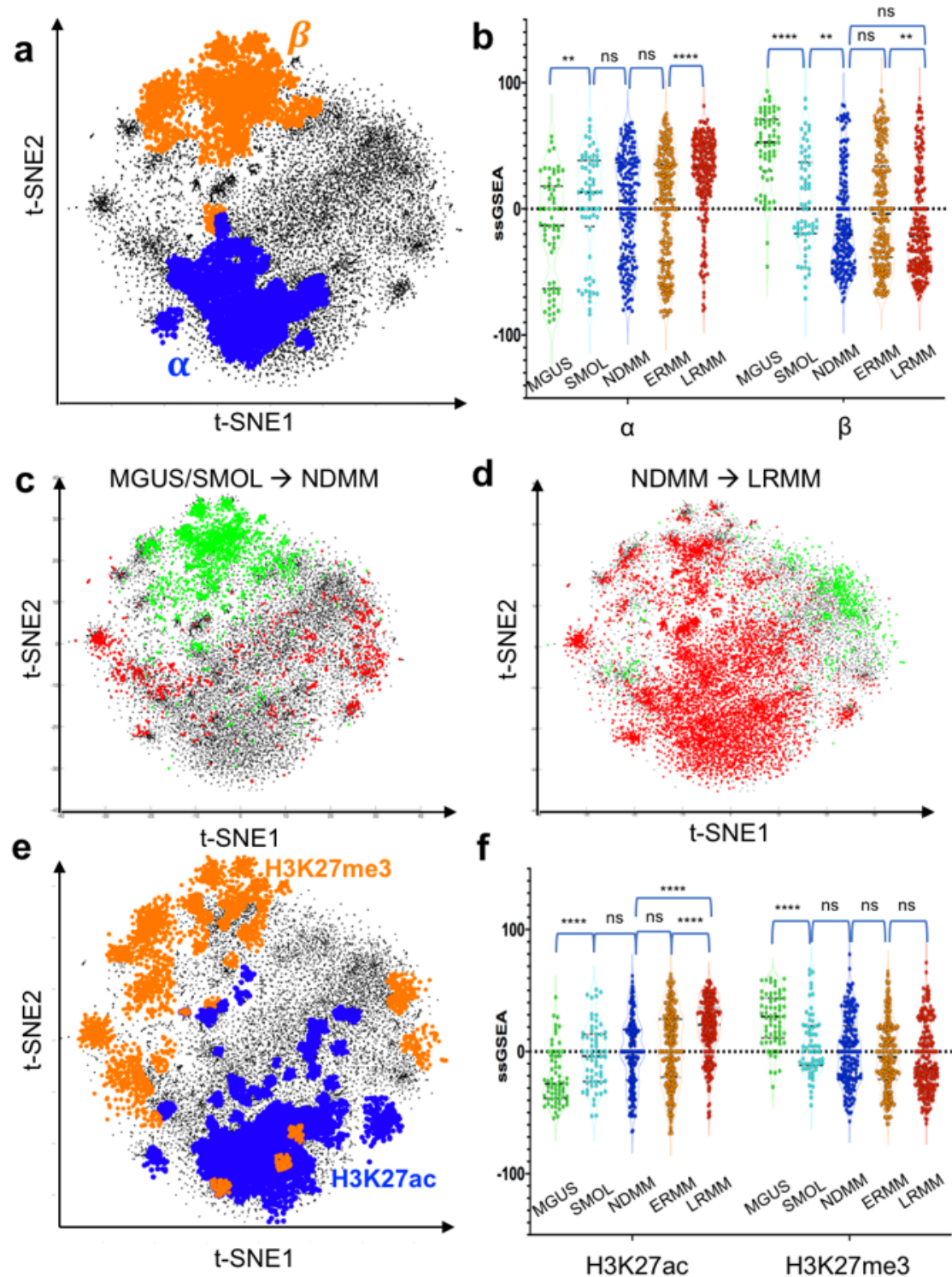


**Figure 3**

Topology of transcriptional regulation of MM biology and disease progression. a, Dimensionality reduction analysis (t-SNE), followed by unsupervised clustering (Fuzzy c-Means), identified 500 clusters of co-expressing genes, across the cohort of 844 RNA-seq MM samples. Unsupervised hierarchical clustering, based on Pearson correlation between ssGSEA of gene clusters and Cancer Hallmarks, identified two superclusters,  $\alpha$  and  $\beta$ , enriched for Hallmarks associated with i) energy metabolism

(OXPHOS and glycolysis, cell cycle, protein and RNA synthesis and degradation), and ii) immune response, EMT, angiogenesis, hypoxia and xenobiotic metabolism, respectively (see Extended Data Figures 2 and 3). b, ssGSEA of superclusters  $\alpha$  and  $\beta$  shows significant increase in expression of the former and decrease in the latter with MM progression and refractory disease. c, Location, in the transcriptional topology, of genes under- (in green) or overexpressed (in red) during MGUS/SMOL→NDMM transition. d, Same as in c, regarding NDMM→LRMM transition. e, Gene clusters individually enriched for histone modifications, according to public databases of Chip-seq experiments, co-localize in the t-SNE, suggesting that such epigenetic modifications are a major contributor to MM transcriptional regulation. f, ssGSEA of two gene sets, H3K27ac and H3K27me3, containing the genes in the clusters highlighted in e, show increase in expression of the former and decrease of the latter, suggesting that both histone alterations become more frequent during MM progression and refractory disease. \* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ .

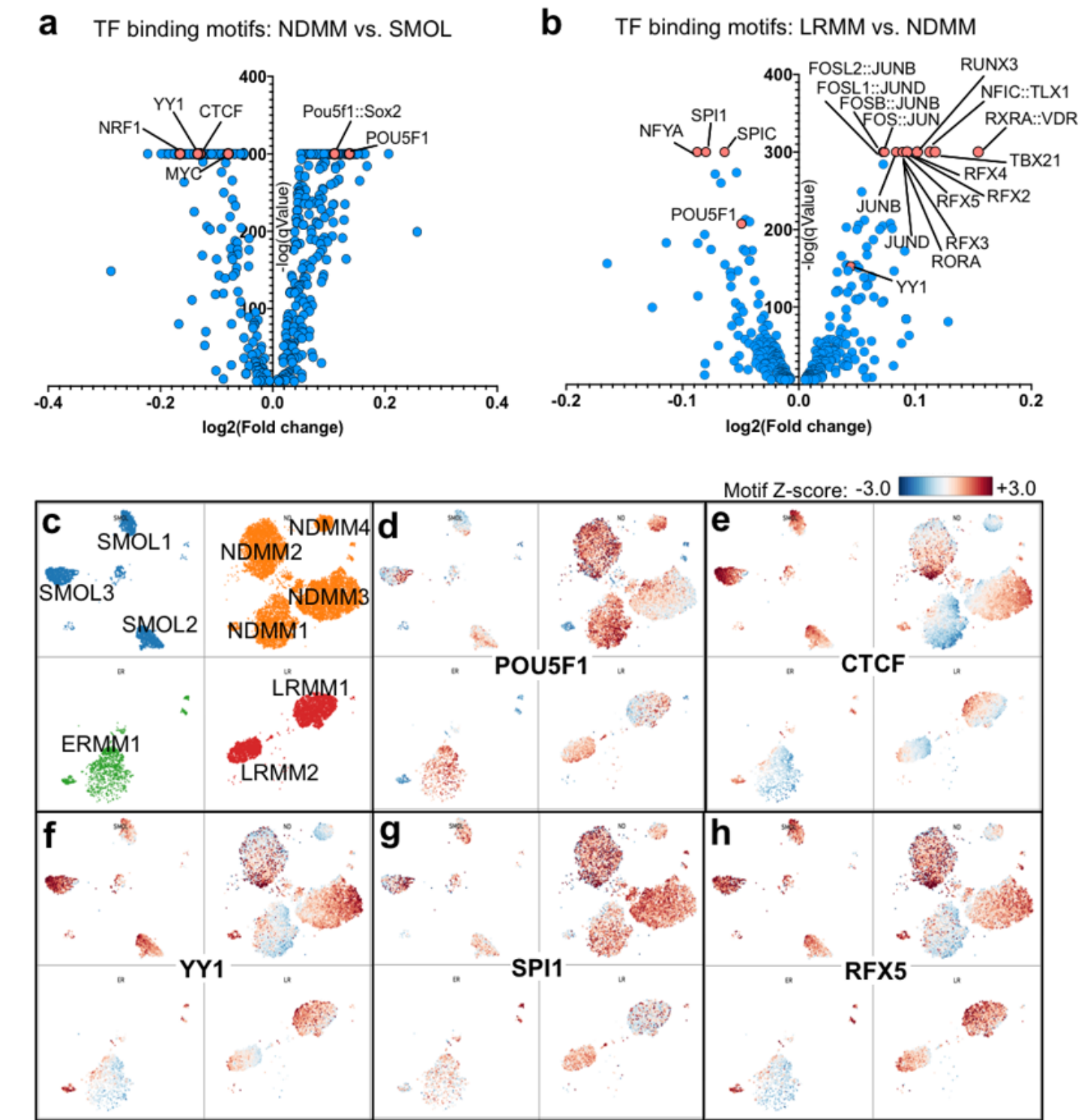


**Figure 4**

Differential chromatin accessibility of DNA-binding protein motifs evidences genome-wide epigenetic changes across MM spectrum. Single cell chromatin accessibility (scATAC-seq) was conducted on 10 MM samples (3 SMOL, 4 NDMM, 1 ERMM, and 2 LRMM), amounting to 34,961 cells. a and b, Differential chromatin accessibility of DNA binding protein (DBP) motifs in NDMM compared to SMOL, and in LRMM in relation to NDMM, respectively. c, scATAC-seq data visualization on Loupe software, depicting

individual samples grouped by disease state to improve visualization. Patient samples within each disease state were numbered accordingly. d – h, Z-normalized chromatin accessibility of the indicated DBP motifs across disease states highlights both inter- and intra-sample heterogeneity. Blue=low accessibility, brown=high accessibility)

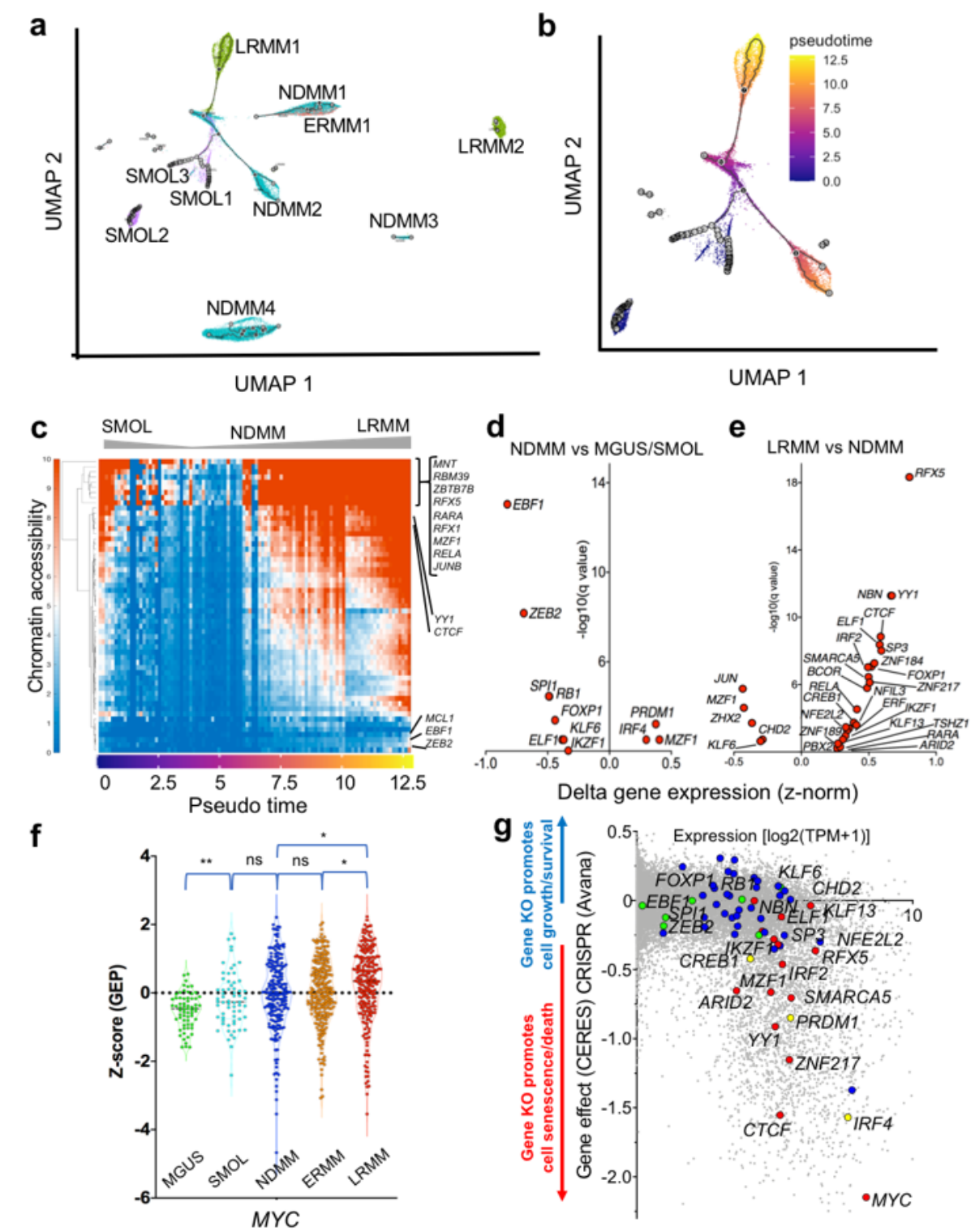


Figure 5

Pseudotime analysis of chromatin accessibility and modeling of transcriptional network identify essential MM SECRC. a, Pseudotime analysis was conducted on scATAC-seq data from 10 MM samples, based on overall (distal + promoter) chromatin accessibility of 1,883 genes whose differential expression across disease states was associated with changes in chromatin accessibility at promoter and distal levels. b, A pseudotime trajectory was generated starting in 3 SMOL, crossing 1 NDMM and ending in 1 LRMM sample. c, Chromatin accessibility of 60 SECRCs, in addition to MCL1, across the pseudotime trajectory depicted in b highlight those with significant increase in chromatin accessibility in LRMM (e.g. RFX5, RELA, JUNB) and those with unaltered low accessibility (e.g. ZEB2, EBF1). Differential expression of SECRC genes, across 844 RNAseq datasets, associated with MM progression (d), and refractory disease (e). f, Despite not passing the false discovery rate, the oncogenic SECRC MYC was overexpressed in LRMM vs. NDMM samples (MYC's  $q=0.015$ ). g, Expression (x-axis) and effect (y-axis) of silencing of 18,300 genes in 20 human MM cell lines confirms that under-expressed SECRC genes in NDMM (e.g., EBF1, ZEB2) have low expression, and low effect when silenced, in MM cell lines, while over-expressed SECRC genes in NDMM and LRMM (e.g., IRF4, RFX5, YY1) are among the most expressed and with highest negative effect when silenced in MM cell lines. Unpaired t-tests were used in d, e and f. Multiple test correction ( $q\text{-value} < 1\%$ ) was used in d and e.

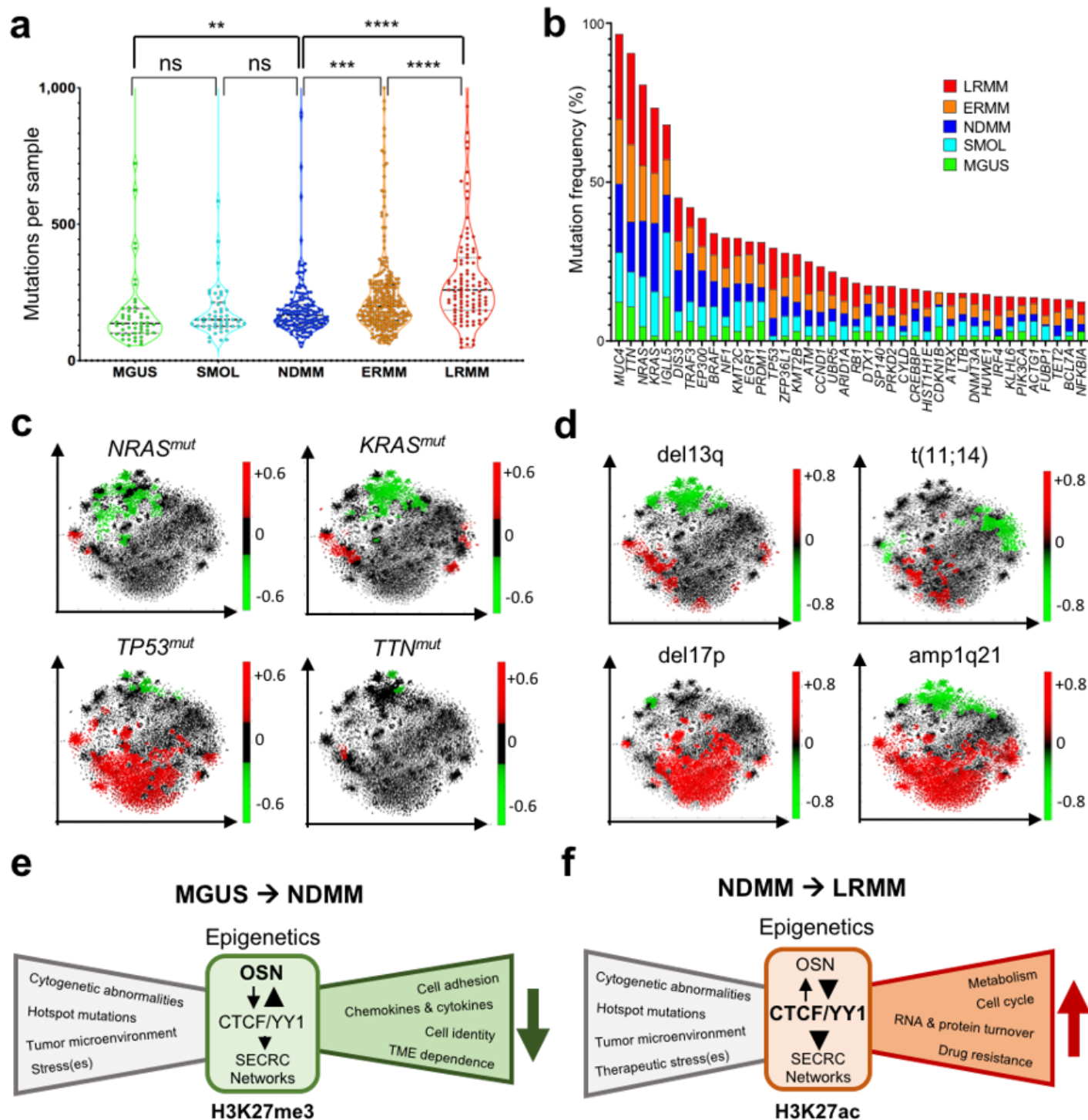


**Figure 6**

Transcriptional dysregulation association with mutations and cytogenetic abnormalities suggests a bow-tie model driving MM through epigenetic dysregulation. WES was performed in a cohort of 816 MM samples. **a**, Mutation load per sample showed across MM spectrum: dashed lines represent median and quartiles for each distribution. No significant difference in the mutation load was detected between pre-malignant states (MGUS and SMOL), but showed increase upon therapy initiation. **b**, Mutation frequency per disease state of the 40 most frequently driver mutations in MM, in addition to MUC4 and TTN (neutral

mutation controls). Differential gene expression (green=under-expression, red=overexpression) associated with hotspot mutations (c) or cytogenetic abnormalities (d). Events with equivalent biology (NRASmut and KRASmut; TP53mut and del17p) show similar transcriptional dysregulation. TTN, the most mutated gene in MM, was not associated with any specific signature, confirming its neutral nature in MM biology. e, A bow-tie model representing MM onset, in which transcriptional changes caused by several distinct inputs (e.g., genetic events and environmental stress) lead to the activation of core SECRCs associated with pluripotency (OSN), leading to genome-wide epigenetic-mediated transcriptional repression (H3K27me3), loss of differentiation and TME-independence. f, Additional genetic events, TME- and therapy-induced stress, cause (and/or select for) the activation of SECRCs (CTCF, YY1, NRF1) responsible for the formation of TADs, triggering the spread of epigenetic activators of gene expression (H3K27ac), leading to the expression of gene sets associated with cell cycle and metabolism in refractory disease (LRMM). Unpaired two-sided t-tests were conducted in a. \* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ .

## Supplementary Files

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

- [CanevaroloetalNatMedExtdataFigsFINAL.pdf](#)
- [SUPPTAB1.pdf](#)