RESEARCH

The DNA methylation landscape of multiple myeloma shows extensive inter- and intrapatient heterogeneity that fuels transcriptomic variability

Jennifer Derrien[1], Catherine Guérin-Charbonnel[1,2], Victor Gaborit[1,3], Loïc Campion[1,2], Magali Devic[1,4], Elise Douillard[1,4], Nathalie Roi[1,4], Hervé Avet-Loiseau[5], Olivier Decaux[6], Thierry Facon[7], Jan-Philipp Mallm[8], Roland Eils[9,10], Nikhil C. Munshi[11], Philippe Moreau[1,4], Carl Herrmann[9], Florence Magrangeas[1,4][†] and Stéphane Minvielle[1,4]*[†]

*Correspondence:
stephane.minvielle@univ-nantes.fr
[1]Université de Nantes, CNRS,
INSERM, CRCINA, F-44000
Nantes, France
Full list of author information is
available at the end of the article
[†]Co-senior

Abstract

Background: Cancer evolution depends on epigenetic and genetic diversity. Historically, in multiple myeloma (MM), subclonal diversity and tumor evolution have been investigated mostly from a genetic perspective.

Results: Here, we combined the notions of epipolymorphism and epiallele switching to analyze DNA methylation heterogeneity in MM patients. We show that MM is characterized by the continuous accumulation of stochastic methylation at the promoters of development-related genes. High entropy change is associated with poor outcomes and depends predominantly on partially methylated domains (PMDs). These PMDs, which represent the major source of inter- and intrapatient DNA methylation heterogeneity in MM, are linked to other key epigenetic aberrations, such as CpG island (CGI)/transcription start site (TSS) hypermethylation and H3K27me3 redistribution as well as 3D organization alterations. In addition, transcriptome analysis revealed that intratumor methylation heterogeneity was associated with low-level expression and high variability.

Conclusion: We propose that disordered methylation in MM is responsible for high epigenetic and transcriptomic instability allowing tumor cells to adapt to environmental changes by tapping into a pool of evolutionary trajectories.

Keywords: multiple myeloma; disordered DNA methylation; epipolymorphism; epiallele switching; inter- and intrapatient heterogeneity; transcriptomic variability

Background

Multiple myeloma (MM) is a neoplasm of plasma cells (PCs) with an incidence rate of approximately 5/100,000 in Europe. The median survival time of patients has improved substantially over the past decade. This is due to the establishment of high-dose therapy followed by autologous stem cell transplantation as a routine procedure, significant improvements in supportive care strategies, and the introduction and widespread use of drugs including immunomodulatory drugs, proteasome inhibitors, histone deacetylase inhibitors and monoclonal antibodies. Nevertheless, almost all patients ultimately relapse due to the emergence of more aggressive subpopulations of myeloma PCs resistant to therapeutic agents.

Several mechanisms have been suggested to explain the capacity of the subpopulations of myeloma cells within an individual to survive the pressure of frontline therapy and proliferate. These mechanisms include: the emergence of myeloma cells that achieve bortezomib resistance by decommitment from immunoglobulin synthesis [1] or by the derepression of growth factor receptors typically not associated with the plasma cell lineage [2], somatic mutations that emerge during disease progression involving key driver genes in MM such as the mono- or biallelic loss of TP53 [3, 4] or the biallelic loss of TRAF3 [5]. These mechanisms facilitate the expansion of proliferative subclonal populations. Genetic intratumor heterogeneity increases the evolutionary fitness potential of a rare subset of myeloma cells harboring a particular combination of molecular aberrations to survive when challenged by multiagent chemotherapy [4]. The disease evolves predominantly through a Darwinian process of clonal expansion, and the population of tumor PCs represents an admixture of competing genetic subclones [6, 5, 7, 8]. In most patients, the treatment pressure causes the profound reorganization and diversification of subclonal populations with complex dynamics of tumor evolution raising the possibility of biologically and clinically important cross-talk between subclones [9]. However, the small number of genetic alterations detected in MM and relevant in relapse mechanisms do not alone explain the profound phenotypic variability across patients [10, 4]. Beyond genetic diversity, other processes generate the intratumoral functional heterogeneity of cancer cells, including global epigenetic changes [11].

To address this question in MM, we conducted an analysis of the DNA methylation landscape of 47 myeloma samples from 26 patients. This study was motivated by the facts that (1) MM is characterized by extreme heterogeneity in median DNA methylation levels compared with normal plasma cells (NPCs) [12], (2) intratumor methylome variability is higher in various cancer types including other B-cell malignancies [13, 14, 15, 16], than in normal cells, and (3) genome-wide analysis of DNA methylation at single-base-pair resolution combined with analysis of the methylation states of neighboring CpGs (in the local sequence context) is now available [17].

Results

Intrapatient epigenetic heterogeneity arises from stochastic DNA methylation

To explore the global DNA methylation spectrum of MM patients, we compared whole-genome bisulfite sequencing (WGBS) datasets of myeloma cells from patients with those of controls (Additional file 1: Table S1). We found a marked global hypomethylation (Fig. 1a), as previously reported in MM and other cancers [12, 13, 15]. Individual CpG sites showed a predominantly bimodal pattern in MM patients (Fig. 1b) with methylation levels depending on the local density of CpG (Fig. 1c and Additional file 2: Table S2). Although MM displayed no evidence of aberrant DNA methylation at many genes (Additional file 3: Figure S1), we observed slight methylation gain at CpG islands (CGIs) and methylation loss at adjacent shores and shelves regions compared to NPCs (Fig. 1d), as illustrated in the genomic region of *DOC2B* (Fig. 1e).

These results show that the MM epigenome is characterized by global DNA hypomethylation and variable intermediate methylation states depending on the local

CpG content. We investigated whether these epigenetic changes could be a source of intrapatient heterogeneity. To assess the degree of epigenetic variability within patients, we employed enhanced reduced representation bisulfite sequencing (eR-RBS) technology on 21 MM patients because eRRBS provides an average depth of 50X per covered CpG and approximately 2.7 million CpGs per sample, with more than 10X sequencing coverage (Additional file 4: Table S3 and Additional file 5: Table S4), and is cost effective for a large number of samples (48 in our study, Additional file 1: Table S1 and Additional file 3: Figure S2). We then applied a computational method that analyzes DNA methylation modifications at a genomic locus defined as a group of four contiguous CpGs covered by single sequence reads [17]. We computed for each genomic locus the epipolymorphism, defined as a probability distribution of the 16 possible methylation patterns (epialleles): a high value of epipolymorphism indicates a stochastic process of DNA methylation [18] (Additional file 3: Figure S3). Almost all genomic loci were located in CGIs (85% in MM samples) and, to a lesser extent, in promoter transcription start sites (TSSs) and gene bodies (74% in MM samples, Additional file 3: Figure S4). We plotted epipolymorphism distribution as a function of methylation level for all loci in NPCs and MM samples (Fig. 2a). Remarkably, loci with modest methylation levels (5% - 25%) showed a higher degree of epipolymorphism in cancer cells than in control cells whereas at higher methylation levels, epipolymorphism distribution was similar in MM and NPCs, indicating enrichment of a fraction of loci-CGIs with a stochastic methylation state at the expense of a bimodal methylation profile in MM samples. In all cases, promoter CGI methylation gain was associated with an increase in epipolymorphism (Additional file 3: Figure S5).

To further characterize epiallele patterns of each sample, we stratified loci according to the frequency of fully methylated or fully unmethylated epialleles (Fig. 2b and Additional file 3: Figure S6). NPC loci displayed near-bimodal distribution, with a large number of homogeneous loci either fully unmethylated or fully methylated (Fig. 2b, top). Conversely, in MM samples, a large number of loci showing high epipolymorphism levels were along the axes, indicating marked enrichment for stochastic (i.e., neither fully methylated nor fully unmethylated) methylation patterns (Fig. 2b, bottom). Notably, this intrasample methylation heterogeneity was already present at the early stage of the disease, recognized as smoldering MM (SMM), and persisted at relapse (Additional file 3: Figure S6). A complementary approach that measures the proportion of discordant reads (PDR) [15] confirmed the increase in epigenetic heterogeneity in MM cells compared to NPCs. We found that the average PDR (i.e. genomic loci with a large number of reads that contain both unmethylated and methylated CpGs over the total number of reads) was significantly higher (p-value < 0.05 (DTK pairwise)) in MM samples, regardless of the disease stage, than in NPCs (Fig. 2c).

Overall, these results revealed that during disease initiation, growth and progression, malignant PCs accumulate randomness in DNA methylation at the expense of a more coherent methylation state, leading to a high degree of intratumoral epigenetic heterogeneity in MM patients.

Entropy changes in MM development

Since we have shown that stochastic DNA methylation gains lead to intratumor heterogeneity, we sought to analyze epiallelic dynamics during MM development and progression. We used an algorithm that allows the identification of epigenetic loci (eloci) that have a significant epiallele composition change (entropy change) between two states (e.g., normal vs cancer or diagnosis vs relapse). To compare global methylome changes across patient samples, the number of eloci was normalized according to the number of covered loci, referred to as eloci per million loci covered (EPM) (Additional file 3: Figure S7)[17]. We first examined the link between the degree of epiallelic shifting and outcome. For this purpose, patients were separated into two groups according to their EPM values. We found that MM patients with high EPM had reduced relapse-free survival (p-value = 0.02 (log-rank test); Fig. 3a). EPM was independent of age, sex and ploidy. Remarkably, the extent of entropy changes varied greatly between patients at diagnosis and persisted at relapse compared to NPCs (Fig. 3b, c). The genome-wide distribution of eloci showed a predominant location at CGIs (average 54% in MM vs NPCs). As a control, we determined eloci by comparing NPCs to NPCs (NPC eloci) and found, on average, 19% of eloci in CGIs (p-value (Wilcoxon) = 0.0018) and at promoter-TSS (average 15% in MM vs NPCs and average 3% in NPC eloci; p-value (Wilcoxon) = 0.0123) (Additional file 3: Figure S8). We observed a similar distribution of eloci in relapse samples compared to NPCs (p-value (Wilcoxon) = 0.75 (CGIs) and 0.82 (promoter-TSS), Additional file 3: Figure S8).

To gain insights into the biological functions of epigenetic perturbations, we performed gene ontology enrichment analysis for genes having eloci (MM diagnosis vs. NPCs) in promoter CGIs (Fig. 3d). Interestingly, associated genes showed marked enrichment for developmental regulators and transcription factor activity, while genes with promoter CGIs, including NPC eloci, were not enriched in any pathways or gene ontology terms. The identification of key signaling pathways rather than specific genes was not surprising considering the high DNA methylation variability across patients. Similar finding were reported by Easwaran et al. across different tumor types [19].

To further examine the impact of treatment on entropy changes, we assessed the epiallele shifting level during time to relapse by comparing diagnosis vs relapse paired samples (Fig. 3e). The degree of difference was highly variable from one patient to another; 24% of patients showed no epigenetic changes (EPM = 0) while 42% showed substantial changes (EPM > 1000) between diagnosis and relapse. We then focused our analysis on the evolution of extreme epiallele patterns (i.e., one predominant epiallele or a mixture of low represented epialleles) during two disease stages (i.e., before and after treatment). For this purpose, we compared epiallele shifts between NPCs and diagnosis samples and between paired samples at diagnosis and relapse. Interestingly, among the four possible epiallele pattern changes (Additional file 3: Figure S9), the selection pattern was significantly enriched at relapse compared to that at diagnosis (p-value (paired Wilcoxon) = 0.0059, Fig. 3f). Although this pattern involved a small number of eloci (median value = 4.2%), this finding suggests that treatment escape is associated with clonal selection at specific genomic loci.

Given the central role of chromatin in stabilizing gene expression and cellular states, we wanted to determined whether chromatin states are associated with genomic loci showing significant entropy changes in MM. To answer this question, we analyzed the distribution of eloci using 127 cell/tissue types at 25-state chromatin state segmentation [20] (Additional file 3: Figure S10a). Hypomethylated eloci were found in quiescent states while hypermethylated epiallele shifts were predominant in bivalent promoter-associated chromatin states (Additional file 3: Figure S10b). In addition, the frequency of these two subtypes of eloci varied widely between patients (Additional file 3: Figure S11).

Abnormal bivalent promoter eloci CGI methylation increases intra- and intertumor heterogeneity and targets Wnt and Ras/MAPK signaling

We found that eloci were predominantly enriched (82.5%) in the bivalent promoters of embryonic stem cells (ESCs) (Additional file 3: Figures S10a and S12). In the embryonic system, bivalent promoter genes are not regulated by DNA methylation but rather by the simultaneous presence of the repressive mark H3K27me3 and the active transcription mark H3K4me3, which allow low basal transcription states that are dynamically inducible to ensure a balance between self-renewal and lineage commitment [21]. We compared the average change in the DNA methylation of eloci between MM samples and NPCs. In patient M#17, who had the highest number of eloci, we observed that 93% of the eloci acquired extensive gains of DNA methylation (higher than 40%) during neoplastic transformation (Fig. 4a and Additional file 3: Figure S13). Methylation gains are maintained during progression (Fig. 4a). This perturbation was associated with a decrease in gene expression in diagnosis and relapse samples compared to NPCs, whereas upregulated genes were present mostly at the time of diagnosis in almost all patients (Fig. 4b and Additional file 3: Figure S14). To further examine the relationship between gene expression and entropy changes in MM patients, we compared the expression levels between bivalent genes containing hypermethylated eloci in their promoter and bivalent genes without eloci. We found that disrupted (i.e., containing hypermethylated eloci) bivalent genes were less expressed but displayed greater interpatient variability (Fig. 4c; only genes with a mean expression level above 1 were taken into account for the coefficient of variation). Gene set enrichment analysis revealed that disrupted bivalent genes were enriched for important cancer-related pathways, including the Wnt and MAPK signaling pathways (Fig. 4d).

Myeloma cells are characterized by predominant stochastic methylation gains at bivalent promoters reflected by a high PDR and epipolymorphism, leading to intratumor heterogeneity (Fig. 4e). We then analyzed the consequences of intratumoral methylation variability on bivalent gene expression levels across the 20 MM samples. We separated the genes into two groups according to the PDR level of their bivalent promoter (lower or higher than the mean PDR) and calculated, in each group and for each patient, the odds ratio (OR) of the association between gene expression (FPM > 1 vs. $\leq$ 1) and bivalent promoter methylation (mean methylation < 25% vs. mean methylation > 75%; Fig. 4f and Additional file 6: Table S5). In all MM samples, promoters with a low PDR tended to maintain the expected opposite relationship between promoter methylation and transcription, whereas in

promoters with a high PDR, for the majority of patients, the link between methylation and expression did not remain significant, and for some patients, we observed a significant link, but the relationship was opposite to the expected result (OR > 1). For example, *LIPG*, which showed comparable methylation levels in two samples (0.61 in M#14 and 0.63 in M#08) coupled with opposite expression levels (FPM of 4.57 in M#14 and 0.29 in M#08), demonstrates the decoupling relationship between promoter methylation and gene expression. M#14 displayed a high promoter PDR (0.80), whereas M#08 displayed a low promoter PDR (0.29, Additional file 3: Figure S15). OR analysis clearly showed the contribution of intratumoral methylation heterogeneity to increased transcriptional variability in MM. These results are concordant with those obtained in chronic lymphocytic leukemia (CLL) [15].

Entropy changes towards hypomethylation preferentially occur in PMDs

Hypomethylated epiallele shifts are localized mainly in quiescent genomic regions (Additional file 3: Figure S10). We identified in most of the MM samples two subsets of hypomethylated eloci compared to normal samples (Fig. 5a, b and Additional file 3: Figure S16). The first type included extensively demethylated loci (Fig. 5b, bottom) and the second type contained epipolymorphic loci emerging due to partial methylation loss (Fig. 5b, top). Genomic regions that have lost their methylated state, termed partially methylated domains (PMDs), were initially discovered in a fibroblast cell line [22]. Several studies have reported cancer and noncancer human primary cells with PMDs [23, 24, 25, 26, 27]. PMDs cover approximately 50 to 75% of the genome of the human primary cell types and tissues investigated, while roughly a quarter are shared which indicates that PMDs retain strong tissue and cell type specificity characteristics [26]. To perform a comprehensive analysis of PMDs in MM, we used MethylSeekR [28] with default parameters to analyze WGBS data from our cohort of MM samples and the available WGBS dataset. We detected PMDs in both NPC samples and in only five MM samples from our cohort (Additional file 3: Figure S17). The PMD structure was highly similar between normal and MM cells despite a very variable level of DNA methylation (Additional file 3: Figure S18a). The base overlap was greater than 80%, the median length distribution was approximately 51 kb and the mean genome coverage was approximately 65% of the genome (Additional file 3: Figure S18b). Given the very good overlap between NPC and MM PMDs, we subsequently studied PMDs determined from NPCs (referred to as PC-PMDs) in MM patients.

PC-PMD borders are not random and coincide with spatial genome organization, as indicated by the closeness of topological associated domain (TAD) borders to PC-PMD borders, which is larger than expected by chance (p < 0.001). In addition, PC-PMDs share key features with PMDs of various tumor types and normal tissues [24, 26, 27], including the correlation with lamina-associated domains (LADs) [29], late-replication timing and low gene density (Additional file 3: Figure S19a,b,c). Except for one sample, PC-PMDs that intersected with LADs displayed a long replication time and significantly low methylation levels (Additional file 3: Figure S19d,e). Importantly, the vast majority of hypomethylated eloci (70%) were located within PC-PMDs.

Severe DNA methylation loss in PC-PMDs contributes to the redistribution of repressive histone marks and perturbations in CGIs/TSSs

We next examined the PC-PMDs methylation status across patients and the associated functional features. As expected, the DNA methylation level was lower within than outside of PC-PMDs (Fig. 5c) and variable between patients (Fig. 5d). In addition, this variability was associated with the PC-PMD length and replication time (Additional file 3: Figure S20). To study the relationship between DNA methylation and other epigenetic features, we integrated available WGBS data together with ChIP-seq data for histone marks. We found that the DNA methylation changes within PC-PMDs were associated with perturbations in both H3K9me3 and H3K27me3. Notably, the H3K9me3 deposit was associated with severe DNA methylation loss in long PC-PMDs (> 1 Mb) (Fig. 6a), as exemplified for the patient MM15548, with an unmethylated 3 Mb PC-PMD that was highly enriched with H3K9me3 (Fig. 6b). To our knowledge, such widespread H3K9me3 deposits associated with DNA methylation loss have so far been observed only in PMDs of cancer cell lines [26, 30, 31]. H3K27me3 was also perturbed inside PMDs. Notably, H3K27me3 enrichment was correlated with DNA methylation erosion (Fig. 6c). One notable example is the *DOCK3*-containing PC-PMD, which is enriched in H3K27me3 and showed H3K27me3 deposits in patients MM15548 and MM22965, who also exhibited DNA methylation erosion in this PMD; howerver, in another patient (MM23977) with a methylation level in this PMD comparable to that in NPCs, the H3K27me3 mark was absent (Fig. 6d). Taken together, these result show that perturbations in repressive histone marks are variable and depend on genomic regions, patients and DNA methylation levels.

We next investigated the impact of PC-PMD methylation erosion on regulatory regions. We found that the normal near-bimodal methylation state observed outside of PC-PMDs, was completely abolished inside PC-PMDs (Fig. 6e). As a result, promoter CGIs lost their hypomethylated state and gained intermediate methylation levels, while adjacent regions became less methylated to reach an intermediate degree of methylation. Notably, the acquisition of DNA methylation inside PC-PMDs resulted in a significant increase in the intermediate methylation state (>20% to <80%) of CGIs at the expense of strictly methylated or unmethylated states in all MM patients examined (Fig. 6f), in agreement with results obtained in breast cancer [24]. We noted that this phenomenon was also observed, albeit to a lesser extent, in NPCs (Additional file 3: Figure S21a); however, the percentage of loci with a high level of discordant reads (> 50%) was higher in all MM compared to normal samples (Additional file 3: Figure S21b) indicating that although PC-PMD methylation is perturbated in NPCs, CGI methylation patterns are more homogeneous (i.e., either fully methylated or unmethylated) (Additional file 3: Figure S21c).

Interestingly, a large proportion of bivalent promoter CGIs with disrupted DNA methylation was embedded within PC-PMDs (39% totally included and 54% with at least 30% of shared bases, as exemplified in Additional file 3: Figure S22). Furthermore, PC-PMD methylation loss was correlated with an overall epiallele shift increase (Spearman p = 0.46) (Additional file 3: Figure S23). These results support the notion that PC-PMD methylation loss may locally fuel epiallele shifts.

Given the interpatient heterogeneity of PMD methylation, we also investigated variations in gene expression. We found more gene expression variability inside PC-PMDs than outside of PC-PMDs; this difference was also maintained at relapse (Fig. 7a). Interestingly, we found that genes inside PC-PMDs were abundant within 186 kb of PC-PMD boundaries (Fig. 7b). We therefore focused our analysis on genes located near boundaries: Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed specific enrichment in immune-related pathways including cytokine-cytokine receptor interaction, complement and coagulation cascades, autoimmune thyroid disease, natural killer cell-mediated cytotoxicity, antigen processing and presentation, regulation of autophagy, cell adhesion molecule and graft-versus-host disease (Fig. 7c). This enrichment pattern in PC-PMDs could provide an explanation for immune evasion mechanisms that occur during tumor progression [32, 33].

Taken together, these results show that MM displays a high epigenomic instability and great transcriptomic variability in PC-PMDs which might be beneficial for tumor cells.

Three-dimensional (3D) reorganization is favored in hypomethylated PMDs

Global DNA methylation loss impacts spatial genome organization in the nucleus [34]. We wondered whether severe DNA hypomethylation that occurs during MM development could reshape 3D chromatin architecture in MM cells. The chromatin fiber of eukaryotic genomes is folded at multiple levels, including large-scale genomic structures to form distinct chromatin compartments A and B, characterized by gene-dense transcriptionally active open chromatin and gene-sparse transcriptionally closed chromatin, respectively [35]. During cancer development the genome is reorganized, and as a consequence, genomic regions of compartment A switch to those of compartment B and vice versa. The proportion of compartment A/B switching varies according to tumor type [36, 37]. We investigated the relationships between PC-PMDs and compartment A/B switching. We determined compartment A/B boundaries at 20-kb resolution from Hi-C data in a lymphoblastoid B cell line (GM12878), with a DNA methylome similar to that of PCs [12], and in the myeloma cell lines U266 and RPMI8226. We found that PC-PMDs were more prone to switch than other genomic regions (Fig. 8a). As a representative example, *IGF1R*, which encodes a major mediator of growth and survival in MM, and is located astride compartments A and B in NPCs, switched entirely to compartment A in both myeloma cell lines (Fig. 8b).

Altogether, these results show that compartment switching might be favored by PMD instability.

Discussion

MM is a heterogeneous disease. However, until now, only studies on genetic abnormalities have demonstrated the presence of subclonal populations that encountered complex evolution during the course of a disease. A recent study highlighted the contribution of intratumor epigenetic heterogeneity to shape CLL evolution after therapy [38]. To assess intrapatient DNA methylation heterogeneity in MM, we used the eRRBS technique, which allows us to capture the methylation status of

individual CpGs in a single cell (more precisely in one single allele), with a large sequencing depth. Combining the notions of epipolymorphism and epiallele switching, we showed that the MM epigenome is characterized by high intratumor heterogeneity with a predominantly stochastic methylation pattern. Furthermore, a high methylation entropy change between malignant and NPCs is associated with a high risk of relapse.

We found that genes that undergo epiallele shifting are enriched in the control of developmental genes. Interestingly, in a model wherein hematopoietic progenitors are proposed to be the cells of origin in MM, an aberrant epigenetic program persisting through normal cell differentiation is implicated in tumor initiation [39]. Further analysis of this aberrant epigenetic program revealed strong enrichment of biological functions associated with developmental regulation (Additional file 3: Figure S24), suggesting that a disruption in developmental pathways could play a key role in the initiation of MM and increase susceptibility to oncogene transformation in response to environmental changes [40].

Entropy changes in bivalent promoters are associated with DNA methylation gains. These gains are stochastic, with an elevated PDR, which leads to a decoupling relationship between promoter methylation and transcription. These results are in line with those obtained in CLL [15]. Moreover, in CLL, single cell RNA-seq analysis showed that a high PDR is correlated with a "noisy" transcriptional landscape and an intermediate transcriptional state that interferes with complete silencing or high-level expression. Several studies have also shown the role of cellular heterogeneity and gene expression noise in overcoming drug resistance or metastatic barriers [41, 42, 43]. Together, these data suggest that increased stochastic methylation variation allows tumor cells to better adapt and find new trajectories in response to environmental changes.

A large majority of entropy changes between normal and malignant plasma cells associated with DNA methylation loss are found in regions that are already partially methylated in normal cells, indicating that PC-PMDs delimit genomic regions in which methylation information is not accurately maintained due to reduced energy consumption and channel capacity, as recently demonstrated for compartment B [44]. These large regions that cover approximately 70% of the genome are the major source of variability within MM patients. The PMDs of naive B cells display strong, progressive DNA methylation loss with differentiation into PCs [25]. During their maturation, PCs undergo successive rounds of division coupled with DNA methylation loss, thus supporting the idea that PMD hypomethylation is linked to mitotic dynamics [45, 46, 47].

In myeloma cells, PC-PMD hypomethylation is associated with other key epigenetic aberrations, such as promoter CGI hypermethylation. Indeed, we observed a loss of the hypomethylated state and gained intermediate methylation levels in regulatory regions within PC-PMDs. This disturbance has already been observed in NPCs; however, disordered methylation at promoters is lower in NPCs than in malignant cells, indicating a homogeneous methylation pattern at the promoter being either fully methylated or unmethylated. Approximately half of bivalent promoters with a high PDR in MM are associated with PC-PMDs. Therefore, the DNA methylation landscape of MM resembles that of the placenta, with stochastic methylated

gain in CGIs embedded in large hypomethylated regions, suggesting that, as in other cancers, myeloma cells coopt placental nuclear programming [48, 49]; this finding is of particular interest since placental and cancerous tissues share relevant features such as immune modulation, angiogenesis and tissue invasion [50]. Interestingly, the specific enrichment of genes in immune-related pathways was revealed when we focused our analysis on genes located near PC-PMD boundaries.

PC-PMDs are also linked to the redistribution of repressive marks. We showed that PC-PMD hypomethylation is associated with a disruption in the H3K9me3 and H3K27me3 marks and variable depending on genomic regions, patients and DNA methylation levels. Promoter CGI controlling developmental genes are silenced by the PRC2 complex, which is responsible for the deposition of H3K27me3 via its catalytic component, EZH2. According to the model of Reddington et al [51], we can hypothesize that DNA methylation prevents PRC2 from binding to inappropriate targets, and that global hypomethylation due to tumorigenesis drives H3K27me3 redistribution responsible for the derepression of target genes and the repression of new genes. This redistribution may partially explain the derepression of *HOXA9* [52] and the de novo bivalent promoters [53] observed in MM. H3K27me3 redistribution could also play an important role in chromatin decompaction, as has been shown in ESCs [54], and could explain the increase in open chromatin regions within the heterochromatin state of MM samples [55]. These data highlight that PMDs must be considered as a separate entity in genomic analyses.

Furthermore, we showed that DNA hypomethylation is variable within the myeloma cells of one individual and that methylation loss modifies H3K27me3 distribution across patients; we can hypothesize that H3K27me3 redistribution is heterogeneous among myeloma cells, leading to cell-to-cell epigenetic variability. Consequently, the tumoral mass in a MM patient would be composed of an admixture of myeloma cells with divergent epigenetic identities, in accordance with what has been demonstrated recently in CLL [56].

Conclusion

Altogether, our results show the importance of genome-wide epigenetic analysis and reveal marked PMD instability in MM patients at presentation. The perturbation in PMDs occurs at the epigenetic, transcriptomic and 3D organization levels and is responsible for interpatient variability. This disturbance in PMDs could explain, in part, some aberrant and heterogeneous phenotypes across MM samples [57, 58]. In addition, the lack of accurate global DNA methylation maintenance also drives intrapatient DNA methylation heterogeneity, which can contribute to intrapatient variability, allowing cell-to-cell diversity in transcriptional programs and opening multiple trajectories in response to therapy (Additional file 3: Figure S25).

Materials and Methods

Patients and samples

Myeloma PCs were derived from bone marrow samples collected at different stages of myeloma: 47 samples form 26 patients were collected at the SMM, diagnosis and relapse stages. Patients (15 women and 11 men, with a median age of 59 years) were monitored at the Intergroupe Francophone du Myélome (IFM) center, and all provided informed consent (Additional file 1: Table S1).

Sample preparation and nucleic acid purification

Myeloma cells were purified using nanobeads and an anti-CD138 antibody (RoboSep, Stem Cell Technologies). After immunomagnetic sorting, the purity of the plasma cell suspension was verified, and only samples with at least 85% of PCs were subjected to genomic analysis. The average cell purity of MM was > 99% (range 90-100%). Control normal plasma cells (NPC1 to NPC3) were obtained from the bone marrow samples of three patients suspected of having monoclonal gammopathy of undetermined significance (MGUS), and the absence of abnormal PCs was assessed by flow cytometry. The average CD138+ cell purity was > 75%. The DNA and RNA of CD138+ cells were purified using the Qiagen protocol and the quality and quantity of the nucleic acids were measured on the Nanodrop, Qubit and Agilent profiles and stored in the MM biobanks of the Nantes and Toulouse Hospitals.

eRRBS

eRRBS is an improvement of the reduced representation bisulfite sequencing (RRBS) protocol, resulting in an increase in CpG detection and coverage. eRRBS library preparations were performed by Integragen and adapted from the protocol described by Garrett-Bakelmanet al. [59]. DNA was digested with the Msp1 enzyme, fragments between 150 bp and 400 bp were selected, and bisulfite conversions were processed. Libraries were sequenced on a HiSeqTM 4000 Illumina machine using 75 bp paired-end reads to an average depth of 50X per covered CpG. The average number of reads sequenced per patient was 55 216 285; and the average alignment rate of uniquely mapped reads was 60.81%, with an average of 2 712 252 CpGs per patients with a coverage of 10X and an average of 1 538 510 CpGs per patient with a coverage of 60X (Additional file 4: Table S3).

eRRBS analysis

The adaptor sequences were removed by Cutadapt (version 1.10) [60]. FastQC (version 0.11.4) was used for quality control of the Illumina paired-end sequencing data [61]. Bisulfite reads were aligned to the bisulfite-converted hg19 genome with the nondirectional model of Bismark alignment software (version 0.14.1) [62]. CpG methylation levels were obtained using the R package methylKit with the default settings: a minimum of 10 reads covering a CpG and at least 20 PHRED quality scores by CpG [63]. We calculated the epigenetic changes between two stages using methclone [17], an algorithm that detects loci of 4 adjacent CpGs (minimum depth of 60 reads), called epialleles (16 possible patterns according to CpG methylation).

Epipolymorphism and epiallele shifting

Epipolymorphism (e) was calculated for each locus of the four adjacent CpGs, to measure intratumor epigenetic heterogeneity, as previously described by Landan et al. [18]:

$$e = 1 - \sum_{i=1}^{16} p_i^2$$

where p_i is the proportion of the epiallele i. The minimal epipolymorphism is 0 (only one pattern represented); epipolymorphism cannot have a value greater than 1. The maximum value of epipolymorphism is 0.9375 (when all 16 patterns are equally distributed). To quantify the degree of epiallele pattern shift, methclone was

used to compute the entropy difference (ΔS) and thus compare the distributions of epialleles between different stages. The entropy difference value can range from 0 (no change) to -144 (maximum change). Loci are characterized as eloci when ΔS < -70 (corresponding to a significant entropy shift). The methclone algorithm allows the discovery, quantification and ranking of subclonal selection based on epiallele shifts. This allows us to measure clonal evolution between disease stages and epigenetic heterogeneity.

To normalize and compare the number of eloci per patient, we computed the number of eloci per million loci sequenced (EPM) as previously described by Sheng Li et al. [17]:

$$EPM = \frac{10^6}{C} \times E$$

where E is the total number of eloci detected between the two stages and C is the total number of loci covered by both samples.

A locus between the NPC and diagnosis samples was defined as an elocus only if it was annotated as an elocus with the 3 NPCs. A hypermethylated or hypomethylated elocus was defined as a DNA methylation difference of at least 25%.

The PDR variable was defined as the percentage of epialleles exhibiting heterogeneous DNA methylation (i.e., not fully methylated or unmethylated).

WGBS

Data quality control and adaptor trimmed were performed with the Trimmomatic tool [64]. Read mapping was carried out with the methylCtools aligner [31]. methylCtools uses an alignment approach similar to that of Bismark by improving the handling of large amounts of data and the speed of alignment. CpG extraction and methylation analyses were carried out with methylCtools and the bsseq R package [65], allowing the analysis, management and storage of WGBS data. As with eRRBS data, we set a coverage threshold of a minimum of 10X for each CpG sequenced for our analyses. The dmrseq R package was used to visualize WGBS data and smooth methylation signals [66].

Detecting PMDs

PMDs were detected in WGBS samples using the R package MethylSeekR [28]. Prior to the detection of PMDs, CpGs overlapping with common single-nucleotide polymorphism (SNPs) were removed from the data (dbSNP137commonhg19, version 1.0.0). The distribution of α values was used to determine the presence or absence of PMDs in samples. When the α distribution was bimodal, the segmentPMDs function was run to identify PMDs on the genome by a hidden Markov model.

RNA sequencing (RNA-seq)

RNA-seq was performed using 200 ng of total RNA by GATC Biotech. Directional libraries were generated after mRNA selection by polyA selection using the UTP method. RNA-seq libraries were sequenced on a HiSeq 2500 Illumina machine using 100 bp paired-end reads. The average number of reads sequenced per sample was 68 454 433. The average alignment rate of uniquely mapped reads was 66% (corresponding to 48 034 921 reads). Read alignment was performed using the STAR aligner (version 2.4.0f1) [67] and human genome hg19 as the reference. FastQC

(version 0.11.4) [61] was used for basic quality control of the Illumina paired-end sequencing data. PCR duplicates were determined and removed using the Picard algorithm [68]. The number of reads mapped to each gene was calculated using HTSeq-count, part of the HTSeq framework [69], version 0.6.0. We then normalized the mapped read counts per million of mapped fragments (FPM) using the robust median ratio method with the DESeq2 R package [70].

Hi-C data analysis and compartments A/B

Hi-C datasets were downloaded for three cell lines: GM12878 (GSM1608505), RPMI-8226 (GSM2334832) and U266 (GSM2334834). HiC-Pro software (version 2.10.0) [71] was used to process Hi-C data, from raw-data to normalized contact maps. All reads were mapped to hg19 using Bowtie2 (global parameters: –very-sensitive -L 30 –score-min L,-0.6,-0.2 –end-to-end –reorder; local parameters:–very-sensitive -L 20 –score-min L,-0.6,-0.2 –end-to-end –reorder). Contact maps were generated at 20 kb resolution and normalized by the iterative correction and eigen-vector decomposition (ICED) technique.

The 20-kb resolution intrachromosomal contact matrices generated by HiC-Pro were used as input to determine compartments A/B and to annotate and visualize interaction maps with R package HiTC version 1.22.1 [72] . Principal component analysis was used to separate chromatin into two compartments: compartment A, with higher gene density, and compartment B, with lower gene density. The determination of compartments A and B was estimated by the analysis of the eigenvectors of the genome contact matrix by the observed-expected method. On the basis that changes in the sign of the eigenvector of the contact matrix correspond to the limits of the genome compartments and taking into account gene density, the compartmentalization of the genome was defined.

TADs

Hi-C data were used to determine the TADs with the HiCExplorer tool, a set of programs used to process, normalize, analyze and visualize Hi-C data [73]. TADs were defined with hicFindTADs and hicPlotTADs was used to visualize the TADs. For the comparison of PMD and TAD borders, we generated a set of randomized PMDs with a size and genomic distributions similar to those of real PMDs. We then calculated the distance between TAD borders and, on the one hand, PMD borders, and, on the other hand, random PMD borders. A t test was carried out to compare the distributions of these distances.

Genomic annotation

RefSeq annotation and CpG islands were obtained from UCSC (`https:/garance/genome.ucsc.edu/`) using the February 2009 (GRCh37/hg19) assembly. CpG shores were defined as regions flanking 2 kb of CpG islands, and CpG shelves were defined as regions flanking 2 kb of CpGs shores.

HOMER was used to annotate loci and eloci, using the annotatePeaks.pl script [74], which determines the genomic type annotation occupied by the center of the loci. We used their basic annotation, including TSSs, transcription termination sites (TTSs), exons, introns and intergenic regions. We defined the promoters as regions

flanking 2 kb of the TSS, and we defined promoter CGIs as promoters that intersect with CpG islands.

All 127 reference epigenomes with 25 chromatin state segmentation annotations were downloaded from the NIH Roadmap Epigenomics Project (`https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/imputed12marks/jointModel/final/;all\OT1\textquotedblleft25_imputed12marks_mnemonics.bed.gz\OT1\textquotedblrightfiles`).

To measure the intersections between regions, bedtools was used (more precisely, the "intersect" option) [75].

Data from the BLUEPRINT Consortium were converted to hg19 coordinates using the liftOver tool [76].

Functional annotation

GREAT tools (version 3.0.0) was used to assign biological functions to noncoding genomic regions by analyzing the annotations of the nearby genes [77] under the following default gene regulatory domain definitions; basal promoter: 5 kb upstream, 1 kb downstream and extension up to 1 Mb. Enrichment statistics were computed using the binomial test and the hypergeometric gene-based test. Pathways were selected as significantly enriched if the false discovery rate (FDR q-value) was < 0.01.

The Database for Annotation, Visualization and Integrated Discovery (DAVID) web interface (version 6.7) was used to perform functional enrichment analysis from a list of genes, especially the KEGG pathways functional database [78]. Enrichment statistics were computed using the Fisher test. For the significance threshold, we have considered genes as greatly enriched if the annotation categories yielded a p-value less than 0.1 (DAVID default threshold)[79].

Survival analysis

Time to relapse data were available for 17 patients. For the relapse-free survival (RFS) analysis, the survival endpoint in this study was the time from diagnosis until relapse. The patients were divided into two groups by the median EPM value: low and high EPMs. Survival curves were estimated using the Kaplan-Meier method and compared with the log-rank test.

For all statistical tests in this study, a two-sided p-value of 0.05 was considered statistically significant. All statistical analyses were performed with software R 3.5.1 [80], in addition to the packages already mentioned in the text, packages data.table [81], viridis, ggbio, GenomicRanges, RColorBrewer, biovizBase, grid, gridBase, MASS, fields, KernSmooth, sp, org.Hs.eg.db, DBI and survival were used.

Abbreviation

MM: Multiple myeloma; PMDs : Partially methylated domains; CGI: CpG island; TSS: Transcription start site; PCs: Plasma cells; NPCs: Normal plasma cells; WGBS: Whole genome bisulfite sequencing; eRRBS: Enhanced reduced representation bisulfite sequencing; SMM: Smoldering multiple myeloma; PDR: Proportion of discordant reads; eloci: Epigenetic loci; EPM: Eloci per million loci covered; ESCs: embryonic stem cells; FPM: mapped read counts per million of mapped fragments; OR: Odd ratio; CLL: Chronic lymphocytic leukemia; TAD: Topological associated domain; LAD: lamina-associated domain; KEGG: Kyoto encyclopedia of genes and genomes; 3D: Three-dimensional; IFM: Intergroupe francophone du myélome; MGUS: Gammopathy of undetermined significance; RRBS: Reduced representation bisulfite sequencing; SNPs: Single-nucleotide polymorphism; RNA-seq: RNA sequencing; ICED: Iterative correction and eigenvector decomposition; TTS: Transcription termination site; FDR: False discovery rate; DAVID: Database for annotation, visualization and integrated discovery; RFS: Relapse-free survival.

Author's contributions
S.M. and F.M. designed the study, collected, and analyzed the data and wrote the paper; J.D. and C.H. collected and analyzed the data and wrote the paper; M.D, E.D, N.R, H.A.-L., O.D., T.F., J-P. M., R.E., M.P. and N.M. collected the data; C.G.-C., V.G. and L.C. analyzed the data.

Availability of data
WGBS, eRRBS and RNA-seq data have been submitted to the European Genome-phenome Archive (https://www.ebi.ac.uk/ega) under dataset accession EGAS00001004346, EGAS00001004348, EGAS00001004347, respectively.
WGBS Blueprint data of normal plasma cells and multiple myeloma patient are available from the EGA under the accessions EGAD00001002322 and EGAD00001002521 respectively. ChIP-Seq data of multiple myeloma patient are available from the EGA under the accessions EGAD00001002379.

Ethics approval and consent to participate
All the samples were obtained from patients enrolled in the different randomized clinical trials implemented by the Intergroupe Francophone du Myelome (IFM); IFM 99-02, IFM 99-06, IFM 2005-01, FIRST/IFM 2007-01, IFM 2007-02, Genomgus Study and IFM 2009 (www.clinicaltrials.gov) or routine practice in France. Samples and data were obtained after an informed consent was signed. All research procedures conformed to the principles of Helsinki Declaration.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Université de Nantes, CNRS, INSERM, CRCINA, F-44000 Nantes, France. [2]Institut de Cancérologie de l'Ouest, Nantes-Saint Herblain, France. [3]LS2N, CNRS, Université de Nantes, Nantes, France. [4] Centre Hospitalier Universitaire, Nantes, France. [5]Institut Universitaire du Cancer, CHU, Centre de Recherche en Cancérologie de Toulouse, INSERM 1037, Toulouse, France. [6]Centre Hospitalier Universitaire, Rennes, France. [7]Centre Hospitalier Universitaire, Lilles, France. [8]Research Group Genome Organization & Function, DKFZ, and BioQuant Heidelberg, 69120 Heidelberg, Germany. [9]Health Data Science Unit, Medical Faculty Heidelberg and BioQuant, 69120 Heidelberg, Germany. [10]Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Charitéplatz 1, 10117 Berlin, Germany Berlin Institute of Health (BIH), Center for Digital Health, Anna-Louisa-Karsch-Strasse 2, 10178 Berlin, Germany. [11]Dana-Farber Cancer Institute, Harvard Medical School, LeBow Institute for Myeloma Therapeutics and Jerome Lipper Center for Multiple Myeloma Research, Boston, MA, United States.

References

1. Leung-Hagesteijn, C., Erdmann, N., Cheung, G., Keats, J.J., Stewart, A.K., Reece, D.E., Chung, K.C., Tiedemann, R.E.: Xbp1s-negative tumor b cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. Cancer cell **24**(3), 289–304 (2013)
2. Zhang, X.-D., Baladandayuthapani, V., Lin, H., Mulligan, G., Li, B., Esseltine, D.-L.W., Qi, L., Xu, J., Hunziker, W., Barlogie, B., *et al.*: Tight junction protein 1 modulates proteasome capacity and proteasome inhibitor sensitivity in multiple myeloma via egfr/jak1/stat3 signaling. Cancer cell **29**(5), 639–652 (2016)
3. Bolli, N., Avet-Loiseau, H., Wedge, D.C., Van Loo, P., Alexandrov, L.B., Martincorena, I., Dawson, K.J., Iorio, F., Nik-Zainal, S., Bignell, G.R., et al.: Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nature communications **5** (2014)
4. Weinhold, N., Ashby, C., Rasche, L., Chavan, S.S., Stein, C., Stephens, O.W., Tytarenko, R., Bauer, M.A., Meissner, T., Deshpande, S., *et al.*: Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. Blood **128**(13), 1735–1744 (2016)
5. Magrangeas, F., Avet-Loiseau, H., Gouraud, W., Lodé, L., Decaux, O., Godmer, P., Garderet, L., Voillat, L., Facon, T., Stoppa, A.M., *et al.*: Minor clone provides a reservoir for relapse in multiple myeloma. Leukemia **27**(2), 473 (2013)
6. Keats, J.J., Chesi, M., Egan, J.B., Garbitt, V.M., Palmer, S.E., Braggio, E., Van Wier, S., Blackburn, P.R., Baker, A.S., Dispenzieri, A., *et al.*: Clonal competition with alternating dominance in multiple myeloma. Blood **120**(5), 1067–1076 (2012)
7. Egan, J.B., Shi, C.-X., Tembe, W., Christoforides, A., Kurdoglu, A., Sinari, S., Middha, S., Asmann, Y., Schmidt, J., Braggio, E., *et al.*: Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. Blood, The Journal of the American Society of Hematology **120**(5), 1060–1066 (2012)

8. Walker, B.A., Wardell, C.P., Melchor, L., Hulkki, S., Potter, N.E., Johnson, D.C., Fenwick, K., Kozarewa, I., Gonzalez, D., Lord, C.J., *et al.*: Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t (4; 14) and t (11; 14) myeloma. Blood, The Journal of the American Society of Hematology **120**(5), 1077–1086 (2012)

9. Maura, F., Bolli, N., Angelopoulos, N., Dawson, K.J., Leongamornlert, D., Martincorena, I., Mitchell, T.J., Fullam, A., Gonzalez, S., Szalat, R., *et al.*: Genomic landscape and chronological reconstruction of driver events in multiple myeloma. Nature communications **10**(1), 1–12 (2019)

10. Lawrence, M.S., Stojanov, P., Mermel, C.H., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., Getz, G.: Discovery and saturation analysis of cancer genes across 21 tumor types. Nature **505**(7484), 495 (2014)

11. Marusyk, A., Janiszewska, M., Polyak, K.: Intratumor heterogeneity: The rosetta stone of therapy resistance. Cancer cell **37**(4), 471–484 (2020)

12. Agirre, X., Castellano, G., Pascual, M., Heath, S., Kulis, M., Segura, V., Bergmann, A., Esteve, A., Merkel, A., Raineri, E., *et al.*: Whole-epigenome analysis in multiple myeloma reveals dna hypermethylation of b cell-specific enhancers. Genome research **25**(4), 478–487 (2015)

13. Berman, B.P., Weisenberger, D.J., Aman, J.F., Hinoue, T., Ramjan, Z., Liu, Y., Noushmehr, H., Lange, C.P., van Dijk, C.M., Tollenaar, R.A., *et al.*: Regions of focal dna hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina–associated domains. Nature genetics **44**(1), 40 (2012)

14. Hansen, K.D., Timp, W., Bravo, H.C., Sabunciyan, S., Langmead, B., McDonald, O.G., Wen, B., Wu, H., Liu, Y., Diep, D., *et al.*: Increased methylation variation in epigenetic domains across cancer types. Nature genetics **43**(8), 768–775 (2011)

15. Landau, D.A., Clement, K., Ziller, M.J., Boyle, P., Fan, J., Gu, H., Stevenson, K., Sougnez, C., Wang, L., Li, S., *et al.*: Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer cell **26**(6), 813–825 (2014)

16. Li, S., Garrett-Bakelman, F.E., Chung, S.S., Sanders, M.A., Hricik, T., Rapaport, F., Patel, J., Dillon, R., Vijay, P., Brown, A.L., *et al.*: Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. Nature medicine **22**(7), 792 (2016)

17. Li, S., Garrett-Bakelman, F., Perl, A.E., Luger, S.M., Zhang, C., To, B.L., Lewis, I.D., Brown, A.L., D'Andrea, R.J., Ross, M.E., *et al.*: Dynamic evolution of clonal epialleles revealed by methclone. Genome biology **15**(9), 472 (2014)

18. Landan, G., Cohen, N.M., Mukamel, Z., Bar, A., Molchadsky, A., Brosh, R., Horn-Saban, S., Zalcenstein, D.A., Goldfinger, N., Zundelevich, A., *et al.*: Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. Nature genetics **44**(11), 1207–1214 (2012)

19. Easwaran, H., Johnstone, S.E., Van Neste, L., Ohm, J., Mosbruger, T., Wang, Q., Aryee, M.J., Joyce, P., Ahuja, N., Weisenberger, D., *et al.*: A dna hypermethylation module for the stem/progenitor cell signature of cancer. Genome research **22**(5), 837–849 (2012)

20. Ernst, J., Kellis, M.: Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. Nature biotechnology **33**(4), 364 (2015)

21. Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., *et al.*: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature **448**(7153), 553–560 (2007)

22. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., *et al.*: Human dna methylomes at base resolution show widespread epigenomic differences. nature **462**(7271), 315–322 (2009)

23. Schroeder, D.I., Blair, J.D., Lott, P., Yu, H.O.K., Hong, D., Crary, F., Ashwood, P., Walker, C., Korf, I., Robinson, W.P., *et al.*: The human placenta methylome. Proceedings of the national academy of sciences **110**(15), 6037–6042 (2013)

24. Brinkman, A.B., Nik-Zainal, S., Simmer, F., Rodriguez-Gonzalez, F.G., Smid, M., Alexandrov, L.B., Butler, A., Martin, S., Davies, H., Glodzik, D., *et al.*: Partially methylated domains are hypervariable in breast cancer and fuel widespread cpg island hypermethylation. Nature communications **10**(1), 1–10 (2019)

25. Durek, P., Nordström, K., Gasparoni, G., Salhab, A., Kressler, C., De Almeida, M., Bassler, K., Ulas, T., Schmidt, F., Xiong, J., *et al.*: Epigenomic profiling of human cd4+ t cells supports a linear differentiation model and highlights molecular regulators of memory development. Immunity **45**(5), 1148–1161 (2016)

26. Salhab, A., Nordström, K., Gasparoni, G., Kattler, K., Ebert, P., Ramirez, F., Arrigoni, L., Müller, F., Polansky, J.K., Cadenas, C., *et al.*: A comprehensive analysis of 195 dna methylomes reveals shared and cell-specific features of partially methylated domains. Genome biology **19**(1), 150 (2018)

27. Zhou, W., Dinh, H.Q., Ramjan, Z., Weisenberger, D.J., Nicolet, C.M., Shen, H., Laird, P.W., Berman, B.P.: Dna methylation loss in late-replicating domains is linked to mitotic cell division. Nature genetics **50**(4), 591–602 (2018)

28. Burger, L., Gaidatzis, D., Schübeler, D., Stadler, M.B.: Identification of active regulatory regions from dna methylation data. Nucleic acids research **41**(16), 155–155 (2013)

29. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., *et al.*: Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature **453**(7197), 948–951 (2008)

30. Hon, G.C., Hawkins, R.D., Caballero, O.L., Lo, C., Lister, R., Pelizzola, M., Valsesia, A., Ye, Z., Kuan, S., Edsall, L.E., *et al.*: Global dna hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome research **22**(2), 246–258 (2012)

31. Hovestadt, V., Jones, D.T., Picelli, S., Wang, W., Kool, M., Northcott, P.A., Sultan, M., Stachurski, K., Ryzhova, M., Warnatz, H.-J., *et al.*: Decoding the regulatory landscape of medulloblastoma using dna methylation sequencing. Nature **510**(7506), 537–541 (2014)

32. Jung, H., Kim, H.S., Kim, J.Y., Sun, J.-M., Ahn, J.S., Ahn, M.-J., Park, K., Esteller, M., Lee, S.-H., Choi, J.K.: Dna methylation loss promotes immune evasion of tumours with high mutation and copy number load.

Nature communications **10**(1), 1–12 (2019)

33. Ryu, D., Kim, S.J., Hong, Y., Jo, A., Kim, N., Kim, H.-J., Lee, H.-O., Kim, K., Park, W.-Y.: Alterations in the transcriptional programs of myeloma cells and the microenvironment during extramedullary progression affect proliferation and immune evasion. Clinical Cancer Research **26**(4), 935–944 (2020)

34. Madakashira, B.P., Sadler, K.C.: Dna methylation, nuclear organization, and cancer. Frontiers in genetics **8**, 76 (2017)

35. Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.*: Comprehensive mapping of long-range interactions reveals folding principles of the human genome. science **326**(5950), 289–293 (2009)

36. Barutcu, A.R., Lajoie, B.R., McCord, R.P., Tye, C.E., Hong, D., Messier, T.L., Browne, G., van Wijnen, A.J., Lian, J.B., Stein, J.L., *et al.*: Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. Genome biology **16**(1), 214 (2015)

37. Wu, P., Li, T., Li, R., Jia, L., Zhu, P., Liu, Y., Chen, Q., Tang, D., Yu, Y., Li, C.: 3d genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. Nature communications **8**(1), 1–11 (2017)

38. Gaiti, F., Chaligne, R., Gu, H., Brand, R.M., Kothen-Hill, S., Schulman, R.C., Grigorev, K., Risso, D., Kim, K.-T., Pastore, A., *et al.*: Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. Nature **569**(7757), 576–580 (2019)

39. Vicente-Dueñas, C., Romero-Camarero, I., González-Herrero, I., Alonso-Escudero, E., Abollo-Jiménez, F., Jiang, X., Gutierrez, N.C., Orfao, A., Marín, N., Villar, L.M., *et al.*: A novel molecular mechanism involved in multiple myeloma development revealed by targeting mafb to haematopoietic progenitors. The EMBO journal **31**(18), 3704–3717 (2012)

40. Vaz, M., Hwang, S.Y., Kagiampakis, I., Phallen, J., Patil, A., O'Hagan, H.M., Murphy, L., Zahnow, C.A., Gabrielson, E., Velculescu, V.E., *et al.*: Chronic cigarette smoke-induced epigenomic changes precede sensitization of bronchial epithelial cells to single-step transformation by kras mutations. Cancer cell **32**(3), 360–376 (2017)

41. Nguyen, A., Yoshida, M., Goodarzi, H., Tavazoie, S.F.: Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. Nature communications **7**(1), 1–13 (2016)

42. Shaffer, S.M., Dunagin, M.C., Torborg, S.R., Torre, E.A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P.A., Xiao, M., *et al.*: Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. Nature **546**(7658), 431–435 (2017)

43. Farquhar, K.S., Charlebois, D.A., Szenk, M., Cohen, J., Nevozhay, D., Balázsi, G.: Role of network-mediated stochasticity in mammalian drug resistance. Nature communications **10**(1), 1–14 (2019)

44. Jenkinson, G., Pujadas, E., Goutsias, J., Feinberg, A.P.: Potential energy landscapes identify the information-theoretic nature of the epigenome. Nature Genetics **49**(5), 719–729 (2017)

45. Barwick, B.G., Scharer, C.D., Bally, A.P., Boss, J.M.: Plasma cell differentiation is coupled to division-dependent dna hypomethylation and gene regulation. Nature immunology **17**(10), 1216 (2016)

46. Lai, A.Y., Mav, D., Shah, R., Grimm, S.A., Phadke, D., Hatzi, K., Melnick, A., Geigerman, C., Sobol, S.E., Jaye, D.L., *et al.*: Dna methylation profiling in human b cells reveals immune regulatory elements and epigenetic plasticity at alu elements during b-cell activation. Genome research **23**(12), 2030–2041 (2013)

47. Caron, G., Hussein, M., Kulis, M., Delaloy, C., Chatonnet, F., Pignarre, A., Avner, S., Lemarié, M., Mahé, E.A., Verdaguer-Dot, N., *et al.*: Cell-cycle-dependent reconfiguration of the dna methylome during terminal differentiation of human b cells into plasma cells. Cell reports **13**(5), 1059–1071 (2015)

48. Smith, Z.D., Shi, J., Gu, H., Donaghey, J., Clement, K., Cacchiarelli, D., Gnirke, A., Michor, F., Meissner, A.: Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. Nature **549**(7673), 543–547 (2017)

49. Lorincz, M.C., Schübeler, D.: Evidence for converging dna methylation pathways in placenta and cancer. Developmental cell **43**(3), 257–258 (2017)

50. Novakovic, B., Saffery, R.: Placental pseudo-malignancy from a dna methylation perspective: unanswered questions and future directions. Frontiers in genetics **4**, 285 (2013)

51. Reddington, J.P., Perricone, S.M., Nestor, C.E., Reichmann, J., Youngson, N.A., Suzuki, M., Reinhardt, D., Dunican, D.S., Prendergast, J.G., Mjoseng, H., *et al.*: Redistribution of h3k27me3 upon dna hypomethylation results in de-repression of polycomb target genes. Genome biology **14**(3), 25 (2013)

52. Chapman, M.A., Lawrence, M.S., Keats, J.J., Cibulskis, K., Sougnez, C., Schinzel, A.C., Harview, C.L., Brunet, J.-P., Ahmann, G.J., Adli, M., *et al.*: Initial genome sequencing and analysis of multiple myeloma. Nature **471**(7339), 467–472 (2011)

53. Agarwal, P., Alzrigat, M., Párraga, A.A., Enroth, S., Singh, U., Ungerstedt, J., Österborg, A., Brown, P.J., Ma, A., Jin, J., *et al.*: Genome-wide profiling of histone h3 lysine 27 and lysine 4 trimethylation in multiple myeloma reveals the importance of polycomb gene targeting and highlights ezh2 as a potential therapeutic target. Oncotarget **7**(6), 6809 (2016)

54. McLaughlin, K.A., Flyamer, I.M., Thomson, J.P., Mjoseng, H.K., Shukla, R., Williamson, I., Grimes, G.R., Illingworth, R.S., Adams, I.R., Pennings, S., et al.: Dna methylation directs polycomb-dependent 3d genome re-organisation in naive pluripotency. bioRxiv, 527309 (2019)

55. Jin, Y., Chen, K., De Paepe, A., Hellqvist, E., Krstic, A.D., Metang, L., Gustafsson, C., Davis, R.E., Levy, Y.M., Surapaneni, R., *et al.*: Active enhancer and chromatin accessibility landscapes chart the regulatory network of primary multiple myeloma. Blood, The Journal of the American Society of Hematology **131**(19), 2138–2150 (2018)

56. Pastore, A., Gaiti, F., Lu, S.X., Brand, R.M., Kulm, S., Chaligne, R., Gu, H., Huang, K.Y., Stamenova, E.K., Béguelin, W., *et al.*: Corrupted coordination of epigenetic modifications leads to diverging chromatin states and transcriptional heterogeneity in cll. Nature communications **10**(1), 1–11 (2019)

57. Bataille, R., Jégo, G., Robillard, N., Barillé-Nion, S., Harousseau, J.-L., Moreau, P., Amiot, M., Pellat-Deceunynck, C.: The phenotype of normal, reactive and malignant plasma cells. identification of' many

and multiple myelomas" and of new targets for myeloma therapy. Haematologica **91**(9), 1234–1240 (2006)

58. Capp, J.-P., Bataille, R.: Multiple myeloma exemplifies a model of cancer based on tissue disruption as the initiator event. Frontiers in oncology **8**, 355 (2018)

59. Garrett-Bakelman, F.E., Sheridan, C.K., Kacmarczyk, T.J., Ishii, J., Betel, D., Alonso, A., Mason, C.E., Figueroa, M.E., Melnick, A.M.: Enhanced reduced representation bisulfite sequencing for assessment of dna methylation at base pair resolution. Journal of visualized experiments: JoVE (96) (2015)

60. Martin, M.: Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal **17**(1), 10 (2011)

61. Andrews, S., et al.: FastQC: a quality control tool for high throughput sequence data (2010)

62. Krueger, F., Andrews, S.R.: Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. bioinformatics **27**(11), 1571–1572 (2011)

63. Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., Mason, C.E.: methylkit: a comprehensive r package for the analysis of genome-wide dna methylation profiles. Genome biology **13**(10), 87 (2012)

64. Bolger, A.M., Lohse, M., Usadel, B.: Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics **30**(15), 2114–2120 (2014)

65. Hansen, K.D., Langmead, B., Irizarry, R.A.: Bsmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. Genome biology **13**(10), 83 (2012)

66. Korthauer, K., Chakraborty, S., Benjamini, Y., Irizarry, R.A.: Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. Biostatistics **20**(3), 367–383 (2019)

67. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R.: Star: ultrafast universal rna-seq aligner. Bioinformatics **29**(1), 15–21 (2013)

68. Picard toolkit. Broad Institute (2018)

69. Anders, S., Pyl, P.T., Huber, W.: Htseq—a python framework to work with high-throughput sequencing data. Bioinformatics **31**(2), 166–169 (2015)

70. Love, M.I., Huber, W., Anders, S.: Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome biology **15**(12), 550 (2014)

71. Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., Barillot, E.: Hic-pro: an optimized and flexible pipeline for hi-c data processing. Genome biology **16**(1), 259 (2015)

72. Servant, N., Lajoie, B.R., Nora, E.P., Giorgetti, L., Chen, C.-J., Heard, E., Dekker, J., Barillot, E.: Hitc: exploration of high-throughput 'c'experiments. Bioinformatics **28**(21), 2843–2844 (2012)

73. Ramírez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Grüning, B.A., Villaveces, J., Habermann, B., Akhtar, A., Manke, T.: High-resolution tads reveal dna sequences underlying genome organization in flies. Nature communications **9**(1), 1–15 (2018)

74. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., Glass, C.K.: Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. Molecular cell **38**(4), 576–589 (2010)

75. Quinlan, A.R., Hall, I.M.: Bedtools: a flexible suite of utilities for comparing genomic features. Bioinformatics **26**(6), 841–842 (2010)

76. Rosenbloom, K.R., Armstrong, J., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al.: The ucsc genome browser database: 2015 update. Nucleic acids research **43**(D1), 670–681 (2015)

77. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G.: Great improves functional interpretation of cis-regulatory regions. Nature biotechnology **28**(5), 495–501 (2010)

78. Huang, D.W., Sherman, B.T., Lempicki, R.A.: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research **37**(1), 1–13 (2009)

79. Sherman, B.T., Lempicki, R.A., et al.: Systematic and integrative analysis of large gene lists using david bioinformatics resources. Nature protocols **4**(1), 44 (2009)

80. Team, R.C., et al.: R: A language and environment for statistical computing (2013)

81. Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Stetsenko, P., Short, T., Lianoglou, S., Antonyan, E., Bonsch, M., Parsonage, H., et al.: Package 'data. table'. Extension of 'data. frame (2015)

Figure 1 Global analysis of the DNA methylome of myeloma samples at diagnosis and NPC samples. (a) Circos plot of DNA methylation levels measured in WGBS data. From the center to the outside: NPC5 (data from the BLUEPRINT project, ERX301127), M#17, M#19, and M#10. Histograms represent CpG methylation levels averaged in 10-Mbp genomic windows. (b) Density distribution of DNA methylation levels measured in WGBS data for NPC (green, data from the BLUEPRINT project, ERX301127,ERX715130) and MM diagnosis (orange) samples. This color code was used in all figures. (c) CpG methylation within 100 bp genomic units as a function of local CpG density. (d) Mean methylation of CGIs and flanking regions in MM patients compared to NPCs. (e) Dispersive DNA methylation surrounding the *DOC2B* promoter in MM samples and in NPCs.

Figure 2 Intrapatient DNA methylation heterogeneity in MM. (a) Epipolymorphism levels as a function of the average DNA methylation at each locus in NPC (green) and diagnosis (orange) samples. The maximal epipolymorphism (continuous black line) and the bimodal epipolymorphism (dotted black line) are represented. (b) Scatterplots showing loci organization for NPC2 and M#19 diagnosis samples. Each point corresponds to a locus of 4 adjacent CpGs. On the left, each point is color coded according to the density of the surrounding points; on the right, each point is color coded according to its epipolymorphism level. (c) Mean proportion of discordant reads per sample.

Figure 3 Evolution of epiallelic changes during the progression of MM and their clinical impact. (a) Time to relapse analysis for patients with high (blue, n=9) or low (gray, n=8) EPM values. (b) EPM values of NPC samples (NPC eloci, in green) and MM samples at diagnosis (MM eloci in orange). (c) Correlation between EPM at diagnosis and relapse compared to that in NPCs. (d) Enrichment analysis of eloci (NPC vs diagnosis) located in promoter CGIs and their associated ontological terms was performed with the GREAT tool [77]. (e) EPM value for each pairwise comparison between diagnosis and relapse (diagnosis vs relapse eloci). (f) Proportion of eloci presenting a selection of a methylation pattern at diagnosis compared to NPCs (in gray) and at relapse compared to diagnosis (in blue).

Figure 4 Methylation disruption in bivalent promoters in MM. (a) Scatterplot of eloci in bivalent promoters as a function of DNA methylation in NPCs and diagnosis (left) and as a function of DNA methylation at diagnosis and relapse (right) for patient M#17. The color gradient corresponds to the point density (low is green; high is red). (b) Scatterplot of DNA methylation of the eloci promoters versus RNA expression of the associated genes. (c) Normalized expression values of genes with a promoter CGI containing a hypermethylated elocus or no elocus (left) and the variation coefficient of these genes in MM samples (right). (d) Ontological analysis of genes with a bivalent promoter CGI affected by hypermethylated eloci. (e) Average methylation epiallele patterns of eloci in promoter CGIs for NPC and diagnosis samples (epi= epipolymorphism). (f) Odds ratios with 95% confidence intervals for the association between gene expression (FPM > 1) and promoter methylation (average methylation > 0.75/average methylation < 0.25) for genes with high (red) or low (blue) PDR levels in the promoter.

Figure 5 DNA methylation loss in partially methylated domains. (a) Scatterplot of hypomethylated eloci as a function of their epipolymorphism in NPC and diagnosis samples (M#09). The color gradient corresponds to the point density (low is green; high is red). Two eloci populations can be distinguished. The first population (top box) had a moderate decrease in methylation associated with a strong epipolymorphism, and the second population (bottom box) had a low heterogeneity and a drastic methylation decrease at diagnosis. (b) DNA methylation levels by patient of eloci with partial demethylation (eloci with an epipolymorphism value between 0.4 and 0.8 in NPCs and remain at the same level in MM) per patient are shown at the top, and DNA methylation levels of eloci with a decreased epipolymorphism value at diagnosis (eloci with an epipolymorphism value between 0.4 and 0.8 in NPCs, and an epipolymorphism value ≤0.2 at diagnosis) per patient are shown at the bottom. Green segments show the average methylation level of NPCs for these loci. The circles under the patient labels represent the size of the population compared to all hypomethylated eloci. Patient M#19, as an example in Fig. 5a, is outlined in orange. (c) DNA methylation levels of WGBS samples inside PC-PMD regions and outside PC-PMD regions (data from the BLUEPRINT project, see Additional file 1: Table S1). (d) Representation of the global DNA methylation of PC-PMDs in WGBS samples.

Figure 6 DNA methylation loss and associated DNA functional features. (a) The average methylation levels inside PC-PMDs according to average H3K9me3 intensity (data from the BLUEPRINT project: ERX1199099 and ERX712768). Each point represents a PC-PMD, and points are colored according to the size of the PC-PMD. (b) Example of a long unmethylated PC-PMD (pink area) associated with an increase in H3K9me3 in patient MM15548 . (c) Average methylation levels in PC-PMDs according to their average H3K27me3 intensity (data from the BLUEPRINT project: ERX1199099 and ERX712769). (d) Example of PC-PMDs (pink areas) with methylation loss at diagnosis associated with an increase in H3K27me3 marks. (e) Mean DNA methylation level at diagnosis in promoter, CGIs, shores and shelves regions according to the presence or absence of PC-PMDs. (f) Proportion of CGIs with partial methylation inside and outside PC-PMDs.

Figure 7 Gene expression variability in hypomethylated PC-PMDs. (a) Variation coefficient of gene expression in PC-PMD and non-PMD regions at diagnosis (left) and relapse (right). (b) Density of the TSS distance to the nearest PC-PMD boundary. The red dotted line corresponds to the 3rd quartile of the TSS distance to the nearest PC-PMD boundary. (c) Ontological analysis of genes in PC-PMDs located less than 186 kb from a PC-PMD boundary.

Figure 8 Hypomethylated PC-PMDs and the reorganization of compartment A/B. (a) Mosaïc plot illustrating the proportions of regions inside or outside PC-PMDs and with or without an intersection with a switch from B to A (left) and from A to B (right). (b) Example of *IGF1R* gene locus switching from compartment B to A (GM12878 cell line versus MM cell lines). Pink indicates PC-PMD regions.

Additional Files
Additional file 1 — Table S1
Data resource.

Additional file 2 — Table S2
Distribution of CpG sites within 100 bp genomic units in the human genome.

Additional file 3 — Figures S1–S25
Supplementary figures.

Additional file 4 — Table S3
Description of sequencing metrics for eRRBS data.

Additional file 5 — Table S4
Average number of CpGs covered by eRRBS with read depths greater than 10X and 60X, given by genomic features.

Additional file 6 — Table S5
Odds ratios and 95% confidence intervals for association between gene expression and promoter methylation according to low or high PDR levels in the promoter.