

RESEARCH

Open Access



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

Jennifer Derrien<sup>1†</sup>, Catherine Guérin-Charbonnel<sup>1,2†</sup>, Victor Gaborit<sup>1,3</sup>, Loïc Champion<sup>1,2</sup>, Magali Devic<sup>1,4</sup>, Elise Douillard<sup>1,4</sup>, Nathalie Roi<sup>1,4</sup>, Hervé Avet-Loiseau<sup>5</sup>, Olivier Decaux<sup>6</sup>, Thierry Facon<sup>7</sup>, Jan-Philipp Mallm<sup>8</sup>, Roland Eils<sup>9,10,11</sup>, Nikhil C. Munshi<sup>12</sup>, Philippe Moreau<sup>1,4</sup>, Carl Herrmann<sup>9</sup>, Florence Magrangeas<sup>1,4†</sup> and Stéphane Minvielle<sup>1,4\*†</sup>

## 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.

**Methods:** Here, we performed an analysis of 42 MM samples from 21 patients by using enhanced reduced representation bisulfite sequencing (eRRBS). We combined several metrics of epigenetic heterogeneity to analyze DNA methylation heterogeneity in MM patients.

**Results:** We show that MM is characterized by the continuous accumulation of stochastic methylation at the promoters of development-related genes. High combinatorial entropy change is associated with poor outcomes in our pilot study and depends predominantly on partially methylated domains (PMDs). These PMDs, which represent the major source of inter- and inpatient 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.

**Conclusions:** We propose that disrupted DNA 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 inpatient heterogeneity, Transcriptomic variability

\*Correspondence: [stephane.minvielle@univ-nantes.fr](mailto:stephane.minvielle@univ-nantes.fr)

†Jennifer Derrien and Catherine Guérin-Charbonnel are co-authors.

†Florence Magrangeas and Stéphane Minvielle are co-senior authors.

<sup>1</sup>Université de Nantes, CNRS, INSERM, CRCINA, F-44000 Nantes, France

<sup>4</sup>Centre Hospitalier Universitaire, Nantes, France

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## 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 [5–8]. 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 does not alone explain the profound phenotypic variability across patients [4, 10]. Beyond genetic diversity, other processes generate the intratumoral functional heterogeneity of cancer cells, including global epigenetic changes [11].

Among epigenetic alterations, aberrant DNA methylation is a common trait of cancer cells characterized by global hypomethylation with focal hypermethylation of CpG islands (CGIs) promoters [12, 13]. Hypomethylation is associated with genomic instability [14] whereas locally hypermethylation likely contributes to the silencing of tumor suppressor genes [15, 16]. Previous studies have shown that genome wide hypomethylation is also

accompanied by a global increase of entropy in most cancers [17, 18]. Furthermore, increased intrasample methylation heterogeneity in cancers compared to their normal counterparts is associated with more pejorative outcomes in hematological malignancies [19–21]. So far, no such study has been carried out in MM. To explore the disrupted DNA methylation landscape in MM, we conducted an analysis of 42 myeloma samples from 21 patients by using enhanced reduced representation bisulfite sequencing (eRRBS). We found that the aberrant DNA methylation is associated with a high intratumor heterogeneity and that combinatorial entropy changes between normal and malignant plasma cells are significantly linked to survival outcome in this pilot study. Moreover, we showed the importance of the DNA methylation disruption on inter- and inpatient heterogeneity and transcriptomic variability.

Collectively, our data highlight the clinical and functional roles of epigenetic heterogeneity during MM development.

## Methods

### Cohorts and samples

Forty-two PC samples derived from 21 MM patients were employed for eRRBS analysis. These bone marrow samples were collected at different stages of the disease: 17 diagnosis (MMD)/relapse (MMR) paired samples, 3 smoldering multiple myeloma (SMM)/MMD paired samples, and 1 SMM/MMR paired sample. Control PCs were purified from bone marrow samples lacking abnormal plasma cells from 3 patients suspected of monoclonal gammopathy of 1 undetermined significance (MGUS). Patients, (11 women and 10 men, with a median age of 59 years) and control patients (1 woman and 2 men, with a median age of 52 years) were monitored at the Intergroupe Francophone du Myélome (IFM) centers, and all provided informed consent. These samples were analyzed using different metrics in order to determine intratumor heterogeneity (epipolymorphism and PDR) and interstates heterogeneity (combinatorial entropy, eloci, EPM) (Additional file 1: Table S1). RNA-seq analysis was also performed for these 42 PC samples.

PC samples derived from 10 MMD patients were employed for WGBS analysis among them 5 were in common with eRRBS analysis. In addition, we used publicly available datasets from BLUEPRINT Consortium [22] including 3 MMD patients and 2 control patients. All these samples were used to PMDs identification through R package MethylSeekR [23]. Chip-seq of histone marks (H3K9me3 and H3K27me3) from the 3 MMD patients (BLUEPRINT Consortium [22]) was used to analyze the redistribution of repressive histone marks in PC-PMDs.

### 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 PCs were purified using the same procedure as for myeloma cells. The average purity of control PCs was 81% (range 74–86%). The absence of abnormal plasma cells in the bone marrow aspiration was assessed by flow cytometry using the 7-color combination CD45, CD19, CD38, CD138, CD28+CD56, kappa, and lambda chains published previously [24]. 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.

### Chromosomal abnormalities analysis

Chromosomal abnormalities present in tumor plasma cells were detected either by fluorescence in situ (FISH) for t(4;14)(p16;q32) and del(17p) using specific probes provided by Abbott Molecular (Paris, France) and Cytocell (Paris, France) according to the manufacturer's instructions. The del(17p) status was retained only if present in at least 55% of the plasma cells or by single nucleotide polymorphism (SNP) array for 1q gain, del(1p32) targeting CDKN2C and hyperdiploidy previously defined as 47 chromosomes or more [25]. Hybridization-based genomic profiling arrays were performed using Genome-Wide Human SNP Array 6.0 or the CytoScan HD array, according to the manufacturer's instructions (Affymetrix, Santa Clara, CA, USA) now part of Thermo Fisher Scientific (Thermo Fisher Scientific, Inc.). Following the procedures of sample preparation, hybridization, and scanning, the CEL file of Genome-Wide Human SNP Array 6.0 was analyzed as previously described [5] and the CEL file of CytoScan HD array was analyzed using the Chromosome Analysis Suite (ChAS) software (Thermo Fisher Scientific, Inc.). Chromosomal changes of each sample were visualized by a diagram of all CNVs, the karyoview, and by visual inspection. Ploidy of tumor plasma cells for each patient was obtained using allele-specific copy number analysis of tumors algorithm (ASCAT 2.2) with default parameters [26].

### 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-Bakelman et al. [27]. 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 HiSeq™ 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 2: Table S2).

### eRRBS analysis

The adaptor sequences were removed by Cutadapt (version 1.10) [28]. FastQC (version 0.11.4) was used for quality control of the Illumina paired-end sequencing data [29]. Bisulfite reads were aligned to the bisulfite-converted hg19 genome with the nondirectional model of Bismark alignment software (version 0.14.1) [30]. 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 [31]. We calculated the epigenetic changes between two stages using methclone [32], an algorithm that detects loci of 4 adjacent CpGs (minimum depth of 60 reads), called epialleles (16 possible patterns according to CpG methylation).

### Metrics used to calculate DNA methylation heterogeneity

Epipolymorphism (Epi) was calculated for each locus of the four adjacent CpGs covered by the same read, to measure intratumor epigenetic heterogeneity, as previously described by Landan et al. [33]:

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

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 ( $\Delta S$ ) 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  $\Delta S < -70$  (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. [32]:

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

where E is the total number of loci 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).

Four possible epiallele pattern changes between two stages were determined as follows: disorder maintenance: no predominant pattern (all patterns below 30%) at either stages; selection: no predominant pattern at stage 1 (all patterns below 30%), one pattern becomes highly predominant (above 70%) at stage 2; disorder: one pattern is highly predominant (above 70%) at stage 1, no predominant pattern (all patterns below 30%) at stage 2; switch: both stages have a highly predominant pattern (above 70%) but patterns of both stages are different.

### WGBS

Data quality control and adaptor trimmed were performed with the Trimmomatic tool [34]. Read mapping was carried out with the methylCtools aligner [35]. 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 [36], allowing the analysis, management, and storage of WGBS data. An average depth of 17.2X per covered CpG was observed. As with eRRBS data, we set a coverage threshold of a minimum of 10X for each CpG sequenced for our analyses, approximately 13.4 million of CpGs per sample with a coverage of 10X were observed. The dmrseq R package was used to visualize WGBS data and smooth methylation signals [37].

### Detecting PMDs

PMDs were detected in WGBS samples using the R package MethylSeekR [23]. 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  $\alpha$  values was used to determine the presence

or absence of PMDs in samples. When the  $\alpha$  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) [38] and human genome hg19 as the reference. FastQC (version 0.11.4) [29] was used for basic quality control of the Illumina paired-end sequencing data. PCR duplicates were determined and removed using the Picard algorithm [39]. The number of reads mapped to each gene was calculated using HTSeq-count, part of the HTSeq framework [40], 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 [41].

### 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) [42] 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 eigenvector 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 [43]. 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 [44]. 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 (<http://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 CpG shores.

HOMER was used to annotate loci and eloci, using the annotatePeaks.pl script [45], 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/>).

To measure the intersections between regions, bedtools was used (more precisely, the “intersect” option) [46].

Data from the BLUEPRINT Consortium were converted to hg19 coordinates using the liftOver tool [22, 47].

### Functional annotation

GREAT tools (version 3.0.0) was used to assign biological functions by analyzing the annotations of the nearby genes [48]. 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 [49]. 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)[50].

### 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 [51], in addition to the packages already mentioned in the text, packages data.table [52], viridis, ggbio, GenomicRanges, RColorBrewer, biovizBase, grid, gridBase, MASS, fields, KernSmooth, sp, org.Hs.eg.db, DBI, and survival were used.

## Results

### Aberrant DNA methylation landscape is associated with an intratumor heterogeneity in MM

Before exploring intratumor DNA methylation heterogeneity, we confirmed using WGBS data that MM is a very hypomethylated tumor with focal hypermethylation (Additional file 1: Table S1; Additional file 3: Figure S1,S2a), as previously reported [8, 53]. Individual CpG sites showed a predominantly bimodal pattern in MM patients (Additional file 3: Figure S2b) with methylation levels depending on the local density of CpG (Additional file 3: Figure S2c; Additional file 4: Table S3). In contrast, control bone marrow plasma cells displayed striking differences in their global methylation status with a significant level of partially methylated regions that progressively increased during the terminal differentiation of germinal center B cells into tonsillar plasma cells and long-lived bone marrow plasma cells (Additional file 3: Figure S3; BLUEPRINT consortium [53]). This is likely to be due to the multiple rounds of mitotic division coupled with DNA demethylation that plasmablasts and plasma cells undergo during differentiation [54, 55]. Hence, an interesting feature of MM is that the global hypomethylation is apparently reactivated as malignant plasma cells actively divide driving a more bimodal DNA methylation pattern. Although MM displayed no evidence of aberrant DNA methylation at many genes (Additional file 3: Figure S4), we observed slight methylation gain at CpG islands (CGIs) and methylation loss at adjacent shores and shelves regions compared to normal plasma cell (NPCs; Additional file 3: Figure S5a), as illustrated in the genomic region of *DOC2B* (Additional file 3: Figure S5b).

We next explored the degree of DNA methylation intratumor heterogeneity by using eRRBS technology (Additional file 1: Table S1; Additional file 3: Figure S1). We obtained an average depth of 50X per covered CpG and

approximately 2.7 million CpGs per sample, with more than 10X sequencing coverage (Additional file 2: Table S2; Additional file 5: Table S4; Additional file 3: Figure S6). As identical average methylation may correspond to different methylation patterns (identical, uniform, disordered), reflecting homogeneous or heterogeneous cell subpopulation mixture (Fig. 1a), we analyzed eRRBS data applying a computational method that investigates DNA methylation modifications at a genomic locus defined as a group of four contiguous CpGs covered by single sequence reads [32]. We computed for each genomic locus the epipolymorphism, derived from the probability distribution of the 16 possible methylation patterns: a high value of epipolymorphism indicates that, for the specific locus, several of the 16 patterns are present in noteworthy proportions among the different reads, i.e., there is a large inter-read variability, thus a stochastic process of DNA methylation [33]. As expected, 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 S7). We plotted epipolymorphism distribution as a function of methylation level for all loci in NPCs and MM samples (Fig. 1b). 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 in MM samples. In all cases, promoter CGI methylation gain was associated with an increase in epipolymorphism in MM samples (Additional file 3: Figure S8). This intrasample heterogeneity can come from two origins: a mixture between uniformly methylated reads (i.e., whereby CpGs in an individual read are fully methylated or fully unmethylated, (Fig. 1a, example with mixture of uniformly methylated reads) or variability within reads (i.e., discordant methylation by which CpGs in an individual read are partly methylated; Fig. 1a, example with disordered patterns). In order to better characterize epiallele patterns of each sample, and differentiate between mixture of uniform reads or mixture of disordered patterns, thus assessing the degree of stochasticity in the DNA methylation process, we computed, for each locus, the proportion of discordant reads (PDR), i.e., of reads with both methylated and unmethylated CpGs [19](Fig. 1a). We found that the average PDR was significantly higher ( $p$  value <0.05, Dunnett-Tukey-Kramer's pairwise multiple comparison test) in MM samples, regardless of the disease stage, than in NPCs (Fig. 1c). These results demonstrate that methylation heterogeneity in MM arises primarily from variability within DNA reads. We also observed stochastic methylation patterns of MM in loci stratified according to the frequency of fully methylated or fully unmethylated

epialleles whereas methylation patterns of NPCs epialleles are almost exclusively uniform (Additional file 3: Figure S9,S10).

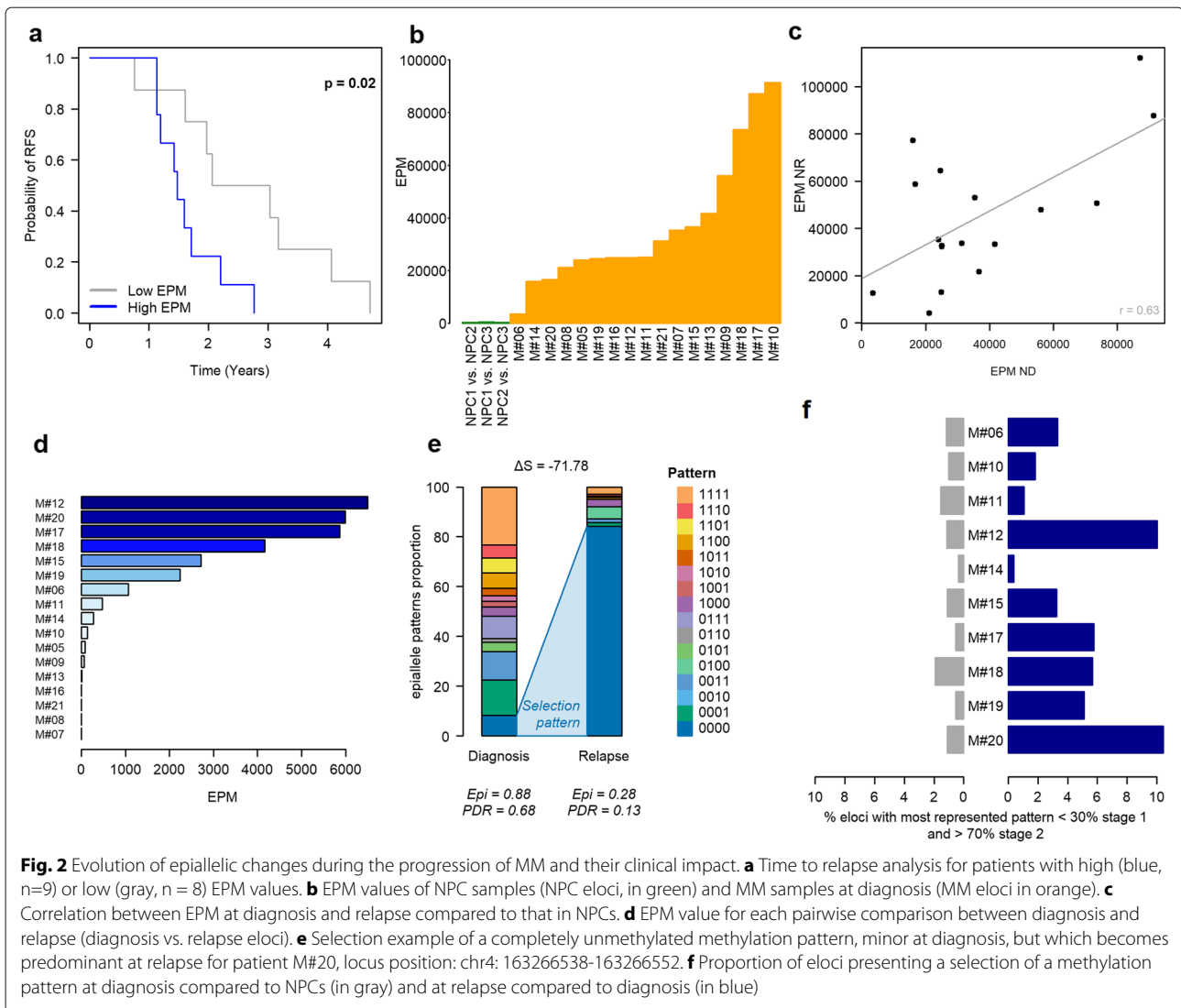
Overall, these results could imply that during disease initiation, growth, and progression, malignant PCs accumulate stochasticity in DNA methylation at the expense of a more coherent methylation state, leading to a high degree of intratumoral epigenetic heterogeneity in MM patients.

### **Combinatorial entropy changes are associated with poor survival outcome in our MM cohort**

Given that, in contrast to diffuse large B cell lymphoma, we did not find a clinical relevance linked to the level of inpatient heterogeneity [21] (Additional file 3: Figure S11), we used another DNA methylation heterogeneity metric which measures shifting between samples rather than variability within individual specimens [32]. The algorithm allows the identification of epigenetic loci (i.e., eloci) that have a significant epiallele composition change (i.e., combinatorial entropy change) between two states (e.g., normal vs. cancer or diagnosis vs. relapse) (Additional file 3: Figure S12). The number of eloci was normalized according to the number of covered loci, referred to as eloci per million loci covered (EPM) to compare genome-wide combinatorial entropy changes across patient samples (Additional file 3: Figure S12)[32] and patients were separated into two groups according to their EPM values. In this pilot study, we found that MM patients with high EPM had a significantly reduced relapse-free survival ( $p$  value = 0.02 (log-rank test); Fig. 2a). In order to assess whether the prognostic impact of EPM was independent of chromosomal abnormalities, we performed multivariable Cox regression analyses. Five specific chromosomal abnormalities, del(17p), t(4;14), del(1p32), 1q gain, and hyperdiploidy, were tested, as they were recently shown to improve classification of MM patients in the high-risk category for death (Additional file 1: Table S1; Additional file 3: Figure S13a,b)[56]. The results showed that in all regression models tested, EPM remained significant in our cohort (Additional file 6: Table S5) and suggested that combinatorial entropy changes may be indicative of more aggressive disease independently of high-risk genetic lesions. EPM was also independent of age, sex, and ploidy. These results will need further validation in an independent cohort.

Remarkably, the extent of combinatorial entropy changes varied greatly between patients at diagnosis and persisted at relapse compared to NPCs (Fig. 2b, c). To further examine the impact of treatment on combinatorial entropy changes, we assessed the epiallele shifting level during time to relapse by comparing diagnosis vs relapse paired samples (Fig. 2d). The degree of difference was highly variable from one patient to another; 24% of





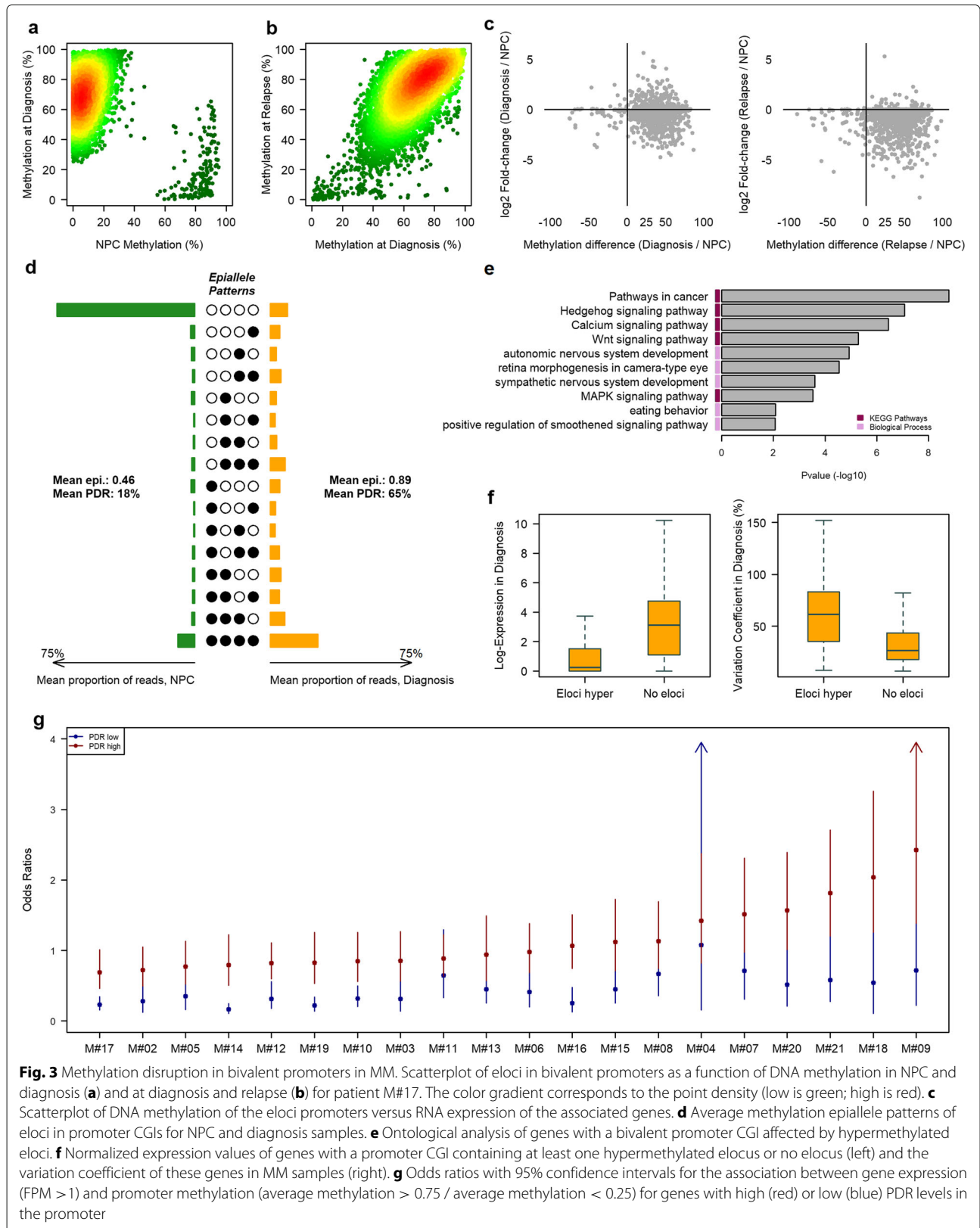
of eloci (median value = 4.2%), this finding suggests that treatment escape is associated with clonal selection at specific genomic loci.

We next sought to investigate the genomic context in which the combinatorial entropy changes occur. To answer this question, we analyzed the distribution of eloci using 127 cell/tissue types at 25-state chromatin state segmentation [57]. We observed an enrichment of eloci in bivalent promoters and quiescent states (Additional file 3: Figure S15a). We obtained similar results by analyzing published ChromHMM data for tonsil plasma cells [58] (Additional file 3: Figure S15b). Bivalent promoters are typically associated with CGIs; as expected, the distribution of eloci in bivalent promoters showed a predominant location at CGIs and promoter-TSS, whereas eloci in quiescent states overlapped with intronic and intergenic regions (Additional file 3: Figure S16a,b).

### Stochastic methylation gains at developmental genes promoters are associated with a decoupling relationship between methylation and gene expression level

Bivalent promoters play an important role during embryonic development. 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 allows low basal transcription states that are dynamically inducible to ensure a balance between self-renewal and lineage commitment [59]. Various studies have reported that these promoters are hypermethylated in cancer cell lines and primary tumors [60, 61]. As expected, 82.5% of combinatorial entropy changes found in bivalent promoters were in bivalent promoters of embryonic stem cells (ESCs) and most of the eloci acquired extensive gains of DNA methylation during neo-





plastic transformation (Fig. 3a and Additional file 3: Figure S17). In addition, DNA methylation mean of promoters with eloci was higher than the mean of promoters without eloci in all patients (Additional file 3: Figure S18). Methylation gains were maintained during progression (Fig. 3b; Additional file 3: Figure S17). These methylation gains were associated with a decrease in gene expression in diagnosis and relapse samples compared to NPCs. Interestingly, upregulated genes were present mostly at the time of diagnosis in almost all patients (Fig. 3c; Additional file 3: Figure S19). These results suggest the existence of a subpopulation which would not have undergone hypermethylation of promoters and which would overexpress certain genes. In this case, a uniform methylation pattern should be predominant. To further characterize these methylation gains, we computed the average epiallele distribution in diagnosis and control samples. Contrary to the expected hypothesis, our results revealed mostly disordered methylation gains at bivalent promoters, reflected by both high PDR and epipolymorphism instead of mixture of uniform states (Fig. 3d).

To investigate the biological relevance of disorder methylation gains, we firstly performed gene ontology (GO) analysis of the genes targeted by hypermethylated eloci. We found a strong enrichment for developmental genes. Moreover, gene set enrichment analysis revealed important cancer-related pathways, including the Wnt and MAPK signaling pathways (Fig. 3e). Then, we examined the relationship between gene expression and combinatorial entropy changes within MM patients, by comparing the expression levels between developmental genes containing hypermethylated eloci in their promoter and bivalent genes without eloci. We found that bivalent genes harboring eloci on their promoter were less expressed but displayed greater interpatient variability (Fig. 3f; only genes with a mean expression level above 1 were taken into account for the coefficient of variation).

Finally, we analyzed the impact of intratumoral methylation variability on developmental genes expression, by separating the genes into two groups according to the PDR level of their 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. 3g; Additional file 7: Table S6). In all MM samples, promoters with a low PDR, i.e., with a greater intra-read homogeneity, tended to maintain the expected opposite relationship between promoter methylation and transcription, whereas in promoters with a high PDR, i.e., with a greater intra-read heterogeneity, for the majority of patients, the link between methylation and expression did not remain significant; for some patients, we even observed a significant link with a relationship 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), demonstrated 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 S20).

Our results demonstrate that disordered DNA methylation gains target the developmental pathway as a whole rather than on specific suppressor genes, and is associated with increased transcriptional variability in MM.

### **Combinatorial entropy changes towards hypomethylation preferentially occur in PMDs**

Genomic regions that have lost their methylated state, termed partially methylated domains (PMDs), in contrast to “highly methylated domains,” were initially discovered in a fibroblast cell line [62]. Several studies have reported cancer and noncancer human primary cells with PMDs [63–67]. 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 [66]. These domains, which coincide with large-scale regions of repressive chromatin, were recently associated with vast epigenetic changes during carcinogenesis [17]. In this context, we sought to determine whether combinatorial entropy changes, especially hypomethylated eloci found in quiescent states, were related to these genomic regions. For this purpose, we performed a comprehensive analysis of PMDs in MM, using MethylSeekR [23], to analyze WGBS data from our cohort of MM samples and the available WGBS dataset (Additional file 3: Figure S1). We detected PMDs in both NPC samples and in only five MM samples from our cohort (Additional file 3: Figure S21). The PMD structure was highly similar between normal and MM cells despite a very variable level of DNA methylation (Additional file 3: Figure S22a). 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 S22b). 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. As expected, PC-PMD borders were not random and coincided with spatial genome organization, as indicated by the closeness of topological associated domain (TAD) borders to PC-PMD borders, which was larger than expected by chance ( $p < 0.001$ ). In addition, PC-PMDs shared key features with PMDs of various tumor types and normal tissues [64, 66, 67], including the correlation with lamina-associated domains (LADs) [68], late-replication timing and low gene density

(Additional file 3: Figure S23a, b, c). Except for one sample, PC-PMDs that intersected with LADs displayed a late replication time and significantly low methylation levels (Additional file 3: Figure S23d,e). We next examined DNA methylation status across patients. As expected, we found that the DNA methylation level was lower within than outside of PC-PMDs (Additional file 3: Figure S24a) and variable between patients (Additional file 3: Figure S24b). In addition, this variability was associated with the PC-PMD length and replication time (Additional file 3: Figure S25). Notably, PC-PMDs were associated with combinatorial entropy changes as the vast majority of hypomethylated loci (70.8%) and 26.7% of hypermethylated loci were located within PC-PMDs. In addition, PC-PMD methylation mean is correlated with an overall epiallele shift increase (Spearman  $p = 0.46$ ) (Additional file 3: Figure S26). These results support the notion that PC-PMD methylation loss may locally fuel epiallele shifts.

#### Severe DNA methylation loss in PC-PMDs is associated with the redistribution of repressive histone marks and perturbations in CGIs/TSSs

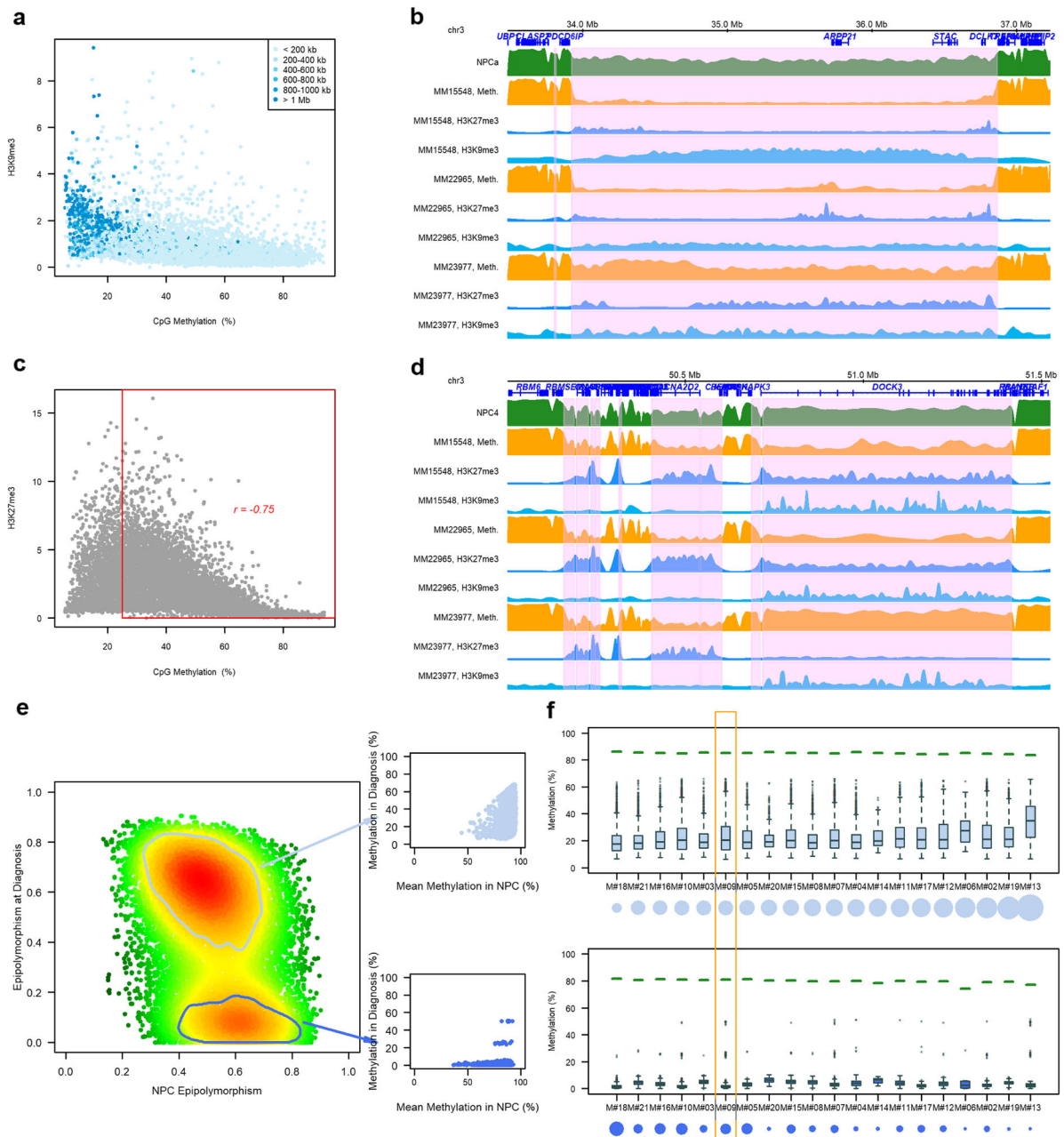
We next sought to study the relationship between DNA methylation and other epigenetic features using available WGBS data together with ChIP-seq data for histone marks (Additional file 3: Figure S1). We found that the DNA methylation losses within PC-PMDs were associated with perturbations in both H3K9me3 and H3K27me3. Notably, the H3K9me3 deposit was associated with severe DNA methylation loss in long PC-PMDs ( $> 1$  Mb) (Fig. 4a), as exemplified for the patient MM15548, with an unmethylated 3 Mb PC-PMD that was highly enriched with H3K9me3 (Fig. 4b). To our knowledge, such widespread H3K9me3 deposits associated with DNA methylation loss have so far been observed only in PMDs of cancer cell lines [35, 66, 69]. H3K27me3 was also perturbed inside PMDs. Notably, H3K27me3 enrichment was correlated with DNA methylation erosion (Fig. 4c). 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; however, in another patient (MM23977) with a methylation level in this PMD comparable to that in NPCs, the H3K27me3 mark was absent (Fig. 4d). Taken together, these results show that perturbations in repressive histone marks are variable and depend on genomic regions, patients and DNA methylation levels.

Given that some regions are hypomethylated in MM compared to NPCs while others are completely demethylated, we next studied combinatorial entropy changes that occurred in these domains using the hypomethylated loci obtained with eRRBS. We identified in most of the MM samples two subsets of hypomethylated loci compared

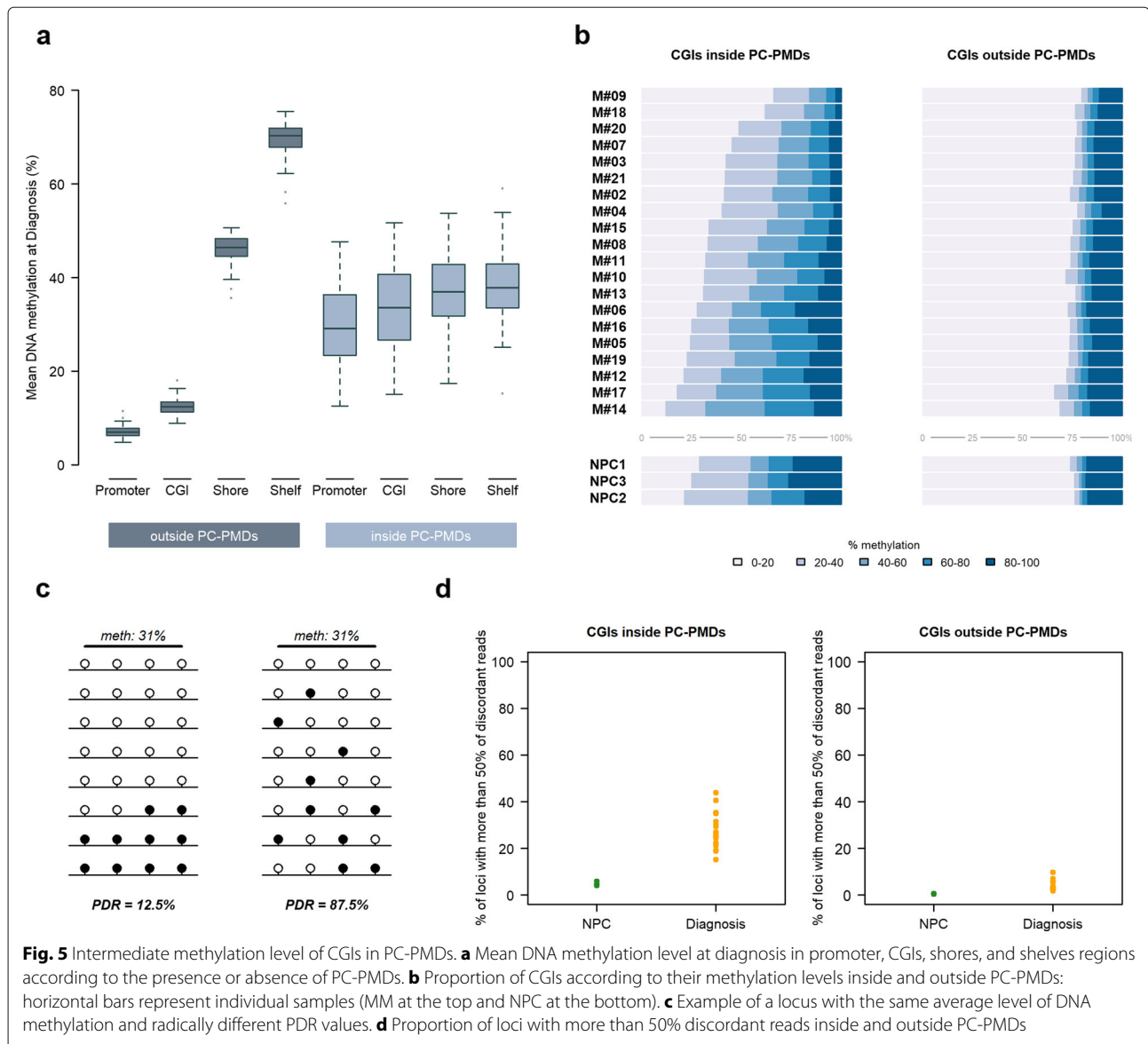
to normal samples (Fig. 4e,f; Additional file 3: Figure S27). The first type included extensively demethylated loci (Fig. 4f, bottom) and the second type contained epipoly-morphic loci emerging due to partial methylation loss (Fig. 4f, top). These loci were variable depending on the patient. Collectively, this suggests that in some genomic regions, combinatorial entropy changes are related to severe DNA methylation loss, which could be replaced by H3K9me3, whereas some domains are characterized by predominant stochastic DNA methylation loss, reflected by high epipoly-morphism associated to intratumor heterogeneity. These domains would be more associated with a redistribution of the H3K27me3 mark.

We next investigated the impact of PC-PMD methylation loss on regulatory regions. We found that the normal near-bimodal methylation state, observed outside of PC-PMDs, was completely abolished inside PC-PMDs (Fig. 5a). 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 of CGIs at the expense of strictly methylated or unmethylated states in all MM patients examined (Fig. 5b), in agreement with results obtained in breast cancer [64]. Interestingly, this phenomenon was also observed, albeit to a lesser extent, in NPCs (Fig. 5b). As identical average methylation may correspond to uniform or disordered states (Fig. 5c), we sought to discriminate between both states by calculating the PDR. We showed that the percentage of loci with a high level of discordant reads ( $> 50\%$ ) was higher in all MM compared to control patients, indicating that although PC-PMD methylation is disturbed in NPCs, CGI methylation patterns are more homogeneous (Fig. 5d). 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 S28).

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. 6a). Interestingly, we found that genes inside PC-PMDs were abundant within 186 kb of PC-PMD boundaries (Fig. 6b). 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



**Fig. 4** Redistribuition of repressive histone marks and combinatorial entropy changes in 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)** 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. **(f)** 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. 5e, is outlined in orange



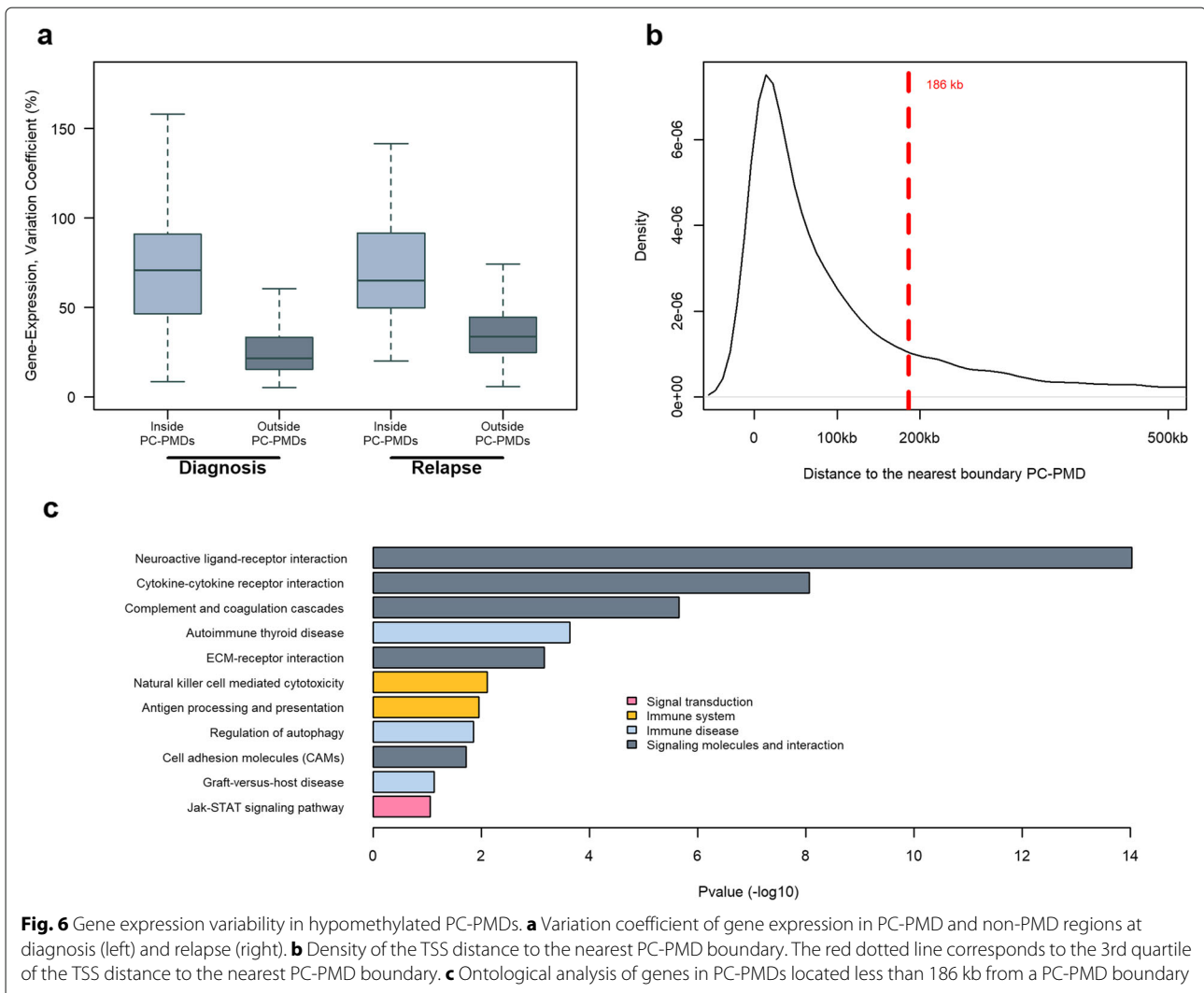
adhesion molecule, and graft-versus-host disease (Fig. 6c). This enrichment pattern in PC-PMDs could provide an explanation for immune evasion mechanisms that occur during tumor progression [70, 71].

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

### 3D genome architecture reorganization is favored in hypomethylated PMDs

Global DNA methylation loss impacts spatial genome organization in the nucleus [72]. 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 [73]. 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 [74, 75]. 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 [76], and in the myeloma cell lines U266 and RPMI8226. We found that PC-PMDs were more prone to switch than other genomic regions (Fig. 7a). 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. 7b).

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

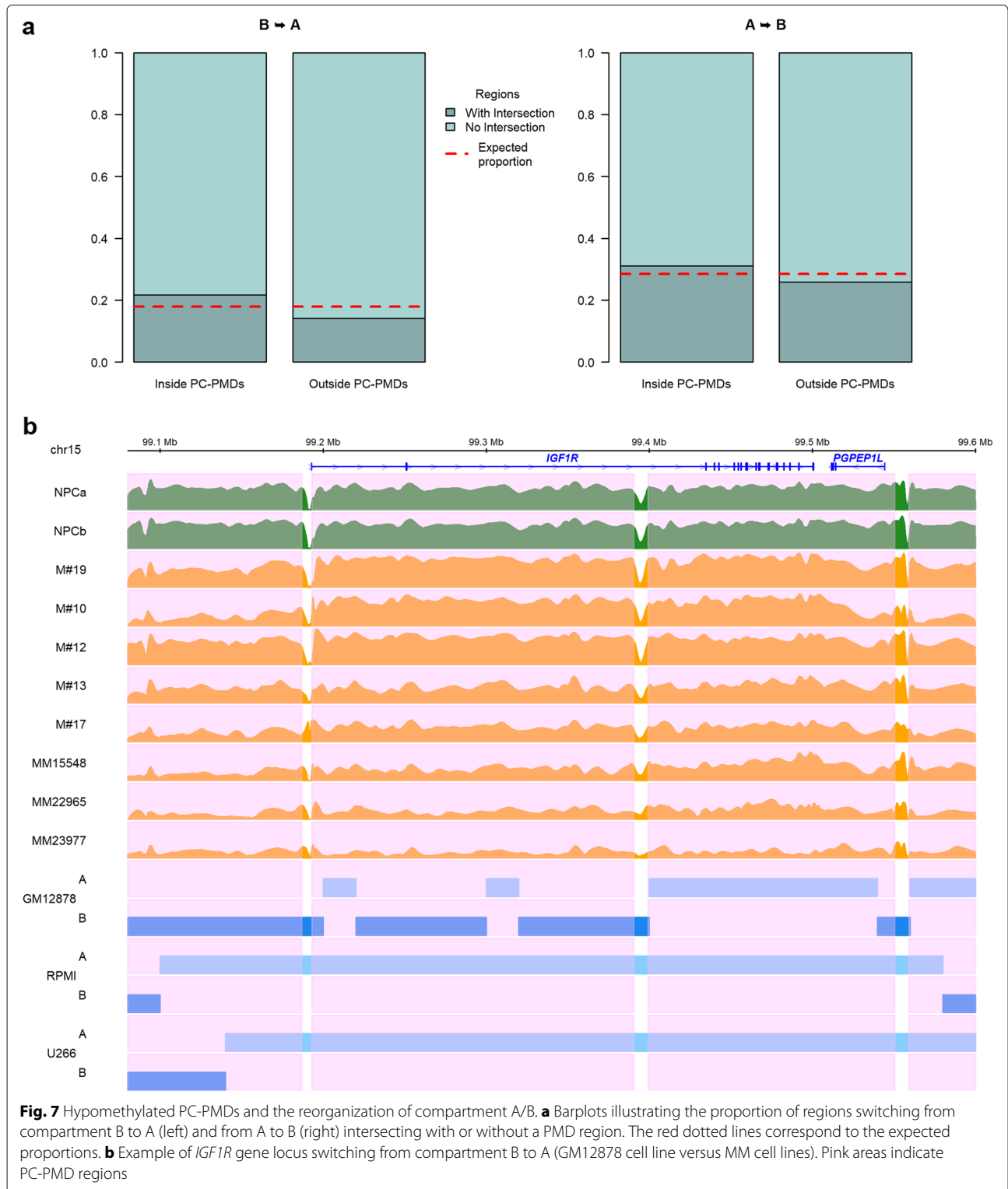
### Discussion

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 chronic lymphocytic leukemia (CLL) evolution after therapy [77]. To assess

inpatient 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 various DNA methylation heterogeneity metrics, we showed that the MM epigenome is characterized by high intratumor heterogeneity with predominantly stochastic methylation patterns.

The fact that the acquisition at the time of diagnosis of a higher EPM confers a poor survival outcome independently of high-risk genetic lesions in our modest cohort provides strong arguments to validate the prognostic impact of epiallele shifting in newly diagnosed MM patients enrolled in phase 3 trials.

We found that loci that undergo epiallele shifting associated with a gain of DNA methylation are enriched within bivalent promoter-associated chromatin states and target developmental genes. De novo DNA methylation of bivalent promoter has already been reported in many other



cancer cell lines and primary tumors and more recently during resetting of hESCs to the naïve state [60, 61, 78]. We demonstrated that in MM these DNA methylation gains are associated with high heterogeneity, in agreement with results obtained in acute myeloid leukaemia [18]. The purpose of hypermethylation of a large number of loci targeting developmental genes remains poorly understood and a topic of debate. To this end, a study has recently established a functional link between the hypermethylation of the CGI promoter of these genes and oncogenic transformation, demonstrating a causal relationship between the hypermethylation and the acceleration of the transformation [79]. 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 [80]. Further analysis of this aberrant DNA methylation program revealed strong enrichment of biological functions associated with developmental regulation (Additional file 3: Figure S29), suggesting that a disruption in developmental pathways does not prevent differentiation into plasma cells but could play a key role in the initiation of MM and increase susceptibility to oncogene transformation in response to environmental changes [81]. Moreover, considering the link between age and cancer predisposition demonstrated by Tao et al. [79], we can assume that disturbance of developmental genes in MM is partially age-related. This warrants further investigation particularly in pre-malignant condition.

During normal development and differentiation, these genes are not regulated by the presence of promoter DNA methylation but by repressive histone modifications [82], which allows low basal transcription states that are dynamically inducible to ensure balance between stem cell self-renewal and lineage/differentiation. We demonstrate that gains of DNA methylation that we observed in promoter of developmental genes in MM are associated 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 [19]. 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 [19]. Several studies have also shown the role of cellular heterogeneity and gene expression noise in overcoming drug resistance or metastatic barriers [83–85]. Together, these data suggest that increased stochastic methylation variation allows tumor cells to better adapt and find new trajectories in response to environmental changes or under treatment pressure.

A large majority of combinatorial 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 [17]. These large regions that cover approximately 70% of the genome are the major source of variability within MM patients.

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 [86, 87]; this finding is of particular interest since placental and cancerous tissues share relevant features such as immune modulation, angiogenesis, and tissue invasion [88]. 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. According to the model of Reddington et al. [89], 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* [90] and the de novo bivalent promoters [91] observed in MM. H3K27me3 redistribution could also play an important role in chromatin decompaction, as has been shown in ESCs [92], and could explain the increase in open chromatin regions within the heterochromatin state of MM samples [93]. These data highlight that PMDs must be considered as a separate entity in genomic analyses.

Finally, 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 [94].

We cannot rule out the possibility that other genomic regions, particularly active enhancers, are also associated with intratumor heterogeneity in MM as it was previously shown in two other hematological malignancies [18, 94]. A comprehensive analysis using WGBS combined with analysis methods that encapsulate underlying epigenetic variability and patient-specific histone marks could address this question.

## Conclusions

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. As numerous studies showed that hypomethylation is associated to genomic instability [14, 95] and given that genetic stability was altered within breast PMDs [64], we can assume that PC-PMDs are domains in which instability at the genetic level is also tolerated. This disturbance in PMDs could explain, in part, some aberrant and heterogeneous phenotypes across MM samples [96, 97]. 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 S30).

Finally, our results confirm that DNA methylation is already disturbed in NPCs but yet remains greatly homogeneous with mostly uniform reads. We can assume that accumulation of aberrant DNA methylation coupled with genetic alterations at the premalignant stages could reach a “point-of-no-return” which would lead to a hypomethylated malignant stage associated with intratumor epigenetic heterogeneity and a mixture of genetic subclones.

## Abbreviations

3D: Three-dimensional; bmPC: Plasma cell from bone marrow; CGI: CpG island; CLL: Chronic lymphocytic leukemia; DAVID: Database for annotation, visualization and integrated discovery; eloci: Epigenetic loci; EPM: Eloci per million loci covered; eRRBS: Enhanced reduced representation bisulfite sequencing; ESCs: embryonic stem cells; FDR: False discovery rate; FPM: Mapped read counts per million of mapped fragments; GC: Germinal center B cell; HPC: Hematopoietic progenitor cell; ICED: Iterative correction and eigenvector decomposition; IFM: Intergroupe francophone du myélome; KEGG: Kyoto Encyclopedia of Genes and Genomes; LAD: Lamina-associated domain; memBC: Memory B cell; MGUS: Gammopathy of undetermined significance; MM: Multiple myeloma; NBC: naïve B cell; NPCs: Normal plasma cells; OR: Odd ratio; PCs: Plasma cells; PDR: Proportion of discordant reads; PMDs: Partially methylated domains; preBC: Precursor B cell; RFS: Relapse-free survival; RNA-seq: RNA sequencing; RRBS: Reduced representation bisulfite sequencing; SMM: Smoldering multiple myeloma; SNPs: Single-nucleotide

polymorphism; TAD: Topological associated domain; tPC: Plasma cell from tonsil; TSS: Transcription start site; TTS: Transcription termination site; WGBS: Whole genome bisulfite sequencing

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-021-00938-3>.

**Additional file 1:** Table S1. Data resource.

**Additional file 2:** Table S2. Description of sequencing metrics for eRRBS data.

**Additional file 3:** Figures S1–S30. Supplementary figures.

**Additional file 4:** Table S3. Distribution of CpG sites within 100 bp genomic units in the human genome.

**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. Multivariable Cox regression models for RFS including EPM along with the 5 high-risk chromosomal changes (a) all variables in a single model, (b) stepwise method.

**Additional file 7:** Table S6. Odds ratios and 95% confidence intervals for association between gene expression and promoter methylation according to low or high PDR levels in the promoter.

## Acknowledgements

This study makes use of data generated by the BLUEPRINT Consortium [22]. A full list of the investigators who contributed to the generation of the data is available from [www.blueprint-epigenome.eu](http://www.blueprint-epigenome.eu). Funding for the project was provided by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 282510 – BLUEPRINT.

## Authors' 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. All authors read and approved the final manuscript.

## Funding

This study was supported by the International Myeloma Foundation (IMF-R16103NN), Fondation Française Pour la Recherche contre le Myélome et les gammopathies monoclonales (FFRMG), the I-SITE NexT (ANR-16-IDEX-0007), and the SIRIC ILIAD (INCA-DGOS-Inserm-12558).

## Availability of data and materials

SNArray, WGBS, eRRBS, and RNA-seq data have been deposited at the European Genome-phenome Archive (EGA, <https://www.ebi.ac.uk/ega>), which is hosted by the EBI and the CRG, under dataset accession EGAD00010002102 (<https://ega-archive.org/datasets/EGAD00010002102>), EGAD00001007821 (<https://ega-archive.org/datasets/EGAD00001007821>), EGAD00001007822 (<https://ega-archive.org/datasets/EGAD00001007822>), and EGAD00001007813 (<https://ega-archive.org/datasets/EGAD00001007813>), respectively.

WGBS Blueprint data of normal plasma cells and multiple myeloma patient are available from the EGA under the accessions EGAD00001002322 (<https://ega-archive.org/datasets/EGAD00001002322>) and EGAD00001002521 (<https://ega-archive.org/datasets/EGAD00001002521>) respectively [22, 53, 76]. ChIP-Seq Blueprint data of multiple myeloma patient are available from the EGA under the accessions EGAD00001002379 (<https://ega-archive.org/datasets/EGAD00001002379>) [22].

## Declarations

### 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 Myélome (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](http://www.clinicaltrials.gov)) or routine

practice in France. All patients provided written informed consent prior to sample and data collection. 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

<sup>1</sup>Université de Nantes, CNRS, INSERM, CRCINA, F-44000 Nantes, France. <sup>2</sup>Institut de Cancérologie de l'Ouest, Nantes-Saint Herblain, France. <sup>3</sup>LS2N, CNRS, Université de Nantes, Nantes, France. <sup>4</sup>Centre Hospitalier Universitaire, Nantes, France. <sup>5</sup>Institut Universitaire du Cancer, CHU, Centre de Recherche en Cancérologie de Toulouse, INSERM 1037, Toulouse, France. <sup>6</sup>Centre Hospitalier Universitaire, Rennes, France. <sup>7</sup>Centre Hospitalier Universitaire, Lille, France. <sup>8</sup>Research Group Genome Organization & Function, DKFZ, and BioQuant Heidelberg, 69120 Heidelberg, Germany. <sup>9</sup>Health Data Science Unit, Medical Faculty Heidelberg and BioQuant, 69120 Heidelberg, Germany. <sup>10</sup>Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Charitéplatz 1, 10117 Berlin, Germany. <sup>11</sup>Berlin Institute of Health (BIH), Center for Digital Health, Anna-Louisa-Karsch-Strasse 2, 10178 Berlin, Germany. <sup>12</sup>Dana-Farber Cancer Institute, Harvard Medical School, LeBow Institute for Myeloma Therapeutics and Jerome Lipper Center for Multiple Myeloma Research, Boston, MA, United States.

Received: 17 August 2020 Accepted: 9 July 2021

Published online: 09 August 2021

#### References

- Leung-Hageteijn C, Erdmann N, Cheung G, Keats JJ, Stewart AK, Reece DE, Chung KC, Tiedemann RE. Xbp1s-negative tumor B cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. *Cancer Cell*. 2013;24(3):289–304. <https://doi.org/10.1016/j.ccr.2013.08.009>.
- Zhang X-D, Baladandayuthapani V, Lin H, Mulligan G, Li B, Esseltine D-LW, Qi L, Xu J, Hunziker W, Barlogie B, Usmani SZ, Zhang Q, Crowley J, Hoering A, Shah JJ, Weber DM, Manasanch EE, Thomas SK, Li B-Z, Wang H-H, Zhang J, Kuitaie S, Tang J-L, Wang H, He J, Yang J, Milan E, Cenci S, Ma W-C, Wang Z-Q, Davis RE, Yang L, Orlowski RZ. Tight junction protein 1 modulates proteasome capacity and proteasome inhibitor sensitivity in multiple myeloma via EGFR/JAK1/STAT3 signaling. *Cancer Cell*. 2016;29(5):639–52. <https://doi.org/10.1016/j.ccell.2016.03.026>.
- Bolli N, Avet-Loiseau H, Wedge DC, Loo PV, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, Hinton JW, Li Y, Tubio JMC, McLaren S, Meara SO, Butler AP, Teague JW, Mudie L, Anderson E, Rashid N, Tai Y-T, Shammam MA, Sperling AS, Fulcinitti M, Richardson PG, Parmigiani G, Magrangeas F, Minvielle S, Moreau P, Attal M, Facon T, Futreal PA, Anderson KC, Campbell PJ, Munshi NC. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun*. 2014;5(1):. <https://doi.org/10.1038/ncomms3997>.
- Weinhold N, Ashby C, Rasche L, Chavan SS, Stein C, Stephens OW, Tytarenko R, Bauer MA, Meissner T, Deshpande S, Patel PH, Buzder T, Molnar G, Peterson EA, van Rhee F, Zangari M, Thanendrarajan S, Schinke C, Tian E, Epstein J, Barlogie B, Davies FE, Heuck CJ, Walker BA, Morgan GJ. Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. *Blood*. 2016;128(13):1735–44. <https://doi.org/10.1182/blood-2016-06-723007>.
- Magrangeas F, Avet-Loiseau H, Gouraud W, Lodé L, Decaux O, Godmer P, Garderet L, Voillat L, Facon T, Stoppa AM, Marit G, Hulin C, Casassus P, Tiab M, Voog E, Randriamalala E, Anderson KC, Moreau P, Munshi NC, Minvielle S. Minor clone provides a reservoir for relapse in multiple myeloma. *Leukemia*. 2012;27(2):473–81. <https://doi.org/10.1038/leu.2012.226>.
- Keats JJ, Chesi M, Egan JB, Garbitt VM, Palmer SE, Braggio E, Wier SV, Blackburn PR, Baker AS, Dispenzieri A, Kumar S, Rajkumar SV, Carpten JD, Barrett M, Fonseca R, Stewart AK, Bergsagel PL. Clonal competition with alternating dominance in multiple myeloma. *Blood*. 2012;120(5):1067–76. <https://doi.org/10.1182/blood-2012-01-405985>.
- Egan JB, Shi C-X, Tembe W, Christoforides A, Kurdoglu A, Sinari S, Middha S, Asmann Y, Schmidt J, Braggio E, Keats JJ, Fonseca R, Bergsagel PL, Craig DW, Carpten JD, Stewart AK. Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood*. 2012;120(5):1060–6. <https://doi.org/10.1182/blood-2012-01-405977>.
- Walker BA, Wardell CP, Melchor L, Hulkki S, Potter NE, Johnson DC, Fenwick K, Kozarewa I, Gonzalez D, Lord CJ, Ashworth A, Davies FE, Morgan GJ. Intracлонаl heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. *Blood*. 2012;120(5):1077–86. <https://doi.org/10.1182/blood-2012-03-412981>.
- Maura F, Bolli N, Angelopoulos N, Dawson KJ, Leongamornlert D, Martincorena I, Mitchell TJ, Fullam A, Gonzalez S, Szalat R, Abascal F, Rodriguez-Martin B, Samur MK, Glodzik D, Roncador M, Fulcinitti M, Tai YT, Minvielle S, Magrangeas F, Moreau P, Corradini P, Anderson KC, Tubio JMC, Wedge DC, Gerstung M, Avet-Loiseau H, Munshi N, Campbell PJ. Genomic landscape and chronological reconstruction of driver events in multiple myeloma. *Nat Commun*. 2019;10(1):. <https://doi.org/10.1038/s41467-019-11680-1>.
- Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505(7484):495–501. <https://doi.org/10.1038/nature12912>.
- Marusyk A, Janiszewska M, Polyak K. Intratumor heterogeneity: the rosetta stone of therapy resistance. *Cancer Cell*. 2020;37(4):471–84. <https://doi.org/10.1016/j.ccell.2020.03.007>.
- Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, Noushmehr H, Lange CPE, van Dijk CM, Tollenaar RAEM, Berg DVD, Laird PW. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina associated domains. *Nat Genet*. 2011;44(1):40–46. <https://doi.org/10.1038/ng.969>.
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*. 2008;454(7205):766–70. <https://doi.org/10.1038/nature07107>.
- Eden A. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*. 2003;300(5618):455. <https://doi.org/10.1126/science.1083557>.
- Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet*. 2002;3(6):415–28. <https://doi.org/10.1038/nrg816>.
- Sproul D, Meehan RR. Genomic insights into cancer-associated aberrant CpG island hypermethylation. *Brief Funct Genom*. 2013;12(3):174–90. <https://doi.org/10.1093/bfpg/els063>.
- Jenkinson G, Pujadas E, Goutsias J, Feinberg AP. Potential energy landscapes identify the information-theoretic nature of the epigenome. *Nat Genet*. 2017;49(5):719–29. <https://doi.org/10.1038/ng.3811>.
- Koldobskiy MA, Abante J, Jenkinson G, Pujadas E, Tetens A, Zhao F, Tryggvadottir R, Idrizi A, Reinisch A, Majeti R, Goutsias J, Feinberg AP. A dysregulated DNA methylation landscape linked to gene expression in MLL-rearranged AML. *Epigenetics*. 2020;15(8):841–58. <https://doi.org/10.1080/15592294.2020.1734149>.
- Landau DA, Clement K, Ziller MJ, Boyle P, Fan J, Gu H, Stevenson K, Sougnez C, Wang L, Li S, Kotliar D, Zhang W, Ghandi M, Garraway L, Fernandes SM, Livak KJ, Gabriel S, Gnirke A, Lander ES, Brown JR, Neuberger D, Kharchenko PV, Hacohen N, Getz G, Meissner A, Wu CJ. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell*. 2014;26(6):813–25. <https://doi.org/10.1016/j.ccell.2014.10.012>.
- Li S, Garrett-Bakelman FE, Chung SS, Sanders MA, Hricik T, Rapaport F, Patel J, Dillon R, Vijay P, Brown AL, Perl AE, Cannon J, Bullinger L, Luger S, Becker M, Lewis ID, To LB, Delwel R, Löwenberg B, Döhner H, Döhner K, Guzman ML, Hassane DC, Roboz GJ, Grimwade D, Valk PJM, Andrea RJD, Carroll M, Park CY, Neuberger D, Levine R, Melnick AM, Mason CE. Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. *Nat Med*. 2016;22(7):792–9. <https://doi.org/10.1038/nm.4125>.
- Pan H, Jiang Y, Boi M, Tabbò F, Redmond D, Nie K, Ladetto M, Chiappella A, Cerchietti L, Shaknovich R, Melnick AM, Inghirami GG, Tam

- W, Elemento O. Epigenomic evolution in diffuse large B-cell lymphomas. *Nat Commun.* 2015;6(1): <https://doi.org/10.1038/ncomms7921>.
22. Fernández JM, de la Torre V, Richardson D, Royo R, Puiggròs M, Moncunill V, Fragkogianni S, Clarke L, Flicek P, Rico D, Torrents D, de Santa Pau EC, Valencia A. The BLUEPRINT data analysis portal. *Cell Syst.* 2016;3(5):491–4955. <https://doi.org/10.1016/j.cels.2016.10.021>.
  23. Burger L, Gaidatzis D, Schübeler D, Stadler MB. Identification of active regulatory regions from DNA methylation data. *Nucleic Acids Res.* 2013;41(16):155. <https://doi.org/10.1093/nar/gkt599>.
  24. Robillard N, Wuillème S, Moreau P, Béné MC. Immunophenotype of normal and myelomatous plasma-cell subsets. *Front Immunol.* 2014;5: <https://doi.org/10.3389/fimmu.2014.00137>.
  25. Avet-Loiseau H, Leleu X, Roussel M, Moreau P, Guerin-Charbonnel C, Caillot D, Marit G, Benboubker L, Voillat L, Mathiot C, Kolb B, Macro M, Campion L, Wetterwald M, Stoppa A-M, Hulin C, Facon T, Attal M, Minvielle S, Harousseau J-L. Bortezomib plus dexamethasone induction improves outcome of patients with t(4,14) myeloma but not outcome of patients with del(17p). *J Clin Oncol.* 2010;28(30):4630–4. <https://doi.org/10.1200/jco.2010.28.3945>.
  26. Loo PV, Nordgard SH, Lingjaerde OC, Russnes HG, Rye IH, Sun W, Weigman VJ, Marynen P, Zetterberg A, Naume B, Perou CM, Borresen-Dale A-L, Kristensen VN. Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci.* 2010;107(39):16910–5. <https://doi.org/10.1073/pnas.1009843107>.
  27. Garrett-Bakelman FE, Sheridan CK, Kacmarczyk TJ, Ishii J, Betel D, Alonso A, Mason CE, Figueroa ME, Melnick AM. Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *J Visualized Experiments.* 2015;96: <https://doi.org/10.3791/52246>.
  28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17(1):10. <https://doi.org/10.14806/ej.17.1.200>.
  29. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2015. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
  30. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics.* 2011;27(11):1571–2. <https://doi.org/10.1093/bioinformatics/btr167>.
  31. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012;13(10):87. <https://doi.org/10.1186/gb-2012-13-10-r87>.
  32. Li S, Garrett-Bakelman F, Perl AE, Luger SM, Zhang C, To BL, Lewis ID, Brown AL, D'Andrea RJ, Ross ME, Levine R, Carroll M, Melnick A, Mason CE. Dynamic evolution of clonal epialleles revealed by methclone. *Genome Biol.* 2014;15(9): <https://doi.org/10.1186/s13059-014-0472-5>.
  33. Landan G, Cohen NM, Mukamel Z, Bar A, Molchadsky A, Brosh R, Horn-Saban S, Zalcenstein DA, Goldfinger N, Zundelevich A, Gal-Yam EN, Rotter V, Tanay A. Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nat Genet.* 2012;44(11):1207–14. <https://doi.org/10.1038/ng.2442>.
  34. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
  35. Hovestadt V, Jones NTW, Picelli S, Wang W, Kool M, Northcott PA, Sultan M, Stachurski K, Ryzhova M, Warnatz H-J, Ralser M, Brun S, Bunt J, Jäger N, Kleinheinz K, Erkek S, Weber UD, Bartholomae CC, von Kalle C, Lawerenz C, Eils J, Koster J, Versteeg R, Milde T, Witt O, Schmidt S, Wolf S, Pietsch T, Rutkowski S, Scheuflen W, Taylor MD, Brors B, Felsberg J, Reifemberger G, Borkhardt A, Lehrach H, Wechsler-Reya RJ, Eils R, Yaspo M-L, Landgraf P, Korshunov A, Zapatka M, Radlwimmer B, Pfister SM, Lichter P. Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature.* 2014;510(7506):537–41. <https://doi.org/10.1038/nature13268>.
  36. Hansen KD, Langmead B, Irizarry RA. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* 2012;13(10):83. <https://doi.org/10.1186/gb-2012-13-10-r83>.
  37. Korthauer K, Chakraborty S, Benjamini Y, Irizarry RA. Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. *Biostatistics.* 2018;20(3): 367–83. <https://doi.org/10.1093/biostatistics/kxy007>.
  38. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2012;29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
  39. Toolkit P. Broad Institute, GitHub Repository. 2018. <http://broadinstitute.github.io/picard/>.
  40. Anders S, Pyl PT, Huber W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics.* 2014;31(2):166–9. <https://doi.org/10.1093/bioinformatics/btu638>.
  41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12): <https://doi.org/10.1186/s13059-014-0550-8>.
  42. Servant N, Varoquaux N, Lajoie BR, 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 Biol.* 2015;16(1): <https://doi.org/10.1186/s13059-015-0831-x>.
  43. Servant N, Lajoie BR, Nora EP, Giorgetti L, Chen C-J, Heard E, Dekker J, Barillot E. HiTC: exploration of high-throughput 'c' experiments. *Bioinformatics.* 2012;28(21):2843–4. <https://doi.org/10.1093/bioinformatics/bts521>.
  44. Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun.* 2018;9(1): <https://doi.org/10.1038/s41467-017-02525-w>.
  45. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell.* 2010;38(4):576–89. <https://doi.org/10.1016/j.molcel.2010.05.004>.
  46. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841–2. <https://doi.org/10.1093/bioinformatics/btq033>.
  47. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haussler M, Harte RA, Heitner S, Hickey G, Hinrichs AS, Hubley R, Karolchik D, Learned K, Lee BT, Li CH, Miga KH, Nguyen N, Paten B, Raney BJ, Smit AFA, Speir ML, Zweig AS, Haussler D, Kuhn RM, Kent WJ. The UCSC genome browser database: 2015 update. *Nucleic Acids Res.* 2014;43(D1):670–81. <https://doi.org/10.1093/nar/gku1177>.
  48. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol.* 2010;28(5):495–501. <https://doi.org/10.1038/nbt.1630>.
  49. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2008;37(1):1–13. <https://doi.org/10.1093/nar/gkn923>.
  50. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocol.* 2008;4(1):44–57. <https://doi.org/10.1038/nprot.2008.211>.
  51. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018. <https://www.R-project.org/>.
  52. Dowle M, Srinivasan A. data.table: Extension of 'data.frame'. R package version 1.11.6. 2018. <https://CRAN.R-project.org/package=data.table>.
  53. Kulis M, Merkel A, Heath S, Queirós AC, Schuyler RP, Castellano G, Beekman R, Raineri E, Esteve A, Clot G, Verdaguer-Dot N, Duran-Ferrer M, Russiñol N, Vilarrasa-Blasi R, Ecker S, Pancaldi V, Rico D, Agueda L, Blanc J, Richardson D, Clarke L, Datta A, Pascual M, Agirre X, Prosper F, Alignani D, Paiva B, Caron G, Fest T, Muench MO, Fomin ME, Lee S-T, Wiemels JL, Valencia A, Gut M, Flicek P, Stunnenberg HG, Siebert R, Küppers R, Gut IG, Campo E, Martín-Subero JI. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet.* 2015;47(7):746–56. <https://doi.org/10.1038/ng.3291>.
  54. Barwick BG, Scharer CD, Bally APR, Boss JM. Plasma cell differentiation is coupled to division-dependent DNA hypomethylation and gene regulation. *Nat Immunol.* 2016;17(10):1216–25. <https://doi.org/10.1038/ni.3519>.
  55. Caron G, Hussein M, Kulis M, Delalay C, Chatonnet F, Pignarre A, Avner S, Lemarié M, Mahé EA, Verdaguer-Dot N, Queirós AC, Tarte K, Martín-Subero JI, Salbert G, Fest T. Cell-cycle-dependent reconfiguration

- of the DNA methylome during terminal differentiation of human B cells into plasma cells. *Cell Rep.* 2015;13(5):1059–71. <https://doi.org/10.1016/j.celrep.2015.09.051>.
56. Perrot A, Lauwers-Cances V, Tournay E, Hulin C, Chretien M-L, Royer B, Dib M, Decaux O, Jaccard A, Belhadj K, Brechignac S, Fontan J, Voillat L, Demarquette H, Collet P, Rodon P, Sohn C, Lifermann F, Orsini-Piocelle F, Richez V, Mohty M, Macro M, Minvielle S, Moreau P, Leleu X, Facon T, Attal M, Avet-Loiseau H, Corre J. Development and validation of a cytogenetic prognostic index predicting survival in multiple myeloma. *J Clin Oncol.* 2019;37(19):1657–65. <https://doi.org/10.1200/jco.18.00776>.
57. Ernst J, Kellis M. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat Biotechnol.* 2015;33(4):364–76. <https://doi.org/10.1038/nbt.3157>.
58. Beekman R, Chapaprieta V, Russiñol N, Vilarasa-Blasi R, Verdaguer-Dot N, Martens JHA, Duran-Ferrer M, Kulis M, Serra F, Javierre BM, Wingett SW, Clot G, Queirós AC, Castellano G, Blanc J, Gut M, Merkel A, Heath S, Vlasova A, Ullrich S, Palumbo E, Enjuanes A, Martín-García D, Beà S, Pinyol M, Aymerich M, Royo R, Puiggros M, Torrents D, Datta A, Lowy E, Kostadima M, Roller M, Clarke L, Flicek P, Agirre X, Prosper F, Baumann T, Delgado J, López-Guillermo A, Fraser P, Yaspo M-L, Guigó R, Siebert R, Martí-Renom MA, Puente XS, López-Otín C, Gut I, Stunnenberg HG, Campo E, Martín-Subero JI. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat Med.* 2018;24(6):868–80. <https://doi.org/10.1038/s41591-018-0028-4>.
59. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim T-K, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature.* 2007;448(7153):553–60. <https://doi.org/10.1038/nature06008>.
60. Easwaran H, Johnstone SE, Neste LV, Ohm J, Mosbrugger T, Wang Q, Aryee MJ, Joyce P, Ahuja N, Weisenberger D, Collisson E, Zhu J, Yegnasubramanian S, Matsui W, Baylin SB. A DNA hypermethylation module for the stem/progenitor cell signature of cancer. *Genome Res.* 2012;22(5):837–49. <https://doi.org/10.1101/gr.131169.111>.
61. Bernhart SH, Kretzmer H, Holdt LM, Jühling F, Ammerpohl O, Bergmann AK, Northoff BH, Doose G, Siebert R, Stadler PF, Hoffmann S. Changes of bivalent chromatin coincide with increased expression of developmental genes in cancer. *Sci Rep.* 2016;6(1):. <https://doi.org/10.1038/srep37393>.
62. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature.* 2009;462(7271):315–22. <https://doi.org/10.1038/nature08514>.
63. Schroeder DI, Blair JD, Lott P, Yu HOK, Hong D, Cray F, Ashwood P, Walker C, Korf I, Robinson WP, LaSalle JM. The human placenta methylome. *Proc Natl Acad Sci.* 2013;110(15):6037–42. <https://doi.org/10.1073/pnas.1215145110>.
64. Brinkman AB, Nik-Zainal S, Simmer F, Rodríguez-González FG, Smid M, Alexandrov LB, Butler A, Martin S, Davies H, Glodzik D, Zou X, Ramakrishna M, Staaf J, Ringné M, Sieuwerts A, Ferrari A, Morganello S, Fleischer T, Kristensen V, Gut M, van de Vijver MJ, Børresen-Dale A-L, Richardson AL, Thomas G, Gut IG, Martens JWM, Foekens JA, Stratton MR, Stunnenberg HG. Partially methylated domains are hypervariable in breast cancer and fuel widespread CpG island hypermethylation. *Nat Commun.* 2019;10(1):. <https://doi.org/10.1038/s41467-019-09828-0>.
65. Durek P, Nordström K, Gasparoni G, Salhab A, Kressler C, de Almeida M, Bassler K, Ulas T, Schmidt F, Xiong J, Glazar P, Klironomos F, Sinha A, Kinkley S, Yang X, Arrigoni L, Amirabad AD, Ardakani FB, Feuerbach L, Gorka O, Ebert P, Müller F, Li N, Frischbutter S, Schlickeiser S, Cendon C, Fröhler S, Felder B, Gasparoni N, Imbusch D, Hutter B, Zipprich G, Tauchmann Y, Reinke S, Wassilew G, Hoffmann U, Richter AS, Sieverling L, Chang H-D, Syrbe U, Kalus U, Eils J, Brors B, Manke T, Ruland J, Lengauer T, Rajewsky N, Chen W, Dong J, Sawitzki B, Chung H-R, Rosenstiel P, Schulz MH, Schultze JL, Radbruch A, Walter J, Hamann A, Polansky JK. Epigenomic profiling of human cd4+ t cells supports a linear differentiation model and highlights molecular regulators of memory development. *Immunity.* 2016;45(5):1148–61. <https://doi.org/10.1016/j.immuni.2016.10.022>.
66. Salhab A, Nordström K, Gasparoni G, Kattler K, Ebert P, Ramirez F, Arrigoni L, Müller F, Polansky JK, Cadenas C, G.Hengstler J, Lengauer T, Manke T, Walter J. A comprehensive analysis of 195 DNA methylomes reveals shared and cell-specific features of partially methylated domains. *Genome Biol.* 2018;19(1):. <https://doi.org/10.1186/s13059-018-1510-5>.
67. Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, Laird PW, Berman BP. DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat Genet.* 2018;50(4):591–602. <https://doi.org/10.1038/s41588-018-0073-4>.
68. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature.* 2008;453(7197):948–51. <https://doi.org/10.1038/nature06947>.
69. Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, Valsesia A, Ye Z, Kuan S, Edsall LE, Camargo AA, Stevenson BJ, Ecker JR, Bafna V, Strausberg RL, Simpson AJ, Ren B. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* 2011;22(2):246–58. <https://doi.org/10.1101/gr.125872.111>.
70. Jung H, Kim HS, Kim JY, Sun J-M, Ahn JS, Ahn M-J, Park K, Esteller M, Lee S-H, Choi JK. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun.* 2019;10(1):. <https://doi.org/10.1038/s41467-019-12159-9>.
71. Ryu D, Kim SJ, 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. *Clin Cancer Res.* 2019;26(4):935–44. <https://doi.org/10.1158/1078-0432.ccr-19-0694>.
72. Madakashira BP, Sadler KC. DNA methylation, nuclear organization, and cancer. *Front Genet.* 2017;8: <https://doi.org/10.3389/fgene.2017.00076>.
73. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326(5950):289–93. <https://doi.org/10.1126/science.1181369>.
74. Barutcu AR, Lajoie BR, McCord RP, Tye CE, Hong D, Messier TL, Browne G, van Wijnen AJ, Lian JB, Stein JL, Dekker J, Imbalzano AN, Stein GS. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol.* 2015;16(1):. <https://doi.org/10.1186/s13059-015-0768-0>.
75. 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. *Nat Commun.* 2017;8(1):. <https://doi.org/10.1038/s41467-017-01793-w>.
76. Agirre X, Castellano G, Pascual M, Heath S, Kulis M, Segura V, Bergmann A, Esteve A, Merkel A, Raineri E, Agueda L, Blanc J, Richardson D, Clarke L, Datta A, Russiñol N, Queirós AC, Beekman R, Rodríguez-Madoz JR, José-Enériz ES, Fang F, Gutiérrez NC, García-Verdugo JM, Robson MI, Schirmer EC, Guruceaga E, Martens JHA, Gut M, Calasanz MJ, Flicek P, Siebert R, Campo E, Miguel JFS, Melnick A, Stunnenberg HG, Gut IG, Prosper F, Martín-Subero JI. Whole-epigenome analysis in multiple myeloma reveals DNA hypermethylation of B cell-specific enhancers. *Genome Res.* 2015;25(4):478–87. <https://doi.org/10.1101/gr.180240.114>.
77. Gaiti F, Chaligne R, Gu H, Brand RM, Kothen-Hill S, Schulman RC, Grigorev K, Rizzo D, Kim K-T, Pastore A, Huang KY, Alonso A, Sheridan C, Omans ND, Biederstedt E, Clement K, Wang L, Felsenfeld JA, Bhavsar EB, Aryee MJ, Allan JN, Furman R, Gnirke A, Wu CJ, Meissner A, Landau DA. Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature.* 2019;569(7757):576–80. <https://doi.org/10.1038/s41586-019-1198-z>.
78. Patani H, Rushton MD, Higham J, Teijeiro SA, Oxley D, Cutillas P, Sproul D, Ficz G. Transition to naïve human pluripotency mirrors pan-cancer DNA hypermethylation. *Nat Commun.* 2020;11(1):. <https://doi.org/10.1038/s41467-020-17269-3>.
79. Tao Y, Kang B, Petkovich DA, Bhandari YR, In J, Brien GS-O, Kong X, Xie W, Zachos N, Maegawa S, Vaidya H, Brown S, Yen R-WC, Shao X, Thakor J, Lu Z, Cai Y, Zhang Y, Mallona I, Peinado MA, Zahnow CA, Ahuja N, Fertig E, Issa J-P, Baylin SB, Easwaran H. Aging-like spontaneous epigenetic silencing facilitates Wnt activation, stemness, and BrafV600e-induced tumorigenesis. *Cancer Cell.* 2019;35(2):315–3286. <https://doi.org/10.1016/j.ccell.2019.01.005>.

80. Vicente-Dueñas C, Romero-Camarero I, González-Herrero I, Alonso-Escudero E, Abollo-Jiménez F, Jiang X, Gutierrez NC, Orfao A, Marín N, Villar LM, Criado MCF, Pintado B, Flores T, Alonso-López D, Rivas JDL, Jiménez R, Criado FJG, Cenador MBG, Lossos IS, Cobaleda C, Sánchez-García I. A novel molecular mechanism involved in multiple myeloma development revealed by targeting MafB to haematopoietic progenitors. *EMBO J*. 2012;31(18):3704–17. <https://doi.org/10.1038/emboj.2012.227>.
81. Vaz M, Hwang SY, Kagiampakis I, Phallen J, Patil A, Hagan HMO, Murphy L, Zahnow CA, Gabrielson E, Velculescu VE, Easwaran HP, Baylin SB. Chronic cigarette smoke-induced epigenomic changes precede sensitization of bronchial epithelial cells to single-step transformation by KRAS mutations. *Cancer Cell*. 2017;32(3):360–3766. <https://doi.org/10.1016/j.ccell.2017.08.006>.
82. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander ES, Koseki H, Bernstein BE. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet*. 2008;4(10):1000242. <https://doi.org/10.1371/journal.pgen.1000242>.
83. Nguyen A, Yoshida M, Goodarzi H, Tavazoie SF. Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. *Nat Commun*. 2016;7(1):. <https://doi.org/10.1038/ncomms11246>.
84. Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, Beqiri M, Sproesser K, Brafford PA, Xiao M, Eggen E, Anastopoulos IN, Vargas-García CA, Singh A, Nathanson KL, Herlyn M, Raj A. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*. 2017;546(7658):431–5. <https://doi.org/10.1038/nature22794>.
85. Farquhar KS, Charlebois DA, Szenk M, Cohen J, Nevozhay D, Balázsi G. Role of network-mediated stochasticity in mammalian drug resistance. *Nat Commun*. 2019;10(1):. <https://doi.org/10.1038/s41467-019-10330-w>.
86. Smith ZD, 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*. 2017;549(7673):543–7. <https://doi.org/10.1038/nature23891>.
87. Lorincz MC, Schübeler D. Evidence for converging DNA methylation pathways in placenta and cancer. *Dev Cell*. 2017;43(3):257–8. <https://doi.org/10.1016/j.devcel.2017.10.009>.
88. Novakovic B, Saffery R. Placental pseudo-malignancy from a DNA methylation perspective: unanswered questions and future directions. *Front Genet*. 2013;4: <https://doi.org/10.3389/fgene.2013.00285>.
89. Reddington JP, Perricone SM, Nestor CE, Reichmann J, Youngson NA, Suzuki M, Reinhardt D, Dunican DS, Prendergast JG, Mjoseng H, Ramsahoye BH, Whitelaw E, Grealley JM, Adams IR, Bickmore WA, Meehan RR. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of polycomb target genes. *Genome Biol*. 2013;14(3):25. <https://doi.org/10.1186/gb-2013-14-3-r25>.
90. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet J-P, Ahmann GJ, Adli M, Anderson KC, Ardlie KG, Auclair D, Baker A, Bergsagel PL, Bernstein BE, Drier Y, Fonseca R, Gabriel SB, Hofmeister CC, Jagannath S, Jakubowiak AJ, Krishnan A, Levy J, Liefeld T, Lonial S, Mahan S, Mfuko B, Monti S, Perkins LM, Onofrio R, Pugh TJ, Rajkumar SV, Ramos AH, Siegel DS, Sivachenko A, Stewart AK, Trudel S, Vij R, Voet D, Winckler W, Zimmerman T, Carpten J, Trent J, Hahn WC, Garraway LA, Meyerson M, Lander ES, Getz G, Golub TR. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467–72. <https://doi.org/10.1038/nature09837>.
91. Agarwal P, Alzrigat M, Párraga AA, Enroth S, Singh U, Ungerstedt J, Österborg A, Brown PJ, Ma A, Jin J, Nilsson K, Öberg F, Kalushkova A, Jernberg-Wiklund H. 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*. 2016;7(6):6809–23. <https://doi.org/10.18632/oncotarget.6843>.
92. McLaughlin K, Flyamer IM, Thomson JP, Mjoseng HK, Shukla R, Williamson I, Grimes GR, Illingworth RS, Adams IR, Pennings S, Meehan RR, Bickmore WA. DNA methylation directs polycomb-dependent 3D genome re-organization in naive pluripotency. *Cell Rep*. 2019;29(7):1974–19856. <https://doi.org/10.1016/j.celrep.2019.10.031>.
93. Jin Y, Chen K, Paeppe AD, Hellqvist E, Krstic AD, Metang L, Gustafsson C, Davis RE, Levy YM, Surapaneni R, Wallblom A, Nahi H, Mansson R, Lin YC. Active enhancer and chromatin accessibility landscapes chart the regulatory network of primary multiple myeloma. *Blood*. 2018;131(19):2138–50. <https://doi.org/10.1182/blood-2017-09-808063>.
94. Pastore A, Gaiti F, Lu SX, Brand RM, Kulm S, Chaligne R, Gu H, Huang KY, Stamenova EK, Béguelin W, Jiang Y, Schulman RC, Kim K-T, Alonso A, Allan JN, Furman RR, Gnirke A, Wu CJ, Melnick AM, Meissner A, Bernstein BE, Abdel-Wahab O, Landau DA. Corrupted coordination of epigenetic modifications leads to diverging chromatin states and transcriptional heterogeneity in CLL. *Nat Commun*. 2019;10(1):. <https://doi.org/10.1038/s41467-019-09645-5>.
95. Sawyer JR, Tian E, Heuck CJ, Johann DJ, Epstein J, Swanson CM, Lukacs JL, Binz RL, Johnson M, Sammartino G, Zangari M, Davies FE, van Rhee F, Morgan GJ, Barlogie B. Evidence of an epigenetic origin for high-risk 1q21 copy number aberrations in multiple myeloma. *Blood*. 2015;125(24):3756–9. <https://doi.org/10.1182/blood-2015-03-632075>.
96. 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*. 2006;91(9):1234–40.
97. Capp J-P, Bataille R. Multiple myeloma exemplifies a model of cancer based on tissue disruption as the initiator event. *Front Oncol*. 2018;8: <https://doi.org/10.3389/fonc.2018.00355>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

