TITLE

DNA methylation variations are required for epithelial-to-mesenchymal transition induced by cancer-associated fibroblasts in prostate cancer cells

Christian Pistore[1^], Elisa Giannoni[2^], Tommaso Colangelo[3^§], Francesca Rizzo[4^], Elena Magnani[1], Livio Muccillo[3], Giorgio Giurato[4], Monica Mancini[1], Samantha Rizzo[1], Mila Riccardi[1], Nora Sahnane[5], Del Vescovo Valerio[7], Kamal Kishore[6], Martina Mandruzzato[1], Filippo Macchi[1], Mattia Pelizzola[6], Denti Michela A.[7], Daniela Furlan[5], Alessandro Weisz[4], Vittorio Colantuoni[3], Paola Chiarugi[2] and Ian Marc Bonapace[1*]

1 Department of Biotechnology and Life Sciences, University of Insubria, Busto Arsizio (VA), Italy
2 Department of Biomedical, Experimental and Clinical Sciences '*Mario Serio*', University of Florence, Florence, Italy
3 Department of Sciences and Technologies, University of Sannio, Benevento, Italy
4 Laboratory of Molecular Medicine and Genomics, Department of Medicine, Surgery and Dentistry '*Schola Medica Salernitana*', University of Salerno, Baronissi (SA), Italy
5 Anatomic Pathology Unit, Department of Surgical and Morphologic Sciences, University of Insubria, Varese, Italy
6 Centre for Genomic Sciences, Italian Institute of Technology, Milan Italy
7 Centre for Integrative Biology, University of Trento, Trento, Italy

^ These authors have equally contributed to the work
* Corresponding author
§ Present address: Institute for Stem-cell Biology, Regenerative Medicine and Innovative Therapies (ISBReMIT), Casa Sollievo della Sofferenza-IRCCS, San Giovanni Rotondo, Italy

Running title: DNA methylation changes are required for EMT in PCa cells

Abstract (**296 words**)

Widespread genome hypo-methylation and promoter hyper-methylation of epithelium-specific genes are hallmarks of stable EMT (epithelial-to-mesenchymal transition), which in prostate cancer (PCa) correlates with castration-resistance, cancer stem-cells generation, chemoresistance and worst prognosis.

Exploiting our consolidated *'ex-vivo'* system, we show that cancer-associated fibroblasts (CAF) released factors play pivotal roles in inducing genome methylation changes required for EMT and stemness in EMT-prone PCa cells.

By global DNA methylation analysis and RNA-seq, we provide compelling evidence that conditioned-media from CAFs explanted from two unrelated patients with advanced PCa, stimulates concurrent DNA hypo- and hyper-methylation required for EMT and stemness in PC3 and DU145, but not in LN-CaP and its derivative C4-2B, PCa cells.

CpG island (CGI) hyper-methylation associates with repression of genes required for epithelial maintenance and invasion antagonism, while activation of EMT markers and stemness genes correlate with CGI hypo-methylation. Remarkably, methylation variations and EMT regulated transcripts almost completely reverse qualitatively and quantitatively during MET.

Unsupervised clustering analysis of the PRAD TCGA dataset with the differentially expressed and methylated EMT signature, identified a gene cluster of differentially expressed genes defined by a CAF+ and AR- phenotype and worst diagnosis. This gene cluster includes the relevant factors for EMT and stemness, which display DNA methylation variations in regulatory regions inversely correlated to their expression changes, thus strongly sustaining the *ex-vivo* data.

DNMT3A dependent methylation is essential for silencing epithelial maintenance and EMT counteracting genes, such as *CDH1* and *GRHL2* that is the direct repressor of *ZEB1*, the key transcriptional factor for EMT and stemness. Accordingly, DNMT3A knock-down prevents EMT entry.

These results shed light on the mechanisms of establishment and maintenance of coexisting DNA hypo- and hyper-methylation patterns during cancer progression, the generation of EMT and cell stemness in advanced PCa, and may pave the way to new therapeutic implications.

Keywords: Epithelial-to-Mesenchymal-Transition (EMT), cancer associated fibroblasts (CAF), DNA methylation, DNMT3A, prostate cancer, UHRF1, DNMT1.

Article (4459 words)

Introduction

Epithelial-to-mesenchymal transition (EMT) and cancer stem cells (CSCs) have been associated with castration-resistant prostate cancer (CRPC), characterized by androgen-independence and have relevant roles during the metastatic process and for drug resistance development[39]. Primary tumour cells can acquire lineage plasticity by activating EMT through multiple autocrine and paracrine signalling networks between cancer cells and several micro-environmental cell types, including cancer-associated fibroblasts (CAFs)[16, 67]. Transition from an epithelial to a complete mesenchymal state occurs via intermediate steps, which define metastable or stable phenotypes, depending on the sustained presence of EMT promoting signals. These regulate transcriptional factors (ZEB1, SNAIL1/2, TWIST), which establish mutual inhibition loops with epithelial integrity modulators, such as GRHL2 and OVOL1/2[15, 31]. Remarkably, EMT can generate cells with stem-like properties in normal and transformed epithelial cells, which in cancer cells adds to metastatization the ability to self-renew, produce large colonies and resist chemotherapy[22, 30, 44, 62, 80].

EMT plasticity suggests that the epigenetic landscape is implicated in the dynamic events underlying metastable or stable transitions. Promoter hyper-methylation and the ensuing repression of *CDH1* and genes involved in cell-cell and cell-basal membrane contacts are hallmarks of stable EMT. Such promoter hyper-methylation is permanent and leads to an irreversible loss of most of the epithelial features, acquisition of a stable mesenchymal-like phenotype and a worse prognosis in many tumours[62]. Likewise, in *in vitro* models of EMT, tumour cells acquire these stable epigenetic marks after sustained cultivation in serum-rich medium[19] or by overexpressing EMT transcriptional factors[43].

How permanent DNA methylation events are acquired is still unknown. Efforts to analyse genome-wide DNA methylation variations during EMT progression led to conflicting

results, without answering the key question as to whether these variations are required for metastable EMT. A reversible EMT induced by TGF-β was associated with no changes in DNA methylation in mouse normal hepatocyte AML12[46] or with reversible DNA hyper-methylation in ovarian cancer cells and in MDCK cells[10, 11]. In contrast, Twist1 overexpression in immortalized human mammary epithelial cells induced stable EMT associated with a widespread DNA hypo-methylation[43]. In an EMT-induction model of benign prostate cells, only a weak correlation between DNA methylation alterations and gene activity was reported[34]. These approaches take into account mainly the role of single diffusible factors (TGF-β) or individual EMT transcriptional factor overexpression. They do not reproduce, instead, the complexity of the autocrine and paracrine signalling among cancer cells and microenvironmental components, their downstream pleiotropic effects, including acquisition of cancer cell stemness.

To address unanswered key questions as to whether DNA methylation variations are required for EMT and how they might be generated during tumour progression, we exploited our consolidated *'ex-vivo'* system that somehow mimics the aforementioned complex signalling networks represented by high-grade androgen-independent prostate carcinoma cells (PC3)[65] and CAFs explanted from patients with aggressive prostate carcinomas. The reciprocal interplay between these cells, but not with fibroblasts from benign prostatic hyperplasia (HPFs), recapitulates all *in vivo* events occurring during metastable EMT, loss of the epithelial and acquisition of a mesenchymal phenotype, generation of cancer stem cells and induction of spontaneous lung metastasis upon xenografting into nude mice[25-27, 63, 64].

Results

Conditioned medium from CAFs induces reversible EMT in PC3 cells.

To induce EMT, aggressive and androgen-independent prostate cancer PC3 cells[65] (low passage, see M&M) were treated for three days with conditioned media from CAFs (CM-CAF), explanted from two unrelated prostate cancer patients (clinical-pathological parameters in Materials and Methods and treatment scheme in Fig.1A). The establishment of EMT was documented by differential modulation of specific EMT markers both at the RNA and protein levels (Fig.1B and 1C)[25]. Replacement of CM-CAF with serum-free medium for further three days (CM-CAF+3) induced the reverse event (mesenchymal to epithelial transition – MET) (Fig.1B and 1C). The EMT process was induced also in the androgen-independent DU145 cells, but not in the less-aggressive, adenocarcinoma-like and androgen-dependent LN-CaP cells[65] or in its xenograft androgen-independent derivative C4-2B, which still retained high E-cadherin expression (SI-Fig2A).

CM-CAFs induces reversible DNA hyper- and hypo-methylation during EMT and MET.

In CM-CAF treated PC3 cells, DNMT1, DNMT3B and UHRF1 were strongly down-regulated both at the mRNA and protein levels, while no changes were observed in LN-CaP cells (Fig.1B and 1C). DNMT3A, instead, was slightly up-regulated during EMT (Fig.1B and 1C). Further exposure of the cells for 3 days to serum-free medium (MET), reverted both mRNAs and proteins to the original levels (Fig.1B and 1C, lanes CM-CAF+3).

Strong CpG methylation reduction of Long Interspersed Nuclear Elements (LINE1)[49] and of an intragenic LINE1 sequence (LINE1-cMet; SI-Fig.1E), were observed by bisulfite-pyrosequencing in PC3 and DU145 (SI-Fig.1D and SI-Fig.2) cells, but not in LN-CaP cells and in its derivative C4-2B (SI-Fig.1D and SI-Fig.2).

Strikingly, DNA methylation variations completely reversed upon switching to serum-free medium (SI-Fig.1D), prompting us to further investigate them by Illumina Infinium HumanMethylation 450K BeadChip array. 22% of the Illumina 450K chip CpGs displayed methylation variations with a $|\Delta\beta|\geq 0.2$ (scatter plots in Fig.2A(a,b,c) and SI-Fig.4) and were distributed along all chromosomes (circos-plot in Fig.2D; SI-Fig.3-1 and SI-Fig.3-2). DNA methylation pattern almost completely reverted during MET: only 11 CpGs did not revert, when considering a $|\Delta\beta|\geq 0.2$ (Fig.2B). This was confirmed by the complementary distribution of the 5meCpGs with respect to EMT (compare Fig.2Ad with 2Ac) and by a high correlation value ($r^2=0.997$) when the EMT 5meCpGs $\Delta\beta$ values were plotted with the MET reciprocal ones (Fig.2Ae). DNA methylation variations were mainly localized to intragenic areas, and, within them, promoter regions were the most affected (SI-Fig.S5). The heat map in Fig.2C, further confirms that variations are highly reproducible during EMT and revert during MET. Conditioned medium from CAFs derived from a second patient induced EMT and DNA methylation alterations that both qualitatively and quantitatively almost overlapped those of the first experiment (SI-Fig.6A, 6B and 6C).

Altogether, these results depict an unprecedented scenario whereby the microenvironmental promotion of EMT induces reproducible and reversible coexisting patterns of hypo- and hyper-methylation at distinct sets of genes and suggests that DNA methylation is regulated by a signalling program.

PC3 cells exposed to CM-CAF activate EMT and stemness transcriptional programs.

To investigate whether the observed DNA methylation changes are associated with gene expression variations instrumental to the EMT and MET processes, we performed expression profiling by RNA-seq. Fold-change variations (FC) ≥1.2 and ≤-1.2 of the differentially expressed (DE) genes were used to define up- and down-regulated genes,

respectively. 13,742 genes changed their expression levels in EMT and reverted during MET (Fig.3A, upper panel).

By volcano plot representation, we selected the common gene sub-sets that were statistically and significantly up- and down-regulated during EMT and MET. Comparing CM-HPF with CM-HPF+3 or CM-HPF with CM-CAF+3, virtually all DE genes were located between -2 and +2 log_2Fold-Change (log_2FC) (Fig.3A, intermediate panels). By contrast, only the 1,924 genes lying between log_2FC values ≤-2 or ≥+2, were specific to EMT (Fig.3A lower left panel; SI-Table 1). Remarkably, these events completely reverted during MET (Fig.3A lower right panel), as shown by the distribution of critical EMT markers (Fig.3A), scatter plot analysis and Venn diagram of the two conditions (Fig.3B and SI-Table 1). Thus, we considered those 1,924 genes as truly EMT-MET common DE genes that, by gene ontology analysis, were enriched among others in the EMT pathway (Fig.3C).

The RNA-seq expression levels of the genes were highly coherent with our Q-PCR results reported in Fig.1C (Fig.3D, upper panel). Interestingly, EMT induced transcriptional activation of ligands, transducers, regulators and targets of the three main pathways of staminality. Stem-like properties of EMT cells were further confirmed by the complete repression of ESRP1 and MBNL3, two EMT alternative splicing regulators required for the isoform switching of CD44[9, 72], essential for the stem-like properties of tumour cells[68]. Indeed, during EMT we observed the transcriptional silencing of three CD44 variants, CD44-004 (699aa), CD44-005 (493aa) and CD44-039 (249aa), and the maintenance of two other isoforms CD44-003 (361aa) and CD44-015 (82aa), respectively (Fig.3, lower panel and SI-Table 2 for transcription variants details). Among the TET genes, only TET1 slightly increased.

CAF conditioning can thus elicit in PC3 cells a transcriptional program that leads to the activation of cellular mobility, mesenchymal transition and, most importantly, cellular staminality. Strikingly, all these events are almost completely rescued to the original transcriptional program during MET.

Differentially expressed and methylated genes induced by CM-CAF in PC3 cells are involved in EMT, MET and cell stemness

To examine whether the acquired EMT phenotype depends on DNA methylation variations that lead to differential gene expression, we separated both DE and Not-DE genes into two sets: 1) not differentially methylated genes with $\Delta\beta$-value between -0.2 and +0.2 (DE&Not-DM, and Not-DE&Not-DM, respectively); 2) differentially methylated genes with $\Delta\beta$-values ≥0.2 or ≤-0.2 (DE&DM and Not-DE&DM, respectively). Out of the 1924 DE genes, only 1443 were annotated on the Illumina array manifest (Fig.4A, complete list in SI-Table 1), 1352 of which were DE&DM and 91 DE&Not-DM. Thus, variations in gene expression during EMT strongly associate with changes in DNA methylation. Control scatter plots are shown in SI-Fig.7A.

Venn diagrams in Fig.4A show that both hypo- and hyper-methylation might simultaneously occur within the promoter and the body of DE&DM genes. To study this complexity in greater detail, we evaluated by scatter plots the distribution of all the DM-5meC with $\Delta\beta$-values ≤-0.2, for the hypo-methylated cytosines, and with $\Delta\beta$-values ≥0.2 for the hyper-methylated cytosines (scatter plots Fig.4Bi and Bii; list in SI-Table 1). Massive DNA methylation variations occurred and hypo-methylation was the dominant trait (upper triangle). The number of hyper-methylated cytosines in DE&DM genes was ca 2-fold higher than in Not-DE&DM genes, suggesting that hyper-methylation was a distinctive feature of DE genes (lower triangle).

To better decipher the data relative to the DE&DM genes involved in EMT (Fig.4Bii), we graphically divided the DM-5meC into four gates: G1, G2, G3 and G4 (Figs. 4Bii and 4C, see legend for details). Considering separately up- and down-regulated genes among those DE&DM, we observed that hypo-methylated cytosines in down-regulated genes (Fig.4B - scatter and red box plots in 'All', 'G1' and 'G2') in CM-HPF treated cells displayed a median and q1 β-values higher than in those up-regulated (Fig.4B). Conversely, G1 and G2 gated cytosines in up-regulated genes are more de-methylated compared to those of down-regulated genes. In down-regulated genes, hyper-methylation influenced the most sparsely methylated cytosines compared to up-regulated genes (Fig.4B scatter and box plots in 'All', CM-HPF). Most interestingly, G4 gated cytosines in down-regulated genes were essentially poorly or not-methylated in CM-HPF and became methylated during EMT (Fig.4B scatter plot and red box plots in 'G4), suggesting a possible role for *de novo* methyltransferases. Remarkably, all these modifications completely reverted during MET (SI-Fig.7B).

Therefore, the methylation changes we observed in the up- and down-regulated genes are those that would be expected if DNA methylation were important to determine transcriptional variations.

Distribution of the hypo- and hyper-methylated cytosines along all genes

We then investigated the distribution of the DM-5meC with respect to the CGIs during EMT. The great majority of hyper-5meC within the CGIs were located in the down-regulated genes (≈91% of the total G3+G4 gated cytosines, histogram at the right of Fig.4c), while hypo-5meC were mainly in up-regulated genes (≈71% of the total G1+G2 gated cytosines, histogram at the right of Fig.4C). A similar behaviour was observed for 5meC variations in the Shores, while in Shelves and Open seas hypo-methylation was

mainly associated with down-regulated genes and hyper-methylation with up-regulated genes (SI-Fig.7C).

Examining the results further, we observed that 98% of all DM-5meC within the CGIs of the up-regulated genes were hypo-methylated (G1+G2, Fig.5A for box plots; 5B for scatter plots) and, strikingly, they were greatly enriched in the promoter regions (Fig.5B). Within the down-regulated genes, the great majority of the hyper-methylated cytosines were G4 gated and located in the promoter regions (Fig.5C for scatter plots; 5A for box plots). Recapitulating comprehensive scatter plots (Fig.5D) indicate that CGIs hypo-methylation is a feature of promoter regions of up-regulated genes during EMT. During MET those CpGs become hyper-methylated in down-regulated genes. Opposite events occur in gene-body regions.

As changes in the methylation status of these CGIs are implicated in gene transcription[18], our results strongly suggest that differential methylation is relevant to the transcriptional control of genes that modulate EMT and stemness, unveiling additionally a pivotal role for *de novo* DNA methyltransferases.

CpG Island DM-5meCs are associated with transcriptional regulation of EMT inducers and opposers, and genes inducing EMT and stemness.

GRHL2 and OVOL1/2 directly repress *ZEB1* promoter[14, 15, 53] that is, instead, activated by microenvironmental factors and, in turn, represses *GRHL2* and *OVOL1/2* genes, required for the maintenance of the epithelial phenotype in several cell types[3]. In our EMT model, *GRHL2*, *OVOL1/2* and *CDH1* mRNAs were strongly down-regulated and their promoter CpGs became methylated in CM-CAF (G4 gated) (SI-Fig.8, SI-Fig.9-1 and SI-Table 1). Conversely, *CDH2* and *ZEB1* mRNAs were up-regulated and their promoters hypo-methylated (G1 gated). (SI-Fig.8 and SI-Fig.9-1).

Strikingly, heavy hypo-methylation (G1 gated) or hyper-methylation (G4 gated) in the promoter and/or gene-body regions of several genes involved in the three key pathways for cancer stemness and in alternative splicing of the CD44 transcript were associated, respectively, with up- or down-regulation of the corresponding gene (SI-Fig.9-2).

Consistent with these results, we found that the transcription variants V1 and V7 of *ZEB1* were specifically transcribed in EMT and their promoter hypo-methylated (G1 gated) (SI-Fig.10A). An opposite situation occurred in the *DNMT3A* gene: hyper-methylation of five cytosines (G4 gated) in the body region of the longest transcript corresponded to transcription repression from the TSS of the *DNMT3A2* promoter (SI-Fig.10B). Most interestingly, moreover, three cytosines within the canonical *VIM* promoter were hyper-methylated (G4 gated), while nine annotated cytosines downstream of the canonical TSS and annotated either as promoter or gene body were strongly hypo-methylated (cytosine 14 to 22 in SI-Fig.10C, G1 gated), and linked with a strong transcriptional activation.

Altogether, these results suggest that transcriptional regulation of several key genes involved in EMT and stemness is associated with a consistent CpG-DM located both in canonical and in alternative promoters.

In vivo validation of the ex-vivo DE-DM signature

To validate the obtained results with *in vivo* data, we surveyed the PRAD TCGA dataset, taking into account the mRNA expression and DNA methylation profiles. To avoid underestimating the results, we selected the cases displaying tumour cell positivity ≥70%, which reduced their number from 499 to 259.

Unsupervised clustering analysis separated the *ex-vivo* gene signature into two clusters based on different gene expression levels. Only the significant the up- or down-regulated

genes were plotted in the heat-map, reducing the number to 488 (Fig.6A). A further reduction to 294 genes was obtained considering only those genes whose expression was consistent with our *ex-vivo* findings. Remarkably, EMT and stemness genes were enriched in the 294 genes (box plots Fig.6B, according genes) identifying the cluster 2 as EMT-like.

According to our ex vivo data, than 52% of Cluster-1 patients were androgen-dependent or androgen receptor positive (AR^+) and the majority of them display lower CAF markers expression[24]. In contrast, over 67% of Cluster-2 patients were androgen-independent or AR^- and the great majority showed higher CAF markers expression (Fig.6C).

The differential methylation profiles of 2144 cytosines, relative to the CGIs of the promoter regions of the 488 selected DE genes, were intersected with the differential expression matrix, to correlate DE_vs_DM genes. We did not find a statistical significant correlation, but only a trend. This is somehow expected, because not all the tumour cells are undergoing EMT and only in some cells the cytosines are permanently methylated. When we consider single EMT markers/regulators and stemness controllers, instead, we found a good inverse correlation between gene expression and DNA methylation of single cytosines within regulatory regions (Fig.6D left, colour map, SI-Table 4). Detailed analysis of selected cytosines of three sample genes (VIM - EMT marker, GRHL2 - EMT regulator and GLI3 - stemness controller) shows significant methylation variations between cluster-1 and 2 (Fig.6D right panel, box and scatter plots, all arrows in gene cartoons). The methylation pattern of most of those cytosines corresponds to our *ex-vivo* experiments (red arrows, Fig.6D right panel).

Finally, we associated Clusters-1/2 with clinico-pathological features of PRAD patients. Remarkably, genes of Cluster-2 were strictly correlated with lymph node metastasis (P=0.047), advanced Pathological T (P=0.00062) and higher Gleason score (P=0.00018) (Table 1).

These *in vivo* results support our *ex-vivo* data suggesting that in tumour specimens DNA methylation variations are associated to changes in gene expression of key EMT markers/regulators and of stemness regulators.

DNMT3A is essential for EMT entry.

Our results suggest that *de novo* DNA methylation is required for EMT. ChIP experiments were thus performed to test the binding of DNMTs to *GRHL2* and *CDH1* promoters (Fig.7A). DNMT3A, but not DNMT3B, was bound to these promoter regions only in CM-CAF (Fig.7B). In the same experimental setting, ChIP experiments conducted with Abs against specific histone modifications revealed the appearance of H3K36me3 and H3K27me3 on the promoters of the two genes. No significant changes were observed for H3K4me3 and H3K9me3 (Fig.7B). Although the later histone modification has been associated with DNA methylation, Suv39H1 was strongly down-regulated during EMT, supporting the absence of changes for H3K9me3 (SI-Fig.12).

As the H3K4me3-H3K27me3 bivalent histone mark is usually found in embryos at the regulatory regions of transcriptionally poised genes[70], capable of getting reactivated or definitively repressed upon defined stimuli, and “frequently bivalent segments” have been found associated with DNA methylation in cancer cells[6, 54], this result indicates that histone modifications, along with DNA methylation variations at these gene promoters appear not to be fixed, providing support to the dynamic nature of the EMT process.

To corroborate the requirement of DNMT3A for EMT, we knocked-down this enzyme in PC3 cells prior to the treatment with conditioned-media. *DNMT3A* silencing did not affect the expression of the various EMT hallmark proteins in CM-HPF treated cells, while it counteracted the EMT process in CM-CAFs by inhibiting the down-regulation of *GRHL2* and *CDH1* and the up-regulation of *ZEB1* and *VIM* (Fig.7D).

Collectively, these experiments demonstrate the requirement of promoter hyper-methylation of the key regulator of ZEB1, *GRHL2*, and *CDH1* for their transcriptional regulation and strongly support the notion that CM-CAF induction of EMT is dependent on DNMT3A activity. As this methyltransferase has been shown to be regulated in lung and gastric cancers by the family of miR-29[17, 20, 47] and miR-29b has been demonstrated to suppress prostate cancer metastasis by regulating EMT signalling[55], we checked the levels of these miRNAs during CM-CAF induced EMT. All three members of the miR-29 family were strongly down-regulated, supporting the increased expression of DNMT3A during EMT (Fig.7E).

Discussion

In the present study we addressed two issues: 1) whether DNA methylation variations are required for metastable EMT and the associated induction of the stemness phenotype; 2) whether and how DNA hypo- and hyper-methylations can be generated simultaneously.

How an extensive genome hypo-methylation along with a restricted CGI hyper-methylation should occur only in stable EMT remains elusive. TIsty and co-authors in an *in vitro* model have shown that sustained cultivation of immortalized human mammary epithelial cells in serum-rich medium produces an irreversible and heritable stable EMT phenotype followed by *de novo* DNA methylation at the *CDH1* promoter and other targeted sites[19]. Issa's group reported that in human mammary epithelial cells overexpressing Twist1, EMT requires prevalently epigenetic remodelling by the Polycomb and Trithorax complexes, although a diffused DNA hypo-methylation and a focal hyper-methylation is observed[43]. We report, instead, compelling evidence that CM from CAFs explanted from two unrelated patients affected by advanced PCa (Gleason 4+5) stimulate rapid, extensive, simultaneous and reversible DNA hypo- and hyper-methylation alterations required for

EMT in androgen-independent prostate PC3 and DU145 cancer cells, strikingly reflecting the *in vivo* CRPC phenotype. These reversible CGIs methylation changes are essential for the transcriptional control of EMT and stemness master genes, whose mRNA levels remarkably revert during the MET process.

EMT induction does not occur in less aggressive and androgen-dependent LN-CaP cells or its androgen-independent xenograft derivative C4-2B, likely due to high expression of E-cadherin and P-cadherin and to adherens junction functionality that render these cells unresponsive to microenvironmental cytokines[23]. This suggests that to allow cells to respond to extracellular EMT inducers (EMT-prone cells), androgen independency must associate with decreased expression of epithelial markers and appearance of epithelial phenotypic instability.

The strong down-regulation of *UHRF1, DNMT1* and *DNMT3B* appears to be instrumental for the hypo-methylation of *Vimentin*, *N-cadherin*, *ZEB1* and other genes, in line with previous reports[75, 76], especially considering the absence of large changes in TETs expression. UHRF1 silencing in PC3 cells is sufficient to generate a strong DNA demethylation[4], further supporting this view. Moreover, global DNA demethylation in naïve embryonic stem cells, primordial germ cells and pre-implantation embryos is dependent on methylation maintenance impairment, due to the reduced expression of UHRF1, strictly required for DNMT1 methylation at the replication fork[8, 59], while TETs contribution to global demethylation appears marginal[71].

At least three main reasons support that DNA hyper-methylation is due to DNMT3A-dependent methylation. First, among the three DNMTs, DNMT3A is the only one that increases during the process, although slightly. Second, in the same conditions, DNMT3A, but not DNMT3B, binds to *CDH1* and *GRHL2* promoters, two key epithelial maintenance genes. Third, and most important, its knock-down impairs EMT entry. Interestingly, Kalland

and co-workers found that DNMT3A was one of the top three most overexpressed epigenetic modifiers in PCa tissues[33]. Strikingly, transcription of *GRHL2*, the direct *ZEB1* regulator[14, 15, 53], is repressed by *de novo* methylation at CGIs *via* DNMT3A. So is the splicing factor ESRP1, which prevents a splice isoform switching of the multifunctional transmembrane glycoprotein CD44 involved in a self-enforcing feedback loop with ZEB1, resulting in EMT maintenance, metastatic colonization and CSC generation in breast, prostate and colon cancer cells[9, 42, 52, 68, 73]. Accordingly, we observe CD44 isoform switching and a strong CGI hyper-methylation leading to complete repression of *ESRP1* and *MBNL3*, another repressed splicing factor during metastatic colonization of PC3 cells[42].

Down-regulation of miR-29s can explain DNMT3A increased expression during EMT, as this miRNA family has been shown to down-regulate DNMT3A in several tissues[20, 47, 50] and miR-29b is down-regulated in prostate cancer tissues[51, 55, 58]. Remarkably, PC3 cells ectopically expressing miR-29b inhibited metastasis spreading *in vivo* and reduced metastasis relevant traits *in vitro* by negatively regulating EMT factors[55].

These unprecedented results highlight that the complex cytokine network released from CAFs - which includes PDGF-α/β, VEGF, SDF-1, MMP2-9-13-14, HGF, IL-6, CXCL14, CCL7, uPA/uPAR and possibly others[16] - can stimulate pleiotropic signalling pathways in target cancer cells that, in turn, activate a programmed pattern of DNA methylation variations and transcriptional regulation. Both these events are almost totally reversible during MET and predict a deterministic program of DNA methylation and gene expression changes, rather than a stochastic one.

Our results might explain the contradictory data previously reported using TGF-β as the only EMT inducer. Indeed, TGF-β-induced EMT does not elicit any DNA methylation change in AML12 normal mouse hepatocytes[46] or in A549 human lung cancer cells[41], and

only very limited, still reversible, variations in MDCK-II dog kidney cells[11] or SKOV3 human ovarian cancer cells[10]. Since a metastable EMT can be induced *in vivo* by multiple means resembling more closely the complex network of cytokines of our *ex-vivo* model, we suggest that DNA methylation alterations could occur also *in vivo* and be required for EMT induction. Indeed, our previous experiments clearly showed that co-injections of PC3 cells with prostate CAFs, but not with HPFs, were able to produce a high tumour incidence in SCID bg/bg mice[25].

EMT induction in our experimental model is remarkably accompanied by generation and amplification of CSCs, in line with literature data [25, 44, 57]. Transcriptional activation of several genes involved in the three main pathways of cell stemness[66] (Hedgehog, Wnt and NOTCH) is associated with strong hypo-methylation at promoter CGIs. Most importantly, it has recently been suggested that the switch from non-CSC to CSC and chemotherapy resistance occur in cancer cells independently from EMT, through a ZEB1-dependent mechanism[78]. A key ZEB1 function is to force *CD44* mRNA splicing conversion and activate the ATM-ZEB1-CHK1 signalling axis[12, 77]. Consistently, transcriptional inactivation associated with *de novo* methylation at promoters and CGIs occurs at a series of ZEB1 target genes[1], such as *CRB3, EPPK1, DSP, ST14* or in tumour invasion regulators (*PTPRG*[13], *SLFN11*[48], *SYK*[36]), modulators and genes involved in the organization of the epithelial phenotype (*CDH1*[7], *OVOL1/2*[38], *LAD1*[45], *FGFR2*[29], *COBL*[28]), or in inhibitors (*DSP*[74] and *BMP4*[79]) and components of the Wnt and NOTCH pathways (*WNT5A, FZD7, NOTCH3*). Of note, DNA hypo- and hyper-methylation are associated with the activation of transcriptional variants of at least *ZEB1*, *Vimentin* and *DNMT3A*.

DNA methylation changes related to gene expression variations have extensively been associated with advanced PCa[35]. TCGA PRAD dataset analysis remarkably shows that the PC3 DE&DM EMT signature intercepts a cluster of differentially expressed genes defined by a CAF+ and AR- phenotype and worst diagnosis. This gene cluster includes the

relevant factors for EMT and stemness, which display DNA methylation variations in regulatory regions inversely correlated to their expression changes, thus greatly sustaining the potential of our approach. Although DNA methylation variations of those select genes are relatively small, our results suggest a few considerations. *Ex-vivo* data show that DNA methylation variations are transient and in the clinical specimens not all the tumour cells would be undergoing metastable EMT and, above all, permanent DNA methylation occurs in aggressive cancer, thus only some cells will hold that stable epigenetic mark. Altogether, the TCGA PRAD dataset analysis significantly supports our *ex-vivo* data, in which DNA methylation changes are required for both metastable EMT induction and generation of CSCs and are likely to take place also PCa tumours, according to the model depicted in Fig.8. CAFs and other microenvironmental components release cytokines that in cancer cells activate signalling pathways that down-regulate *UHRF1*, *DNMT1* and *DNMT3B*. This down-regulation leads to a passive and diffuse DNA hypo-methylation achieved through inefficient methylation maintenance[40] and not necessarily associated with transcriptional activation of the corresponding genes[21], likely due to the underlying chromatin configuration. Such a mechanism has been suggested for the global demethylation events occurring in naïve ESC, where loss of H3K9me2 impairs UHRF1 binding to chromatin[71] thereby inhibiting DNMT1 recruitment to hemimethylated DNA[8, 59]. DNMT3A, possibly locally recalled by H3K36me3, induces *de novo* methylation at CGIs of several genes, including *GRHL2*, *OVOL1/2* and *ESRP1* and thus entailing their repression. As a consequence, *ZEB1* is activated and *CD44* undergoes isoform switching, resulting in induction of EMT and cancer stemness. Demethylation of stem cell genes occurs, contributing to the EMT process also through activation of *ZEB1* and *Vimentin* transcript variants.

As CRPC may be linked to the acquisition of a stable mesenchymal-like phenotype together with permanent changes in DNA methylation, according to the well-defined EMT

program associated with the development of drug resistance and adverse prognosis[62], our results open new perspectives on how these permanent DNA methylation variations may be originated and pave the way to new therapeutic implications.

Methods

The aim, design and setting of the study

The aim of the study was to address unanswered questions concerning DNA methylation variations required for EMT, and how these are generated during tumour progression.

To address these questions, we exploited our *'ex-vivo'* system, in which interrelations between prostate carcinoma cells and cancer-associated fibroblasts (CAF) mimic the *in vivo* events occurring during metastable EMT.

For the setting of the study, see Fig.1.

Cell Culture, EMT-MET processes induction

Human prostate cancers cell lines, PC3, DU145 and LN-CaP, were originally obtained from American type culture collection (ATCC - Manassas, VA, USA) and authenticated by PCR/short tandem repeat (STR) analysis. The clones actually used in this study, come from frozen vials of a 10th to 15th passage in culture and from the same set of vials that have been used for all the previous published studies cited in this paper[25-27, 63, 64]. LN-CaP C4-2B are a kind gift of Dr. Frank Claessens (Department of Cellular and Molecular Medicine, University of Leuven, Belgium)[60]. Regular mycoplasma tests were performed with the kit N-GARDE Mycoplasma PCR Reagent set (EuroClone®). PC3, DU145, cells were cultured in DMEM and LN-CaP and LN-Cap C4-2B in RPMI 1640 (EuroClone®) medium, both containing 10% heat-inactivated foetal bovine serum (FBS, EuroClone®), 2mM Glutammine and 100 U/ml of penicillin/streptomycin (EuroClone®).

Human prostate fibroblasts (HPFs and CAFs) were isolated from surgical explantation after patient informed consent in accord with the Ethics Committee of the Azienda Ospedaliera Universitaria Careggi. Briefly, HPFs and CAFs were extracted from prostate

of benign prostate hyperplasia-bearing patients and from intra-tumoural regions of prostate of PCa-bearing patients respectively. Pathological staging of the two patients undergoing radical prostatectomy revealed prostate carcinomas with extracapsular invasion (pT3a), but without lymphatic invasion (N0) (average Gleason = 4 + 5). Pre-operative PSA levels were 8.3 and 9.42 ng/ml, respectively.

Tissues were digested overnight in 1 mg/ml collagenase I, and cells plated in DMEM containing 10% FBS. PC3 and LN-CaP EMT induction was performed with a treatment for 72h with conditioned media (CM-HPF or CM-CAF) prepared following the protocol previously published [70]. (CM from HPFs and CAFs was obtained by 48-h serum-starved cells, clarified by centrifugation, and freshly used). MET process was induced in PC3 cells in EMT with further 72h of treatment with serum-free medium.

RNAi transfection

RNAi protocol was performed using siGENOME DNMT3A (D-006672 - Dharmacon-GE Heathcare Life Science), siGENOME ctrl (M-005834 - Dharmacon- GE Heathcare Life Science) and INTERFERin® (Polyplus Transfection) as transfection reagent. PC3 cells were transfected 24h before EMT induction by following INTERFERin® manufacturer's instruction.

Western blot analysis

Western blots on total cellular extracts of human prostate cancer cell lines in EMT and MET were performed with antibodies against CDH1 (1:3.000 - 610181 - BD Bioscience™), CDH2 (1:3000 - 610920 - BD Bioscience™), *Vimentin* (1:5000 - GTX100619 - GeneTex®), Snail (C15D3, 1:1000 - #3879 - Cell Signalling Technology®), Slug (C19G7, 1:1000 - #9585 - Cell Signalling Technology®), TWIST1 (1:200 - ab50887 - Abcam), TCF8/ZEB1 (D80D3, 1:1000 - #3396 - Cell Signalling Technology®), UHRF1 (1:5000 - kindly gifted by IGBMC Institute of Strasburg), DNMT1 (1:500 - ab13537 - Abcam), DNMT3A (D23G1,

1:1000 - #3598 - Cell Signalling Technology®), DNMT3B (1:1000 - ab122932 - Abcam), GRHL2 (1:500 - HPA004820 - Sigma Aldrich®), β-Actin (1:3000 - ab49900 - Abcam) and GAPDH (1:3000 - MAB347 - Merk Millipore). Chemiluminescence reactions on secondary anti-mouse and anti-rabbit antibodies (315-035-045 and 211-035-109 - Jackson ImmunoResearch Laboratories, Inc.) were activated with ECL (RPN 2209 - Amersham-GE Heathcare Life Science) and the signals were visualized after developing of autoradiography films (28906836 - Amersham-GE Heathcare Life Science).

qRT-PCR

Total RNA was isolated from cells with TRIzol® Reagent (Thermo Fisher Scientific). For mRNA quantification, cDNAs were prepared retro-transcribing 1µg of total RNA with iScript™ Reverse Transcription supermix for RT-qPCR (Bio-Rad, Hercules, CA). qRT-PCR reactions were performed with iQ SYBR Green supermix (Bio-Rad, Hercules, CA). The PCR was monitored in real-time using a MJ-Chromo4 (Bio-Rad, Hercules, CA) that allowed determination of the threshold cycle (Ct) at the beginning of the exponential amplification of the PCR products. The average Ct, from triplicate assays, was used for further calculations. Gene primers are indicated in SI-1: Table S3. Results are expressed as the mean value of triplicate experiments.

TaqMan MicroRNA Reverse Transcription kit (Life Technologies) was used for miRNAs quantification. Starting from total RNA (10 ng), miRNAwas converted to cDNA using reverse transcriptase and miRNA-specific stem-loop primers: hsa-miR-29a (002112), hsa-miR-29b (000413), hsa-miR-29c (000587). For relative miRNA quantification, snoRNA U48 (RNU48; 001006) was used as a reference gene [37, 56]. The PCR reaction (20 µl), containing 1.33 µl of cDNA, 10 µl of TaqMan 2X Universal PCR Master Mix, 1 µl of TaqMan MicroRNA Assay (20X) containing probes specific for the miRNAs of interest, was

incubated at 95°C for 10 min, and then at 95°C for 15 s and 60°C for 60 s for 40 cycles. The relative expression of miRNAs was calculated by using the comparative Ct method.

ChIP experiment

ChIP experiments in PC3 cells undergoing EMT were performed as previously published [5] with antibodies against DNMT1 (8µl - ab13537 - Abcam), DNMT3A (5µl - 39206 - Active Motif®), DNMT3B (5µl - 39207 - Active Motif®), H3K27me3 (5µl - 39155 - Active Motif®), H3K9me3 (5µl - 39765 - Active Motif®), H3K4me3 (5µl - 39159 - Active Motif®), H3K36me3 (5µl - 61102 - Active Motif®) and IgG (5µl - sc-2027 - Santa Cruz Biotechnology). IgG antibody was used as negative control of chromatin immunoprecipitation. DNA samples were analysed by performing PCR experiments with GoTaq (Promega). Primers for CDH1 and GRHL2 gene promoters are reported SI-Table 3.

Genome-wide DNA methylation analysis

Genomic DNA extraction and purification from treated cells were performed using SNET lysis buffer (10mM Tris pH 8.0, 0.1M EDTA, 0.5% SDS) and phenol/chloroform/isoamyl alcohol (24:25:1 – P1944 Sigma-Aldrich®) protocols. High quality DNA (500 ng) was bisulfite converted using EZ DNA methylation kit (Zymo Research, Irvine, CA). Bisulfite converted DNA (200 ng) was used for analysis of whole-genome methylation using the HumanMethylation 450K BeadChip (Illumina), which contains 485,577 probes covering 21,231 (99%) RefSeq genes. In brief, bisulfite converted DNA was whole-genome amplified for 20 hours followed by end-point fragmentation. Fragmented DNA was precipitated, denatured, and hybridized to the BeadChips for 20 hours at 48°C. The BeadChips were washed and the hybridized primers were extended and labelled prior to scanning the BeadChips using the Illumina iScan system. GenomeStudio® software (version 2011.1; Illumina Inc.) was used for the extraction of DNA methylation signals from scanned arrays. The methylation level for each cytosine was expressed as a Beta value

calculated as the fluorescence intensity ratio of the methylated to unmethylated versions of the probes: Beta values ranged between 0 (unmethylated) and 1 (methylated). The annotation relating to CpG islands (CGIs) uses the following categorization: "shore", each of the 2 kb-sequences flanking a CGI; "shelf", each of the 2 kb-sequences next to a shore; "open sea", DNA not included in any of the previous sequences or in CGIs 4. TSS200 and TSS1500 indicate the region between position -200 bp and -1,500 bp from the Transcription Start Site (TSS), respectively. The significant methylation difference between two given loci is indicated by a Delta-Beta value and determined with GenomeStudio Methylation Module using Illumina custom algorithm for calculating DiffScores (DiffScore ≤ -30.0 (≈ pval < 0.001) = hypomethylation; DiffScore ≥ 30.0 (≈ pval < 0.001) = hypermethylation). Raw microarray data have been deposited in the EBI ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) according to the Minimum Information About a Microarray Experiment (MIAME) guidelines with Accession Number E-MTAB-4753.

RNA-Sequencing

Indexed libraries were prepared from 1 ug purified RNA with TruSeq Stranded Total RNA Sample Prep Kit (Illumina) according to the manufacturer's instructions. Libraries were sequenced (paired-end, 2x100 cycles) at a concentration of 8pmol/L per lane on HiSeq2500 platform (Illumina) with a coverage of >70 million sequence reads/sample. The raw sequence files generated (.fastq file) underwent quality control analysis using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and the quality checked reads were then aligned to the human genome (hg19 assembly) using TopHat version 2.0.10 [69], with standard parameters. Isoforms quantification was performed using Cufflink v2.1.1 [69]. A given mRNA was considered expressed when detected by at least ≥ 10 reads. Differentially expressed mRNAs were identified using DESeq2 version 1.8.2. Firstly, gene annotation was obtained for all known genes in the human genome, as provided by Enseble (GRCh37)

(https://support.illumina.com/sequencing/sequencing_software/igenome.ilmn). Using the reads mapped to the genome, we calculated the number of reads mapping to each transcript with HTSeq-count [2]. These raw read counts were then used as input to DESeq [2] for calculation of normalized signal for each transcript in the sample and differential expression was reported as fold-change along the associate adjusted p-values (computed according to Benjamini-Hochberg). Raw RNA sequencing data have been deposited in the EBI ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) according to the Minimum Information about high-throughput SEQuencing Experiment (MINSEQE) guidelines with Accession Number E-MTAB-4752.

miRNA analysis

TaqMan MicroRNA Reverse Transcription kit (Life Technologies) was used for miRNAs quantification. Starting from total RNA (10 ng), miRNAs were converted to cDNA using reverse transcriptase and miRNA-specific stem-loop primers: hsa-miR-29a (002112), hsa-miR-29b (000413), hsa-miR-29c (000587). For relative miRNA quantification, snoRNA U48 (RNU48; 001006) was used as a reference gene [37, 56]. The PCR reaction (20 µl), containing 1.33 µl of cDNA, 10 µl of TaqMan 2X Universal PCR Master Mix, 1 µl of TaqMan MicroRNA Assay (20X) containing probes specific for the miRNAs of interest, was incubated at 95°C for 10 min, and then at 95°C for 15 s and 60°C for 60 s for 40 cycles. The relative expression of miRNAs was calculated by using the comparative Ct method.

DNA pyrosequencing

DNA bisulfite conversion was performed using Epitect kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The quantification of the methylation percentages of four consecutive CpG sites in the LINE-1 promoter region (GenBank accession number X58075) was performed by pyrosequencing using previously published protocol and primers [61] (Fig.1a and SI-1). Levels of local LINE1 methylation were analysed

using a LINE1-cMET assay amplifying a region between exon 2 and exon 3 of c-MET oncogene [32]. Three consecutive CpG sites were analysed by pyrosequencing. The primers are reported in SI-1: Table S3.

Independent PRAD TCGA dataset analysis.

The Cancer Genome Atlas (TCGA) of primary prostate adenocarcinoma (PRAD) tissues was enquired (https://gdc-portal.nci.nih.gov). Data from the Illumina Infinium DNA methylation (HumanMethylation450 array) and the IlluminaHiSeq mRNA expression profiles carried out on n=499 patients were considered and an unsupervised analysis of Differential Expressed (DE) and Differentially Methylated (DM) genes previously found in PC3-CAF3 signature performed. Among all PRAD samples, only those exhibiting tumour cell positivity ≥70% were analysed (n=259). Suitable DE and DM genes were selected according to the PC3 signature as mentioned above. Normalized genes expression levels 3 data were Z-score transformed to reduce skewness. The heat-map in Figure 6A was depicted using the publicly available software Morpheus (https://software.broadinstitute.org/morpheus/) and the "One minus Pearson correlation" hierarchical clustering of genes performed.

Statistical analyses

All experiments were performed at least three times, and all samples were analysed in triplicate. Results are presented as mean±s.d. The statistical differences of gene expression and DNA pyrosequencing were assessed, respectively, by unpaired Student's t-tests and two-way ANOVA using for both Graphpad PRISM® 6.0 software. We used the Benjamini-Hochberg method to calculate the statistical significance of Ingenuity Pathway Analysis. For box plot analysis we calculated the *p* value by Wilcoxon-Mann-Whitney test.

All analyses were performed using the R Bioconductor software and SPSS v17 Statistical Package (SPSS, Chicago, IL, USA).

Association between PRAD clusters 1/2 (differential gene expression levels; Figure 6A) and phenotypic characteristics was evaluated by the Pearson's chi-squared test (Table 1).

List of abbreviations

EMT: epithelial-to-mesenchymal transition
MET: mesenchymal-to-epithelial transition
PCa: prostate carcinoma
CRPC: castration-resistant prostate cancer
HPF: fibroblasts from benign prostatic hyperplasia
CAF: cancer associated fibroblast
CM-HPF: conditioned medium from fibroblasts from benign prostatic hyperplasia
CM-CAF: conditioned medium from cancer associated fibroblast
DE: differentially expressed
DM: differentially methylated
FC: fold change

Declarations

- **Ethics approval and consent to participate**

The Ethics approval and consent to participate at the study named 'Ruolo del microambiente stromale nella immunomodulazione e nella progression del carcinoma prostatico (Role of the stromal microenvironment for immunomodulation and for cancer progression in prostatic carcinoma)' was released with reference number 'BIO 15.016' from the Ethics committee 'Area Vasta Centro' of the "Azienda Ospedaliera Universitaria Correggi" in Florence – Italy on the 26th of June 2015 (see attached document).

- **Availability of data and materials**

Raw microarray data have been deposited in the EBI ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) according to the Minimum Information About a Microarray Experiment (MIAME) guidelines with Accession Number E-MTAB-4753.

Raw RNA sequencing data have been deposited in the EBI ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) according to the Minimum Information about high-throughput SEQuencing Experiment (MINSEQE) guidelines with Accession Number E-MTAB-4752.

- **Conflict of interest**

"The authors declare that they have no competing interests"

- **Acknowledgements**

Work was supported by: Epigenomics Flagship Project (EPIGEN-CNR-IT: to IMB, VC, DF and AW); Italian Association for Cancer Research (AIRC: grants IG-8797 to PC and IG-17426 to AW). T.C. is now supported by a three year AIRC fellowship (Project code: 19548)

“Supplementary Information accompanies the paper on the Oncogene webs site (http://www.nature.com/onc)”

References

1 Aigner K, Dampier B, Descovich L, Mikula M, Sultan A, Schreiber M *et al*. The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. Oncogene 2007; 26: 6979-6988.

2 Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015; 31: 166-169.

3 Aue A, Hinze C, Walentin K, Ruffert J, Yurtdas Y, Werth M *et al*. A Grainyhead-Like 2/Ovo-Like 2 Pathway Regulates Renal Epithelial Barrier Function and Lumen Expansion. J Am Soc Nephrol 2015; 26: 2704-2715.

4 Babbio F, Castiglioni I, Cassina C, Gariboldi MB, Pistore C, Magnani E *et al*. Knock-down of methyl CpG-binding protein 2 (MeCP2) causes alterations in cell proliferation and nuclear lamins expression in mammalian cells. BMC Cell Biol 2012; 13: 19.

5 Babbio F, Pistore C, Curti L, Castiglioni I, Kunderfranco P, Brino L *et al*. The SRA protein UHRF1 promotes epigenetic crosstalks and is involved in prostate cancer progression. Oncogene 2012.

6 Bernhart SH, Kretzmer H, Holdt LM, Jühling F, Ammerpohl O, Bergmann AK *et al*. Changes of bivalent chromatin coincide with increased expression of developmental genes in cancer. Sci Rep 2016; 6: 37393.

7 Bhatt T, Rizvi A, Batta SP, Kataria S, Jamora C. Signaling and mechanical roles of E-cadherin. Cell Commun Adhes 2013; 20: 189-199.

8 Bostick M, Kim JK, Estève PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 2007; 317: 1760-1764.

9 Brown RL, Reinke LM, Damerow MS, Perez D, Chodosh LA, Yang J *et al*. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. J Clin Invest 2011; 121: 1064-1074.

10 Cardenas H, Vieth E, Lee J, Segar M, Liu Y, Nephew KP *et al*. TGF-β induces global changes in DNA methylation during the epithelial-to-mesenchymal transition in ovarian cancer cells. Epigenetics 2014; 9: 1461-1472.

11 Carmona FJ, Davalos V, Vidal E, Gomez A, Heyn H, Hashimoto Y *et al*. A comprehensive DNA methylation profile of epithelial-to-mesenchymal transition. Cancer Res 2014; 74: 5608-5619.

12 Chaffer CL, Marjanovic ND, Lee T, Bell G, Kleer CG, Reinhardt F *et al*. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell 2013; 154: 61-74.

13 Cheung AK, Ip JC, Chu AC, Cheng Y, Leong MM, Ko JM *et al*. PTPRG suppresses tumor growth and invasion via inhibition of Akt signaling in nasopharyngeal carcinoma. Oncotarget 2015; 6: 13434-13447.

14 Cieply B, Riley P, Pifer PM, Widmeyer J, Addison JB, Ivanov AV *et al*. Suppression of the epithelial-mesenchymal transition by Grainyhead-like-2. Cancer Res 2012; 72: 2440-2453.

15 Cieply B, Farris J, Denvir J, Ford HL, Frisch SM. Epithelial-mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. Cancer Res 2013; 73: 6299-6309.

16 Cirri P, Chiarugi P. Cancer-associated-fibroblasts and tumour cells: a diabolic liaison driving cancer progression. Cancer Metastasis Rev 2012; 31: 195-208.

17 Cui H, Wang L, Gong P, Zhao C, Zhang S, Zhang K *et al*. Deregulation between miR-29b/c and DNMT3A is associated with epigenetic silencing of the CDH1 gene, affecting cell migration and invasion in gastric cancer. PLoS One 2015; 10: e0123926.

18 Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev 2011; 25: 1010-1022.

19 Dumont N, Wilson MB, Crawford YG, Reynolds PA, Sigaroudinia M, Tlsty TD. Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers. Proc Natl Acad Sci U S A 2008; 105: 14867-14872.

20 Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E *et al*. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci U S A 2007; 104: 15805-15810.

21 Ficz G. New insights into mechanisms that regulate DNA methylation patterning. J Exp Biol 2015; 218: 14-20.

22 Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST *et al*. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature 2015; 527: 472-476.

23 García de Herreros A. Epithelial to mesenchymal transition in tumor cells as consequence of phenotypic instability. Front Cell Dev Biol 2014; 2: 71.

24 Gascard P, Tlsty TD. Carcinoma-associated fibroblasts: orchestrating the composition of malignancy. Genes Dev 2016; 30: 1002-1019.

25 Giannoni E, Bianchini F, Masieri L, Serni S, Torre E, Calorini L *et al*. Reciprocal activation of prostate cancer cells and cancer-associated fibroblasts stimulates epithelial-mesenchymal transition and cancer stemness. Cancer Res 2010; 70: 6945-6956.

26 Giannoni E, Bianchini F, Calorini L, Chiarugi P. Cancer associated fibroblasts exploit reactive oxygen species through a proinflammatory signature leading to epithelial mesenchymal transition and stemness. Antioxid Redox Signal 2011; 14: 2361-2371.

27 Giannoni E, Taddei ML, Parri M, Bianchini F, Santosuosso M, Grifantini R *et al*. EphA2-mediated mesenchymal-amoeboid transition induced by endothelial progenitor cells enhances metastatic spread due to cancer-associated fibroblasts. J Mol Med (Berl) 2013; 91: 103-115.

28 Grega-Larson NE, Crawley SW, Erwin AL, Tyska MJ. Cordon bleu promotes the assembly of brush border microvilli. Mol Biol Cell 2015; 26: 3803-3815.

29 Guo M, Liu W, Serra S, Asa SL, Ezzat S. FGFR2 isoforms support epithelial-stromal interactions in thyroid cancer progression. Cancer Res 2012; 72: 2017-2027.

30 Hollier BG, Evans K, Mani SA. The epithelial-to-mesenchymal transition and cancer stem cells: a coalition against cancer therapies. J Mammary Gland Biol Neoplasia 2009; 14: 29-43.

31 Hong T, Watanabe K, Ta CH, Villarreal-Ponce A, Nie Q, Dai X. An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States. PLoS Comput Biol 2015; 11: e1004569.

32 Hur K, Toiyama Y, Takahashi M, Balaguer F, Nagasaka T, Koike J *et al*. MicroRNA-200c modulates epithelial-to-mesenchymal transition (EMT) in human colorectal cancer metastasis. Gut 2013; 62: 1315-1326.

33 Ke XS, Qu Y, Rostad K, Li WC, Lin B, Halvorsen OJ *et al*. Genome-wide profiling of histone h3 lysine 4 and lysine 27 trimethylation reveals an epigenetic signature in prostate carcinogenesis. PLoS One 2009; 4: e4687.

34 Ke XS, Qu Y, Cheng Y, Li WC, Rotter V, Øyan AM *et al*. Global profiling of histone and DNA methylation reveals epigenetic-based regulation of gene expression during epithelial to mesenchymal transition in prostate cells. BMC Genomics 2010; 11: 669.

35 Kgatle MM, Kalla AA, Islam MM, Sathekge M, Moorad R. Prostate Cancer: Epigenetic Alterations, Risk Factors, and Therapy. Prostate Cancer 2016; 2016: 5653862.

36 Krisenko MO, Geahlen RL. Calling in SYK: SYK's dual role as a tumor promoter and tumor suppressor in cancer. Biochim Biophys Acta 2015; 1853: 254-263.

37 Lawlor H, Meunier A, McDermott N, Lynch TH, Marignol L. Identification of suitable endogenous controls for gene and miRNA expression studies in irradiated prostate cancer cells. Tumour Biol 2015; 36: 6019-6028.

38 Lee B, Villarreal-Ponce A, Fallahi M, Ovadia J, Sun P, Yu QC *et al*. Transcriptional mechanisms link epithelial plasticity to adhesion and differentiation of epidermal progenitor cells. Dev Cell 2014; 29: 47-58.

39 Li P, Yang R, Gao WQ. Contributions of epithelial-mesenchymal transition and cancer stem cells to the development of castration resistance of prostate cancer. Mol Cancer 2014; 13: 55.

40 Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E *et al*. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Mol Cell Biol 2002; 22: 480-491.

41 Liu F, Zhou Y, Zhou D, Kan M, Niu X, Zhang Z *et al*. Whole DNA methylome profiling in lung cancer cells before and after epithelial-to-mesenchymal transition. Diagn Pathol 2014; 9: 66.

42 Lu ZX, Huang Q, Park JW, Shen S, Lin L, Tokheim CJ *et al*. Transcriptome-wide landscape of pre-mRNA alternative splicing associated with metastatic colonization. Mol Cancer Res 2015; 13: 305-318.

43 Malouf GG, Taube JH, Lu Y, Roysarkar T, Panjarian S, Estecio MR *et al*. Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. Genome Biol 2013; 14: R144.

44 Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY *et al*. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008; 133: 704-715.

45 Marinkovich MP, Taylor TB, Keene DR, Burgeson RE, Zone JJ. LAD-1, the linear IgA bullous dermatosis autoantigen, is a novel 120-kDa anchoring filament protein synthesized by epidermal cells. J Invest Dermatol 1996; 106: 734-738.

46 McDonald OG, Wu H, Timp W, Doi A, Feinberg AP. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nat Struct Mol Biol 2011; 18: 867-874.

47 Morita S, Horii T, Kimura M, Ochiya T, Tajima S, Hatada I. miR-29 represses the activities of DNA methyltransferases and DNA demethylases. Int J Mol Sci 2013; 14: 14647-14658.

48 Nogales V, Reinhold WC, Varma S, Martinez-Cardus A, Moutinho C, Moran S *et al*. Epigenetic inactivation of the putative DNA/RNA helicase SLFN11 in human cancer confers resistance to platinum drugs. Oncotarget 2016; 7: 3084-3097.

49 Ohka F, Natsume A, Motomura K, Kishida Y, Kondo Y, Abe T *et al*. The global DNA methylation surrogate LINE-1 methylation is correlated with MGMT promoter methylation and is a better prognostic factor for glioma. PLoS One 2011; 6: e23332.

50 Parpart S, Roessler S, Dong F, Rao V, Takai A, Ji J *et al*. Modulation of miR-29 expression by α-fetoprotein is linked to the hepatocellular carcinoma epigenome. Hepatology 2014; 60: 872-883.

51 Porkka KP, Pfeiffer MJ, Waltering KK, Vessella RL, Tammela TL, Visakorpi T. MicroRNA expression profiling in prostate cancer. Cancer Res 2007; 67: 6130-6135.

52 Preca BT, Bajdak K, Mock K, Sundararajan V, Pfannstiel J, Maurer J *et al*. A self-enforcing CD44s/ZEB1 feedback loop maintains EMT and stemness properties in cancer cells. Int J Cancer 2015; 137: 2566-2577.

53 Roca H, Hernandez J, Weidner S, McEachin RC, Fuller D, Sud S *et al*. Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer. PLoS One 2013; 8: e76773.

54 Rodriguez J, Muñoz M, Vives L, Frangou CG, Groudine M, Peinado MA. Bivalent domains enforce transcriptional memory of DNA methylated genes in cancer cells. Proc Natl Acad Sci U S A 2008; 105: 19809-19814.

55 Ru P, Steele R, Newhall P, Phillips NJ, Toth K, Ray RB. miRNA-29b suppresses prostate cancer metastasis by regulating epithelial-mesenchymal transition signaling. Mol Cancer Ther 2012; 11: 1166-1173.

56 Sauer E, Babion I, Madea B, Courts C. An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensic organ tissue identification. Forensic Sci Int Genet 2014; 13: 217-223.

57 Scheel C, Weinberg RA. Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. Semin Cancer Biol 2012; 22: 396-403.

58 Schubert M, Spahn M, Kneitz S, Scholz CJ, Joniau S, Stroebel P *et al*. Distinct microRNA expression profile in prostate cancer patients with early clinical failure and the impact of let-7 as prognostic marker in high-risk prostate cancer. PLoS One 2013; 8: e65064.

59 Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA *et al*. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 2007; 450: 908-912.

60 Spans L, Helsen C, Clinckemalie L, Van den Broeck T, Prekovic S, Joniau S *et al*. Comparative genomic and transcriptomic analyses of LNCaP and C4-2B prostate cancer cell lines. PLoS One 2014; 9: e90002.

61 Stefanoli M, La Rosa S, Sahnane N, Romualdi C, Pastorino R, Marando A *et al*. Prognostic relevance of aberrant DNA methylation in g1 and g2 pancreatic neuroendocrine tumors. Neuroendocrinology 2014; 100: 26-34.

62 Steinestel K, Eder S, Schrader AJ, Steinestel J. Clinical significance of epithelial-mesenchymal transition. Clin Transl Med 2014; 3: 17.

63 Taddei ML, Parri M, Angelucci A, Bianchini F, Marconi C, Giannoni E *et al*. EphA2 induces metastatic growth regulating amoeboid motility and clonogenic potential in prostate carcinoma cells. Mol Cancer Res 2011; 9: 149-160.

64 Taddei ML, Cavallini L, Comito G, Giannoni E, Folini M, Marini A *et al*. Senescent stroma promotes prostate cancer progression: the role of miR-210. Mol Oncol 2014; 8: 1729-1746.

65 Tai S, Sun Y, Squires JM, Zhang H, Oh WK, Liang CZ *et al*. PC3 is a cell line characteristic of prostatic small cell carcinoma. Prostate 2011; 71: 1668-1679.

66 Takebe N, Miele L, Harris PJ, Jeong W, Bando H, Kahn M *et al*. Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: clinical update. Nat Rev Clin Oncol 2015; 12: 445-464.

67 Tam WL, Weinberg RA. The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat Med 2013; 19: 1438-1449.

68 Todaro M, Gaggianesi M, Catalano V, Benfante A, Iovino F, Biffoni M *et al*. CD44v6 is a marker of constitutive and reprogrammed cancer stem cells driving colon cancer metastasis. Cell Stem Cell 2014; 14: 342-356.

69 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR *et al*. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 2012; 7: 562-578.

70 Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. Curr Opin Cell Biol 2012; 24: 374-386.

71 von Meyenn F, Iurlaro M, Habibi E, Liu NQ, Salehzadeh-Yazdi A, Santos F *et al*. Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. Mol Cell 2016; 62: 983.

72 Warzecha CC, Sato TK, Nabet B, Hogenesch JB, Carstens RP. ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. Mol Cell 2009; 33: 591-601.

73 Warzecha CC, Shen S, Xing Y, Carstens RP. The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events. RNA Biol 2009; 6: 546-562.

74 Yang L, Chen Y, Cui T, Knösel T, Zhang Q, Albring KF *et al*. Desmoplakin acts as a tumor suppressor by inhibition of the Wnt/β-catenin signaling pathway in human lung cancer. Carcinogenesis 2012; 33: 1863-1870.

75 Yaqinuddin A, Qureshi SA, Qazi R, Abbas F. Down-regulation of DNMT3b in PC3 cells effects locus-specific DNA methylation, and represses cellular growth and migration. Cancer Cell Int 2008; 8: 13.

76 Yaqinuddin A, Qureshi SA, Qazi R, Farooq S, Abbas F. DNMT1 silencing affects locus specific DNA methylation and increases prostate cancer derived PC3 cell invasiveness. J Urol 2009; 182: 756-761.

77 Zhang P, Wei Y, Wang L, Debeb BG, Yuan Y, Zhang J *et al*. ATM-mediated stabilization of ZEB1 promotes DNA damage response and radioresistance through CHK1. Nat Cell Biol 2014; 16: 864-875.

78 Zhang P, Sun Y, Ma L. ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. Cell Cycle 2015; 14: 481-487.

79 Zhang Y, Yeh LK, Zhang S, Call M, Yuan Y, Yasunaga M *et al*. Wnt/β-catenin signaling modulates corneal epithelium stratification via inhibition of Bmp4 during mouse development. Development 2015; 142: 3383-3393.

80 Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H *et al*. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature 2015; 527: 525-530.

FIGURE LEGENDS

Figure 1. *Ex-vivo* model, qPCR and Western blot analysis of EMT markers and factors, and DNA methylation factors in LN-CaP and PC3 cells treated with the indicated conditioned media.

A) Schematic representation of the experimental procedure to induce EMT in both LN-CaP and PC3 cells, followed by the induction of the reverse process, MET only in PC3 cells. To induce EMT, LN-CaP and PC3 cells were treated for 72 hours with conditioned medium produced by human fibroblasts isolated from patients with benign prostatic tumour (CM-HPF) or prostate cancer with a scale of Gleason ≥ 4 (CM-CAF). 48h after seeding, fibroblasts were serum starved for 24 hours, and the medium than used to treat LN-CaP and PC3 cell lines. To induce MET, PC3 cells were further treated for 72h with serum free medium, called, respectively, CM-HPF+3 or CM-CAF+3. Conditioned medium from PC3 was prepared keeping the cells in culture with medium of starvation for 24 hours.

B) RT-PCR analysis. mRNA levels of LN-CaP, PC3 (EMT) and PC3 (MET) were analysed performing quantitative PCR with selective primers for EMT-markers (CDH1, CDH2, and VIM), EMT-factors (SNAI1, SNAI2, TWIST1 and ZEB1) and DNA Methylation Factors (UHRF1, DNMT1, DNMT3A and DNMT3B). In PC3 cells, CDH1, CDH2, VIM, ZEB1, UHRF1, DNMT1 and DNMT3B showed significant mRNA level variations during EMT process. Statistical analysis was performed on three independent biological experiments (unpaired Student's t-test; *: $p < 0.05$). GAPDH was used as a control housekeeping gene.

C) Western blot analysis and relative densitometry. EMT and MET processes were analysed using antibodies against specific markers: CDH1, CDH2 and VIM (EMT-markers), SNAIL1, SLUG, TWIST1 and ZEB1 (EMT-Factors), UHRF1, DNMT1, DNMT3A

and DNMT3B (DNA Methylation Factors). GAPDH was used as a control housekeeping gene. In PC3 cells, CDH1, CDH2, VIM, SNAI1, ZEB1, UHRF1, DNMT1 and DNMT3B showed a significant increase in protein levels in EMT process. Statistical analysis was performed on three independent biological experiments (unpaired Student's t-test; *: $p < 0.05$). The densitrometric analysis is the average of three independent experiments.

Figure 2. DNA methylation analysis in PC3 cells induced in EMT and MET.

A) Scatter plots of β-values and correlation coefficient (r^2) associated with variation in DNA methylation levels compared to controls ((a) CM-HPF vs CM-HPF+3, (b) CM-HPF vs CM-CAF+3), EMT ((c) CM-HPF vs CM-CAF) and MET ((d) CM-CAF vs CM-CAF+3) processes, and control of the reciprocal (e) CM-CAF+3 vs 1/CM-CAF.

B) Table of the CpG sites in which significant variations in DNA methylation levels are observed (pValue ≤0.01/Diff. Score ≥30 and Δβ ≥0.2 or Δβ ≥0.5) during EMT (CM-HPF_vs_CM-CAF), MET (CM-CAF_vs_CM-CAF+3) and controls (CM-HPF_vs_CM-HPF+3).

C) Heat-maps of DNA methylation levels on promoter and gene body regions relative to each of the three experimental biological replicates analysed by Illumina 450K arrays on PC3 cells treated with CM-HPF, CM-CAF and CM-CAF+3.

D) The circos plot represents the methylation levels on CpG sites (β-value) distributed along the entire length of the chromosomes in control (4th circos-plot in purple), EMT (2nd circos-plot in blue) and MET (1st circos-plot in orange) samples, whereas CpG hypo- (3rd circos-plot in green) and hyper-methylated (3rd circos-plot in red), observed during EMT process, are the result of the subtractions of β-values between CM-HPF and CM-CAF treated cells. Analysis was performed on three independent biological experiments.

Figure 3. Global transcriptional profiling during EMT and MET processes.

A) Flow chart, scatter plot and volcano plots of the RNA-seq analysis (p_adj≤0.05 and $\log_2$ Fold-Change ≥2; ≤-2) of genes differentially expressed (DE) during EMT (CM-HPF *vs* CM-CAF), MET (CM-CAF *vs* CM-CAF+3) and in controls (CM-HPF *vs* CM-HPF+3, CM-HPF *vs* CM-CAF+3). The results presented are the average of three independent biological experiments. The p_adj value was calculated by Students t-test method.

B) The common 1,924 DE genes (p_adj ≤ 0.05 and $\log_2$ Fold-Change ≥ 2; ≤ -2) in EMT and MET, up- (1,082 and 842) and down-regulated (842 and 1,082), respectively, are illustrated in the scatter plot (upper panel; coefficient of correlation r^2 = 0.986) and Venn diagram (lower panel).

C) Top ten scoring canonical pathways enriched by the common 1,924 DE genes obtained by Ingenuity Pathway Analysis (IPA). The stacked bar chart displays the percentage of up-regulated (red) and down-regulated (green) genes and those not-overlapping with our data-set (white) in each canonical pathway (the total number of genes is reported on the top of each bar). The secondary x-axis reports the statistical significance (log-P value) calculated by the Benjamini-Hochberg method.

D) Gene expression analysis ($\log_2$ Fold-Change) carried out on EMT Factors (ZEB1, TWIST1, SNAI1 and SNAI2), EMT Markers (VIM, CDH1, CDH2, CDH3, CRB3, EPPK1, PKP3, DSP, MPZL2, FGFR2 and HOXB13), EMT regulated genes (TGFB1I1, TGFBI, SERPINE and GREM1), EMT invasion/antagonists (GRHL2, ST14, PTPRG, OVOL1, OVOL2 and EHF), stem cell regulatory proteins (SHH, SMO, HHIP, GLI3, FZD4, DKK1, LGR5, HEY1 and JAG1), alternative splicing regulators (ESRP1 and MBNL3), TET proteins (TET1, TET2 and TET3) and differential expression of CD44 variants in control (CM-HPF), EMT (CM-CAF) and MET (CM-CAF+3) cells. The levels of the CD44 variants

are expressed in FPKM (Fragments Per Kilobase Million). For CD-44 variants see also Figure 8. The results presented are the average of three independent biological experiments.

Figure 4. Differentially expressed genes and DNA methylation changes analysis in EMT.

A) The flow chart shows the genes selection method employed to overlap RNA-Seq (ENSEMBLE) and Illumina methylation 450k array data. In detail, we recognized different groups of genes: Not-DE (-1.2< FC <1.2) and DE (log_2FC ≥2; ≤-2), respectively, which can be divided in not-differentially methylated (Not-DM) (-0.2 <Δβ< 0.2) and differentially methylated (DM) (Δβ ≤-0.2; ≥0.2). The distribution of DE-DM common genes (EMT vs MET, 1352) on the Promoter and/or Gene Body regions is illustrated in the Venn diagram; they are further split in up- and down-regulated. The results presented are the average of three independent biological experiments.

B) Scatter plots show the distribution of hyper-methylated (Δβ-value ≥0.2; lower triangle) and hypo-methylated (Δβ-value ≤-0.2; upper triangle) cytosine on 10151 Not-DE & DM **(i)** and on 1,443 DE & DM genes **(ii)** during EMT. Scatter- and Box-plots of the latter group are divided into up- (Blue) and down- (Red) regulated genes. The hyper- and hypo-methylated cytosines are divided into four gates: G1, G2, G3 and G4. Gate 1 (G1) includes cytosines experiencing strong hypo-methylation up to complete de-methylation, reaching β-values not exceeding 0.2 in CM-CAF treated cells (x axis), *i.e.* cytosines that were heavily methylated and became hypo-methylated or almost fully de-methylated during EMT; Gate 2 (G2) hypo-methylated cytosines that achieved β-values >0.2 in CM-CAF treated cells; Gate 3 (G3) hyper-methylated cytosines with β-values >0.2 in CM-HPF treated cells (y axis); Gate 4 (G4) cytosines that were almost unmethylated and that

became hyper-methylated during EMT, *i.e.* cytosines experiencing *de novo* methylation. The *P* Value is calculated by Wilcoxon-Mann-Whitney test.

C) Scatter-plots of the methylation status of CpGs associated with island regions in Not-DE & DM and DE & DM genes using the same criteria as in **B**.

Figure 5. CpG islands (CGIs) promoter hyper-methylation induces gene repression in the EMT program.

A) Box plots of DM CGIs in up- and down-regulated genes (DE & DM) located in the promoter and gene body regions. The *P* Value is calculated by Wilcoxon-Mann-Whitney test.

B and **C)** Scatter plots of CGIs in up- **(B)** and down-regulated **(C)** genes (DE & DM) located in the promoter and gene body regions, showing the distribution of hyper-methylated ($\Delta\beta$-value ≥0.2; Lower triangle composed by G3+G4 gate) and hypo-methylated ($\Delta\beta$-value ≤-0.2; upper triangle composed by G1+G2 gate) cytosines on 1,352 DE & DM genes during EMT. The histograms represent the CpG island percentage clustering Gene Body and/or Promoter regions for hyper-methylated ($\Delta\beta$-value ≥0.2; G3 and G4 gate) and hypo-methylated ($\Delta\beta$-value ≤-0.2; G1 and G2 gate) genes.

D) Scatter plots of DE *vs* DM in promoter and gene body CGIs that combine $\Delta\beta$ variation of DNA methylation and mRNA levels, during EMT and MET processes.

Figure 6. Analysis of the TCGA dataset (PRAD) and correlation with the *ex-vivo* PC3 DE-DM derived signature.

A) Heat-map of the unsupervised clustering analysis of TCGA dataset. The unsupervised clustering analysis of TCGA RNA sequencing data from 259 TCGA prostate cancer

samples (tumour cell content: ≥70%) separated the *ex-vivo* PC3 DE-DM signature into two clusters (1 and 2) based on differential gene expression. The significant variant genes (row n=488) that discriminate the two identified clusters are plotted in the heat-map. PC3 UP and DOWN regulated genes of the *ex-vivo* DE-DM signature are indicated on the right. The 294 genes whose expression levels changes according to the PC3 signature, are boxed in the heat-map.

B) Box plots for the comparison of the expression levels between cluster-1 and 2 of EMT markers/regulators and of stem cell regulators. The box-plots show the RNA levels (z-score) of EMT Markers (*VIM, CDH1, CDH2, CRB3, DSP, PKP3*), EMT regulators (*ZEB1, GRHL2, OVOL1, ST14*), stem cell regulatory proteins (*LGR5, GLI3, DKK1, HEY1 and HHIP*), alternative splicing regulators (*ESRP1*) in cluster-1 (Brown) and cluster-2 (Green-EMT like). The *P*-value is calculated by Wilcoxon-Mann-Whitney test (*p<0.05).

C) Analysis of the expression levels of the androgen receptor and of CAF markers in cluster-1 and 2. Upper table. PRAD samples distribution (relative percentage) according to the AR or CAF markers levels (up,+, or down,-, calculated with respect to the median) in Cluster-1 or 2, respectively. Intermediate and lower box-plots. RNA levels (Z-score) of AR (androgen receptor) and CAF markers (*ASPN, COL11A1, MFAP5*) in cluster-1 or 2. The *P*-value is calculated by Wilcoxon-Mann-Whitney test (*p<0.05).

D) Analysis of the correlation between RNA expression and DNA methylation levels of the EMT markers/regulators and of stem cells regulators in the PRAD dataset. Left: correlation between DNA methylation and RNA level expression. Colour code: from inverse correlation (-1, high RNA expression and low DNA methylation and *vice-versa*, blue) to direct correlation (+1, low RNA expression and low DNA methylation and *vice-versa*, red) for all the cytosines annotated on the Illumina Infinium HumanMethylation 450K BeadChip array of the indicated genes. Right: Box and scatter plots for the comparison of the DNA methylation levels between cluster-1 and 2 of the indicated cytosines (bleu arrows in the

drawings) along three sample genes (VIM - EMT marker, GRHL2 - EMT regulator and GLI3 - stemness controller). The numbers in the plots indicate the specific cytosines reported in the overlying drawings.

Figure 7. Role of DNMT3A in CM-CAF induction of EMT.

A) Cartoons of CDH1, GRHL2 and β-ACTIN promoter and related promoter regions analysed in ChIP experiments.

B) ChIP experiments performed after immunoprecipitation with antibodies against DNA methyltransferases (DNMT1, -3A and -3B) and the indicated histone modifications (transcriptional repression H3K9me3, H3K27me3 and transcriptional activation H3K4me3, H3K36me3). Quantitative ChIP analysis of all antibodies is the results of three technical replicates of three independent experiments (*****: $p<0.05$; ******: $p<0.01$; and *******: $p<0.001$ unpaired Student's t-test).

C) Western Blot analysis performed on PC3 cells silenced for DNMT3A 24h before the EMT induction with conditioned or control mediums, using antibodies against CDH1, DNMT3, GRHL2, ZEB1 and VIM proteins (β-ACTIN as control housekeeping gene). The levels of the proteins are calculated by densitometric analysis on the displayed bands (left image).

D) TaqMan MicroRNA Reverse Transcription for the miR-29a, b and c on the RNA extracted form PC3 cells treated with CM-HPF and CM-CAF. snoRNA U48 was used for relative miRNA quantification. All the samples are the result of three technical replicates of two independent experiments (*****: $p<0.01$; **º**: $p<0.05$, unpaired Student's t-test).

Figure 8. Schematic model of the molecular pathway implicated in EMT processes induced by CM-CAF in PC3 cells.

In PC3 cells treated with CM-HPF (control, left panel), epithelial factors such as OVOL1, OVOL2 and GRHL2 are expressed and inhibit the transcription of ZEB1 gene; CDH1 gene is transcribed and ESRP1 regulate the splicing of CD44 isoforms.

The conditioning of PC3 cells with CM-CAF (right panel) is able to induce EMT process through DNA methylation variations and down-regulation of GRHL2, CDH1 and ESRP1 (and consequently isoform-switching of CD44); conversely, DNA hypo-methylation determines activation of ZEB1 and stemness genes.

TABLE 1. Association between PC3 EMT–signature clusters and patients' clinico-pathological features in TCGA PRAD dataset.

FIGURE 1

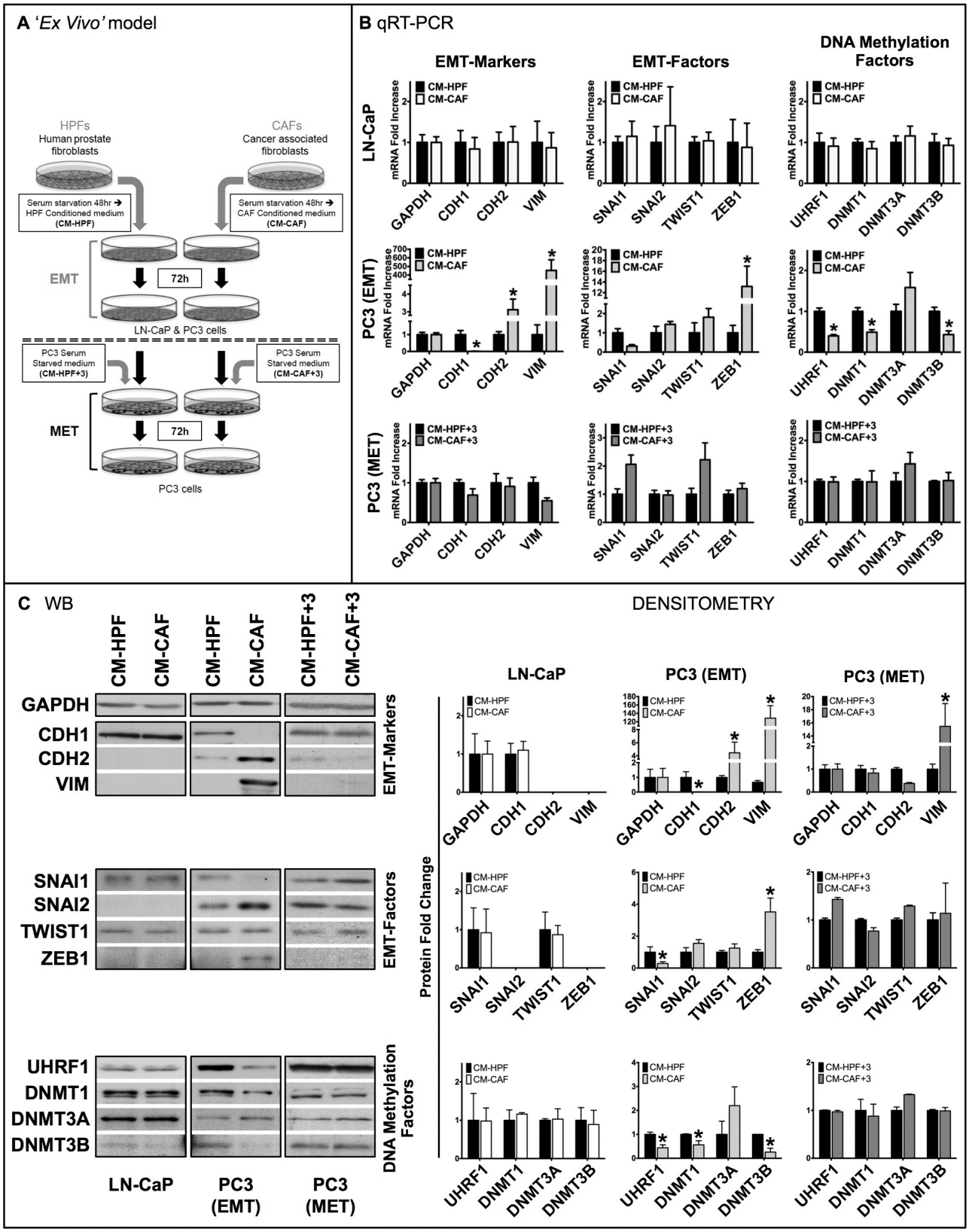
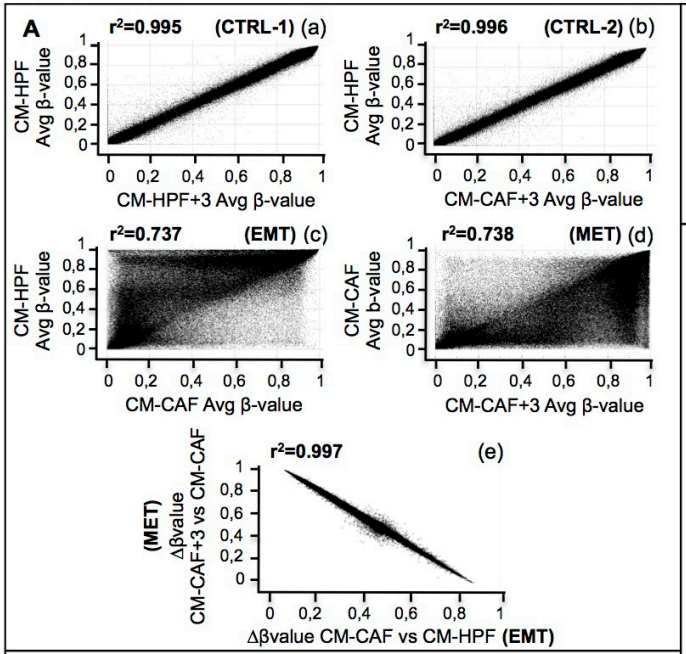
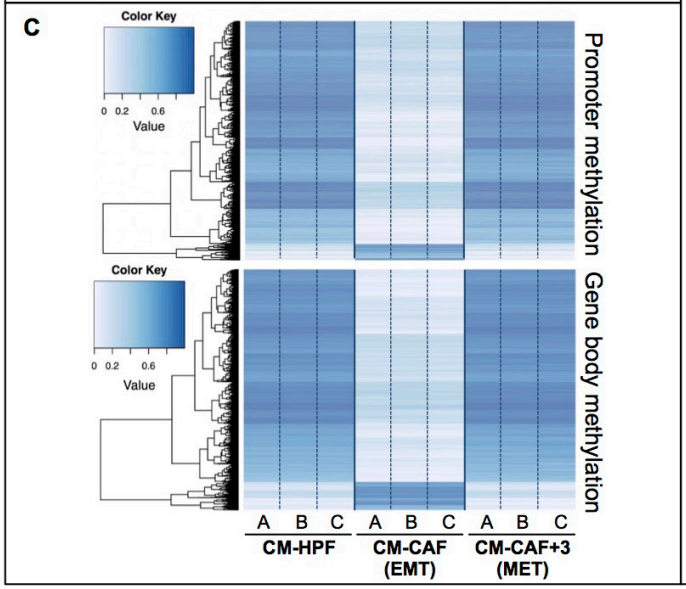

A 'Ex Vivo' model
HPFs
Human prostate fibroblasts
CAFs
Cancer associated fibroblasts
Serum starvation 48hr → HPF Conditioned medium (CM-HPF)
Serum starvation 48hr → CAF Conditioned medium (CM-CAF)
EMT
72h
LN-CaP & PC3 cells
PC3 Serum Starved medium (CM-HPF+3)
PC3 Serum Starved medium (CM-CAF+3)
MET
72h
PC3 cells
B qRT-PCR
EMT-Markers
EMT-Factors
DNA Methylation Factors
LN-CaP
PC3 (EMT)
PC3 (MET)
mRNA Fold Increase
CM-HPF
CM-CAF
CM-HPF+3
CM-CAF+3
GAPDH
CDH1
CDH2
VIM
SNAI1
SNAI2
TWIST1
ZEB1
UHRF1
DNMT1
DNMT3A
DNMT3B
C WB
DENSITOMETRY
Protein Fold Change
LN-CaP
PC3 (EMT)
PC3 (MET)

FIGURE 2





B

	Detection pValue≤0,01 Diff. Score ≥30	Δβ≥0,2 Hypo	Δβ≥0,2 Hyper	Δβ≥0,5 Hypo	Δβ≥0,5 Hyper
CM-HPF_vs_CM-CAF+3	15	4*	1	0	0
CM-CAF_vs_CM-HPF (EMT)	159010	96534	11148	21852	1963
CM-CAF+3_vs_CM-HPF (MET)	22	8**	3	0	0

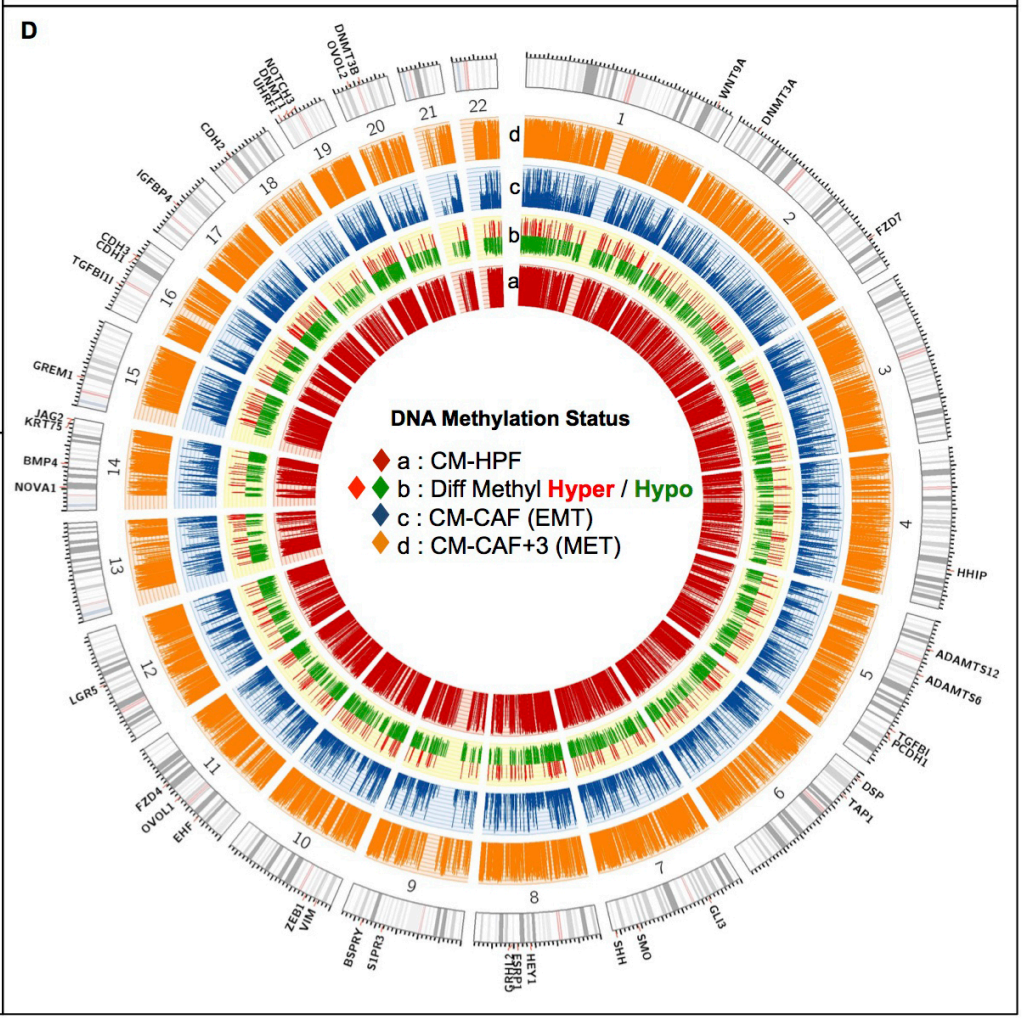


FIGURE 3

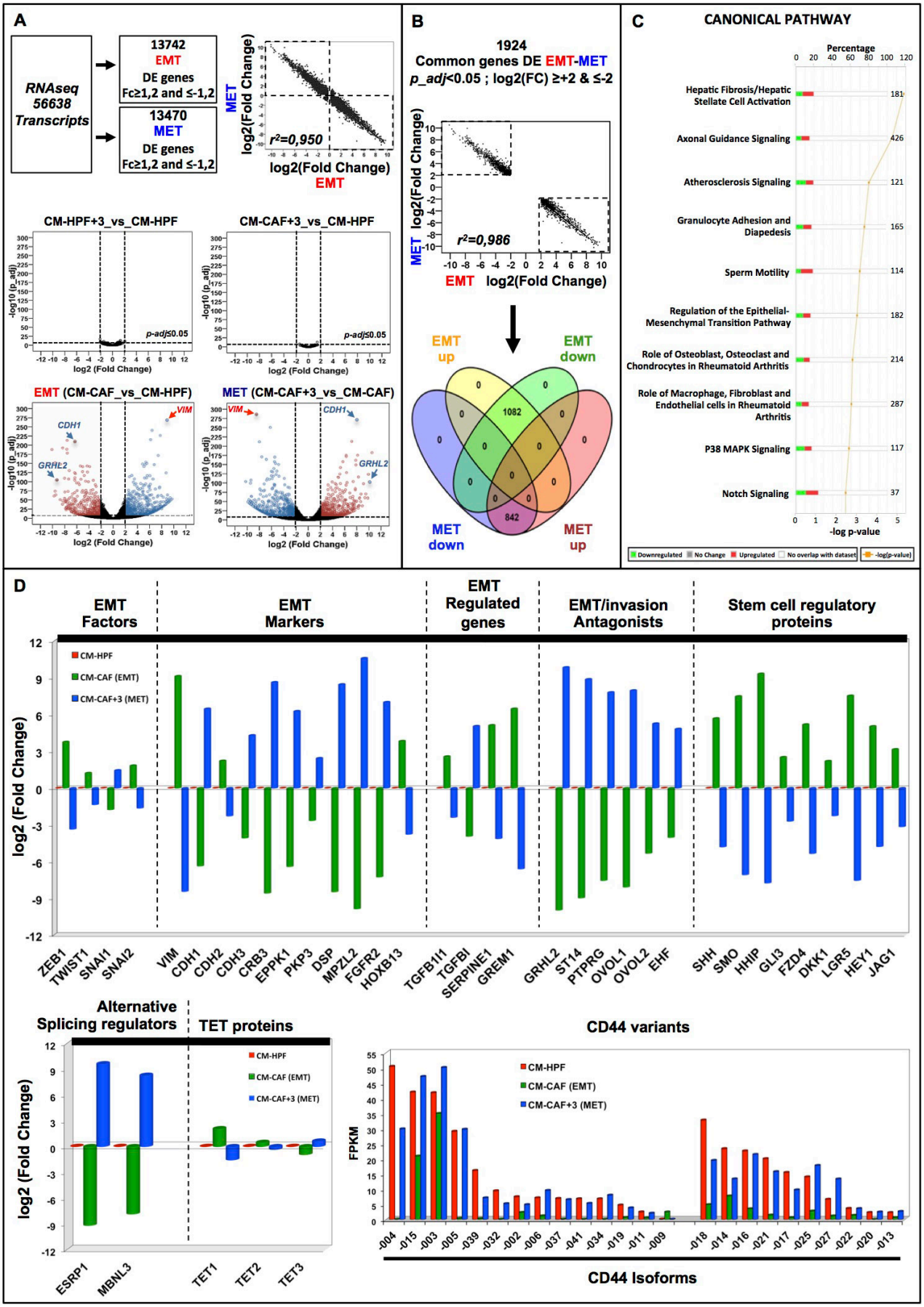

A
RNAseq 56638 Transcripts
13742 EMT DE genes Fc≥1,2 and ≤-1,2
13470 MET DE genes Fc≥1,2 and ≤-1,2
MET log2(Fold Change)
r²=0,950
log2(Fold Change) EMT
CM-HPF+3_vs_CM-HPF
CM-CAF+3_vs_CM-HPF
-log10 (p_adj)
p-adj≤0.05
log2 (Fold Change)
EMT (CM-CAF_vs_CM-HPF)
MET (CM-CAF+3_vs_CM-CAF)
CDH1
VIM
GRHL2
B
1924
Common genes DE EMT-MET
p_adj<0.05 ; log2(FC) ≥+2 & ≤-2
MET log2(Fold Change)
r²=0,986
EMT log2(Fold Change)
EMT up
EMT down
MET down
MET up
1082
842
C
CANONICAL PATHWAY
Percentage
Hepatic Fibrosis/Hepatic Stellate Cell Activation
Axonal Guidance Signaling
Atherosclerosis Signaling
Granulocyte Adhesion and Diapedesis
Sperm Motility
Regulation of the Epithelial-Mesenchymal Transition Pathway
Role of Osteoblast, Osteoclast and Chondrocytes in Rheumatoid Arthritis
Role of Macrophage, Fibroblast and Endothelial cells in Rheumatoid Arthritis
P38 MAPK Signaling
Notch Signaling
181
426
121
165
114
182
214
287
117
37
-log p-value
Downregulated
No Change
Upregulated
No overlap with dataset
-log(p-value)
D
EMT Factors
EMT Markers
EMT Regulated genes
EMT/invasion Antagonists
Stem cell regulatory proteins
CM-HPF
CM-CAF (EMT)
CM-CAF+3 (MET)
log2 (Fold Change)
ZEB1 TWIST1 SNAI1 SNAI2
VIM CDH1 CDH2 CDH3 CRB3 EPPK1 PKP3 DSP MPZL2 FGFR2 HOXB13
TGFB1I1 TGFBI SERPINE1 GREM1
GRHL2 ST14 PTPRG OVOL1 OVOL2 EHF
SHH SMO HHIP GLI3 FZD4 DKK1 LGR5 HEY1 JAG1
Alternative Splicing regulators
TET proteins
ESRP1 MBNL3
TET1 TET2 TET3
CD44 variants
FPKM
-004 -015 -003 -005 -039 -032 -002 -006 -037 -041 -034 -019 -011 -009
-018 -014 -016 -021 -017 -025 -027 -022 -020 -013
CD44 Isoforms

FIGURE 4

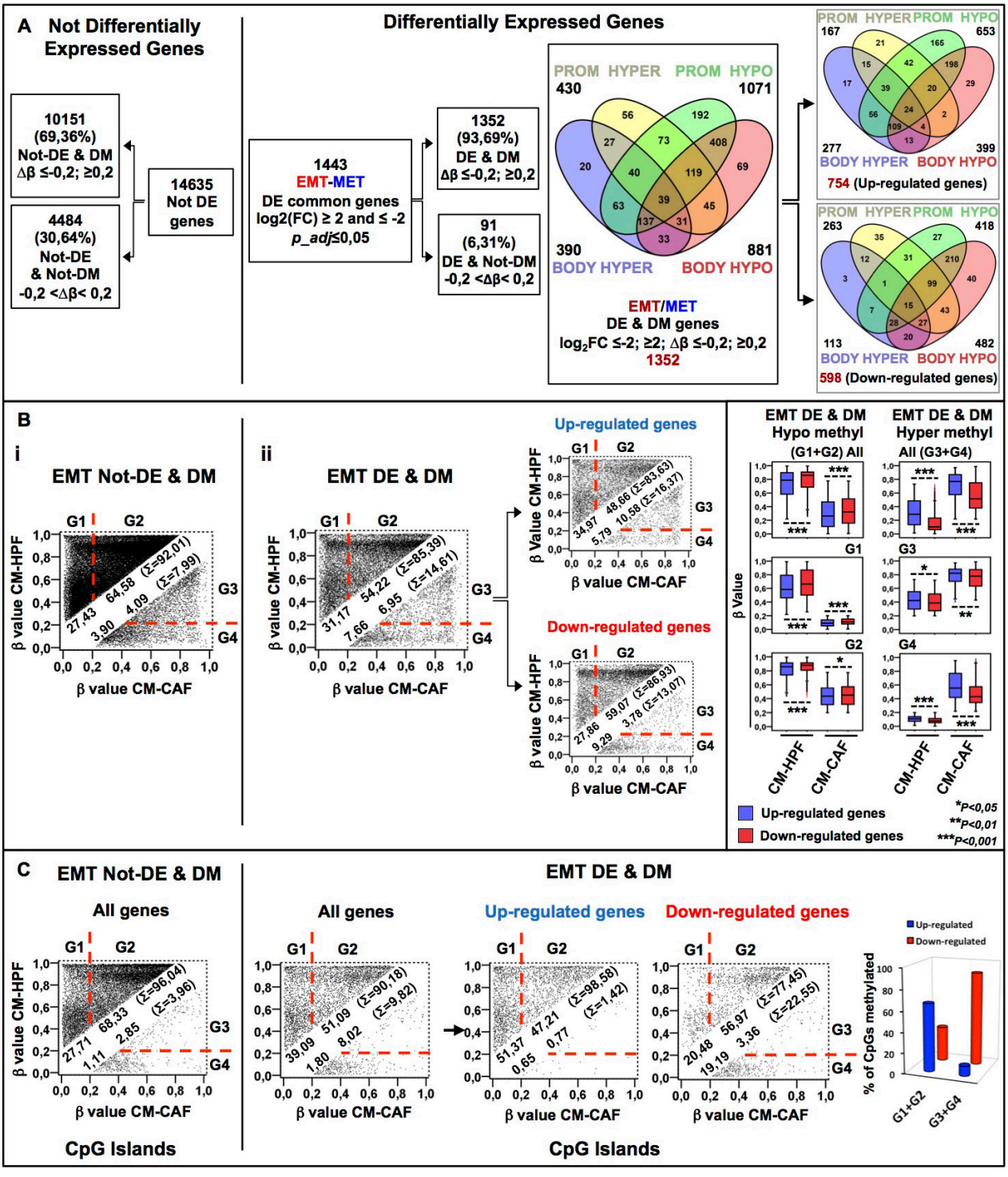

A Not Differentially Expressed Genes
10151 (69,36%) Not-DE & DM Δβ ≤-0,2; ≥0,2
4484 (30,64%) Not-DE & Not-DM -0,2 <Δβ< 0,2
14635 Not DE genes
Differentially Expressed Genes
1443 EMT-MET DE common genes log2(FC) ≥ 2 and ≤ -2 p_adj≤0,05
1352 (93,69%) DE & DM Δβ ≤-0,2; ≥0,2
91 (6,31%) DE & Not-DM -0,2 <Δβ< 0,2
PROM HYPER 430
PROM HYPO 1071
BODY HYPER 390
BODY HYPO 881
EMT/MET DE & DM genes log2FC ≤-2; ≥2; Δβ ≤-0,2; ≥0,2 1352
PROM HYPER 167
PROM HYPO 653
BODY HYPER 277
BODY HYPO 399
754 (Up-regulated genes)
PROM HYPER 263
PROM HYPO 418
BODY HYPER 113
BODY HYPO 482
598 (Down-regulated genes)
B
i
EMT Not-DE & DM
ii
EMT DE & DM
Up-regulated genes
Down-regulated genes
β value CM-HPF
β value CM-CAF
G1
G2
G3
G4
EMT DE & DM Hypo methyl (G1+G2) All
EMT DE & DM Hyper methyl All (G3+G4)
β Value
CM-HPF
CM-CAF
Up-regulated genes
Down-regulated genes
*P<0,05
**P<0,01
***P<0,001
C
EMT Not-DE & DM
EMT DE & DM
All genes
Up-regulated genes
Down-regulated genes
CpG Islands
% of CpGs methylated
G1+G2
G3+G4

FIGURE 5

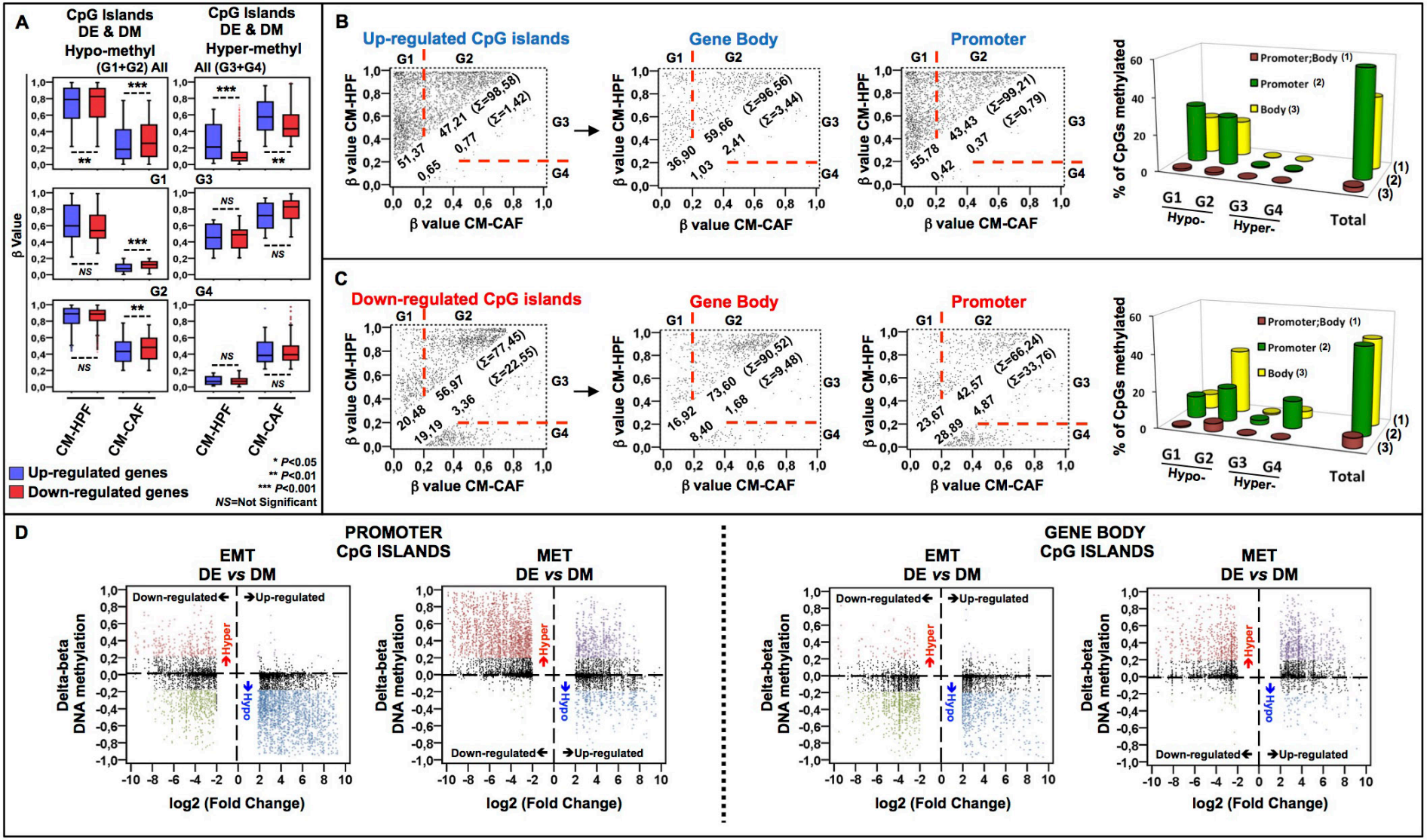

A
CpG Islands DE & DM Hypo-methyl (G1+G2) All
CpG Islands DE & DM Hyper-methyl All (G3+G4)
β Value
G1
G2
G3
G4
CM-HPF
CM-CAF
Up-regulated genes
Down-regulated genes
* P<0.05
** P<0.01
*** P<0.001
NS=Not Significant
B
Up-regulated CpG islands
Gene Body
Promoter
β value CM-HPF
β value CM-CAF
% of CpGs methylated
Promoter;Body (1)
Promoter (2)
Body (3)
Hypo-
Hyper-
Total
C
Down-regulated CpG islands
D
PROMOTER CpG ISLANDS
GENE BODY CpG ISLANDS
EMT DE vs DM
MET DE vs DM
Down-regulated
Up-regulated
Hyper
Hypo
Delta-beta DNA methylation
log2 (Fold Change)

Figure 6

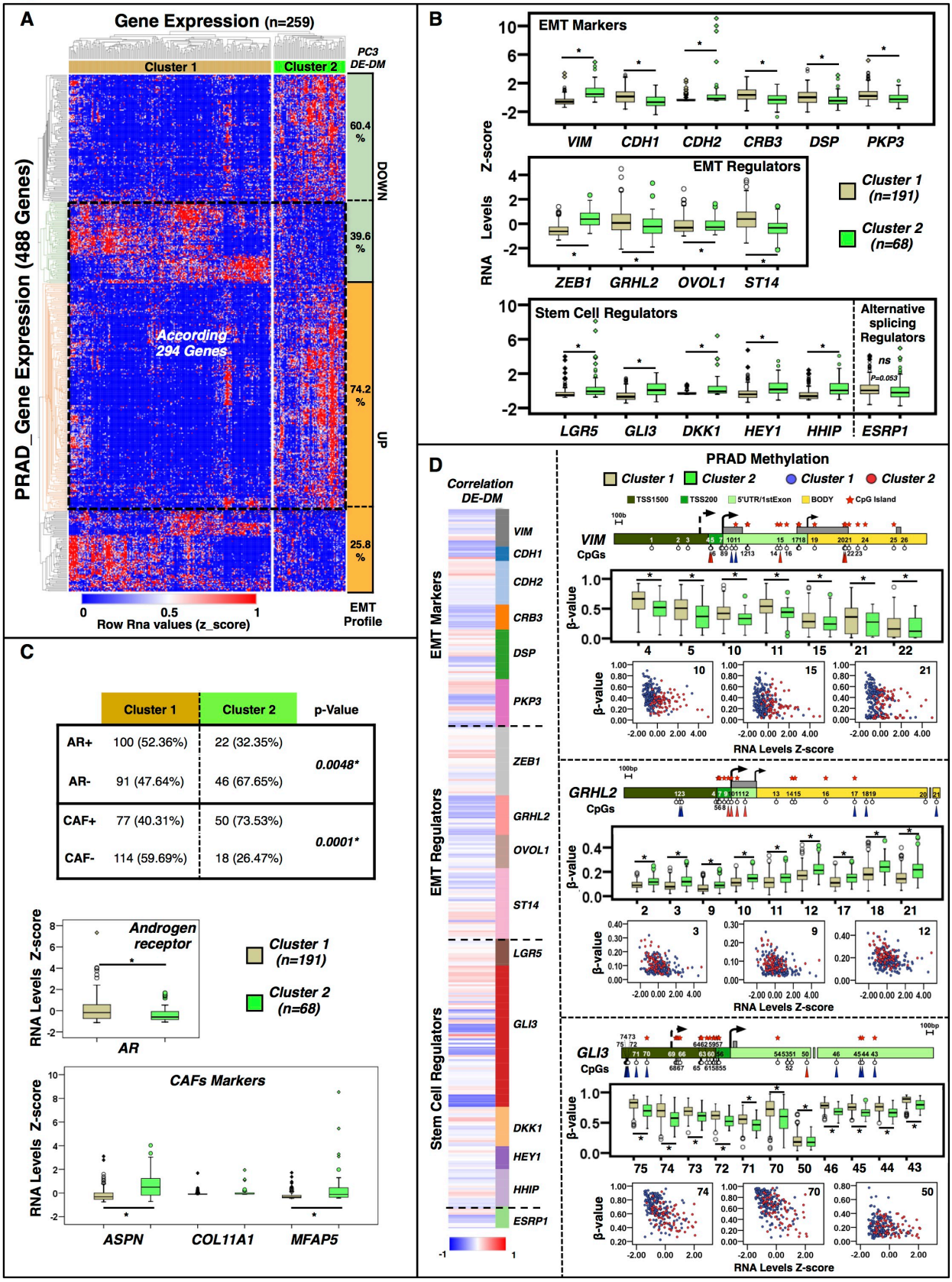

A
Gene Expression (n=259)
PC3 DE-DM
Cluster 1
Cluster 2
PRAD_Gene Expression (488 Genes)
According 294 Genes
60.4 %
39.6 %
74.2 %
25.8 %
DOWN
UP
0
0.5
1
Row Rna values (z_score)
EMT Profile
B
EMT Markers
VIM CDH1 CDH2 CRB3 DSP PKP3
EMT Regulators
ZEB1 GRHL2 OVOL1 ST14
Z-score
RNA Levels
Cluster 1 (n=191)
Cluster 2 (n=68)
Stem Cell Regulators
LGR5 GLI3 DKK1 HEY1 HHIP
Alternative splicing Regulators
ESRP1
ns
P=0.053
C
Cluster 1
Cluster 2
p-Value
AR+ 100 (52.36%) 22 (32.35%) 0.0048*
AR- 91 (47.64%) 46 (67.65%)
CAF+ 77 (40.31%) 50 (73.53%) 0.0001*
CAF- 114 (59.69%) 18 (26.47%)
Androgen receptor
RNA Levels Z-score
AR
Cluster 1 (n=191)
Cluster 2 (n=68)
CAFs Markers
ASPN COL11A1 MFAP5
D
Correlation DE-DM
EMT Markers
EMT Regulators
Stem Cell Regulators
VIM CDH1 CDH2 CRB3 DSP PKP3 ZEB1 GRHL2 OVOL1 ST14 LGR5 GLI3 DKK1 HEY1 HHIP ESRP1
PRAD Methylation
Cluster 1 Cluster 2 Cluster 1 Cluster 2
TSS1500 TSS200 5'UTR/1stExon BODY CpG Island
VIM CpGs
GRHL2 CpGs
GLI3 CpGs
β-value
RNA Levels Z-score

FIGURE 7

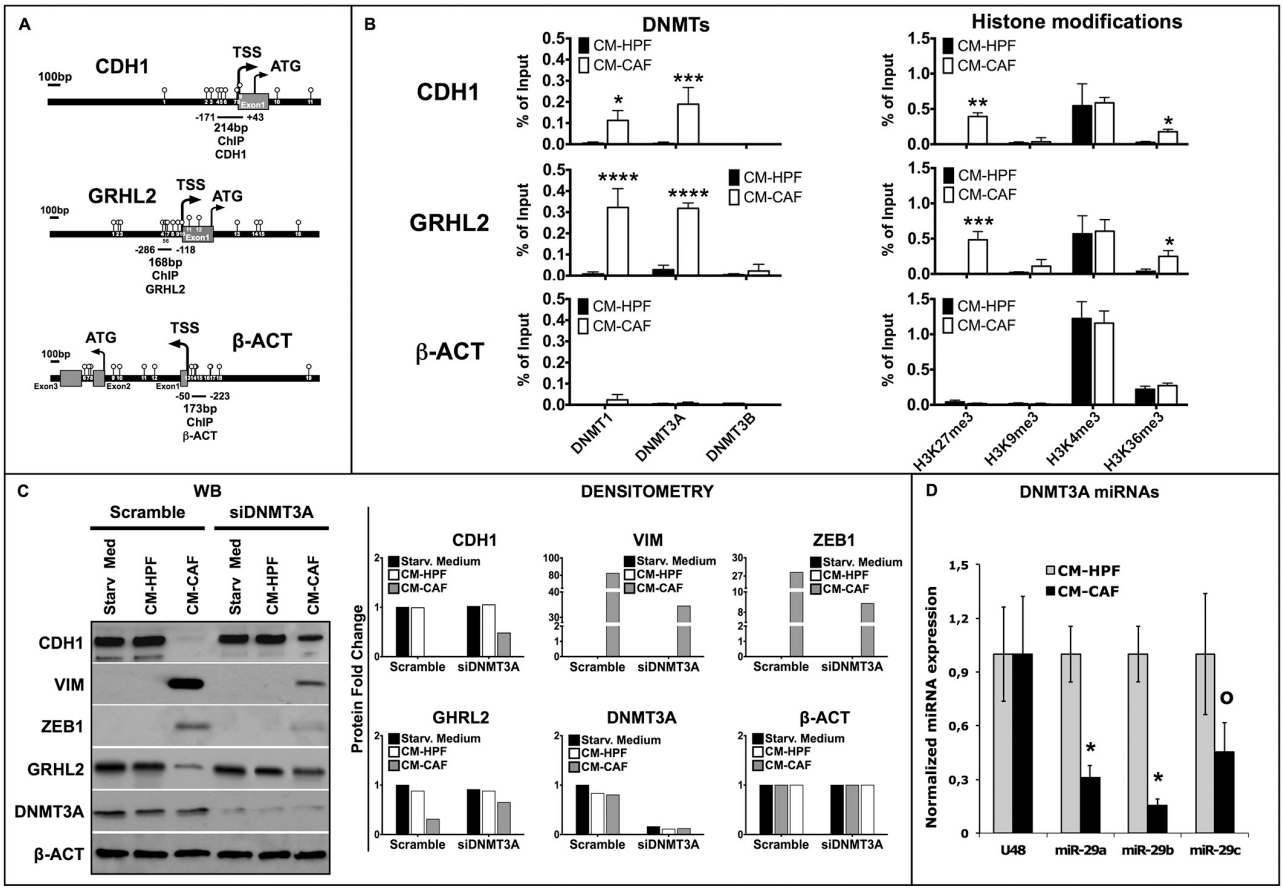
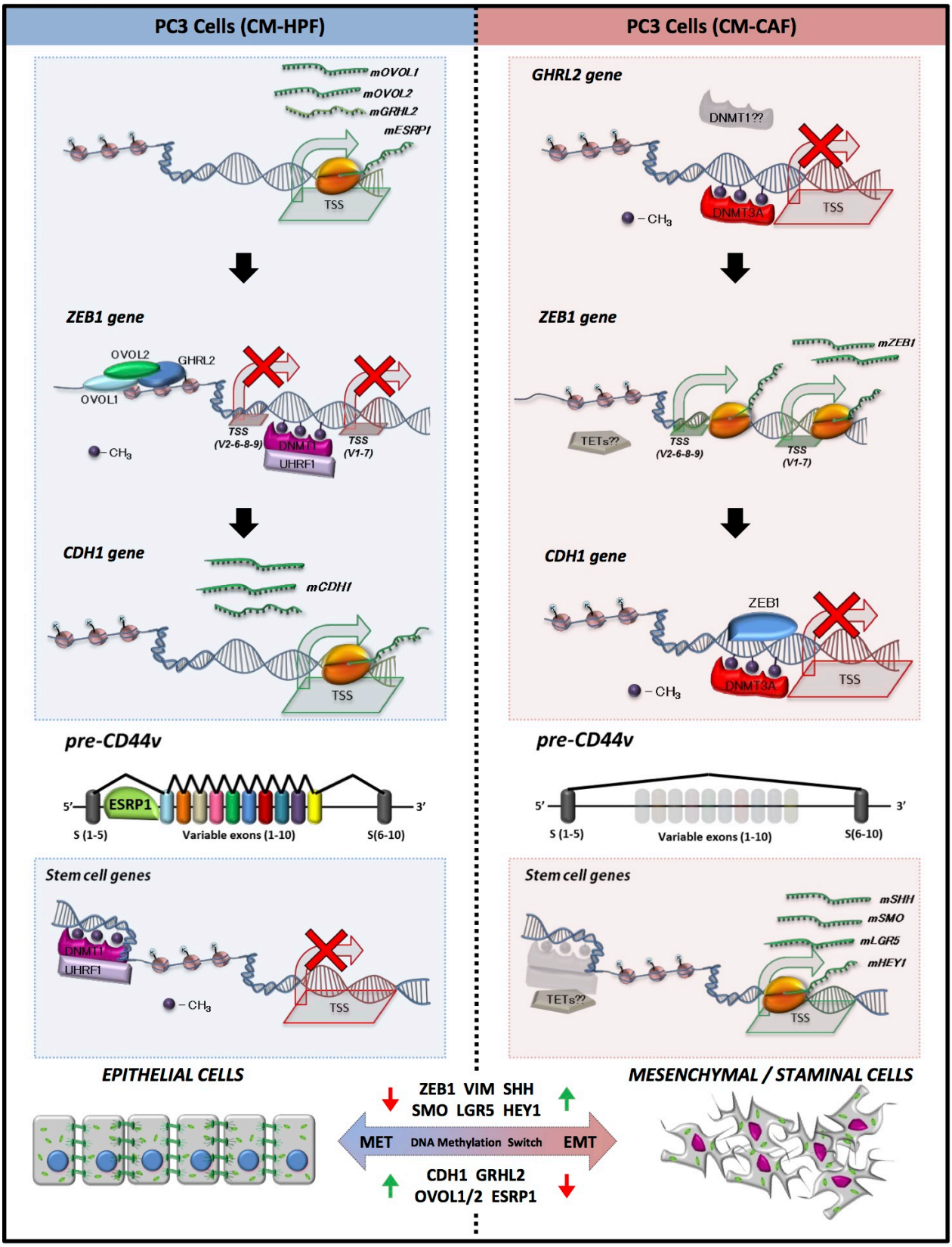

A
CDH1
TSS
ATG
100bp
-171
+43
214bp
ChIP
CDH1
GRHL2
-286
-118
168bp
ChIP
GRHL2
β-ACT
-50
-223
173bp
ChIP
β-ACT
B
DNMTs
Histone modifications
CM-HPF
CM-CAF
% of Input
CDH1
GRHL2
β-ACT
DNMT1
DNMT3A
DNMT3B
H3K27me3
H3K9me3
H3K4me3
H3K36me3
C
WB
Scramble
siDNMT3A
Starv Med
CM-HPF
CM-CAF
CDH1
VIM
ZEB1
GRHL2
DNMT3A
β-ACT
DENSITOMETRY
Protein Fold Change
Starv. Medium
Scramble
siDNMT3A
D
DNMT3A miRNAs
Normalized miRNA expression
U48
miR-29a
miR-29b
miR-29c

FIGURE 8


PC3 Cells (CM-HPF)
PC3 Cells (CM-CAF)
mOVOL1
mOVOL2
mGRHL2
mESRP1
TSS
GHRL2 gene
DNMT1??
DNMT3A
TSS
– CH3
ZEB1 gene
OVOL2
GHRL2
OVOL1
TSS (V2-6-8-9)
DNMT1
UHRF1
TSS (V1-7)
mZEB1
TETs??
CDH1 gene
mCDH1
ZEB1
pre-CD44v
5′
ESRP1
3′
S (1-5)
Variable exons (1-10)
S(6-10)
Stem cell genes
mSHH
mSMO
mLGR5
mHEY1
EPITHELIAL CELLS
MESENCHYMAL / STAMINAL CELLS
ZEB1 VIM SHH
SMO LGR5 HEY1
MET
DNA Methylation Switch
EMT
CDH1 GRHL2
OVOL1/2 ESRP1

VARIABILE		CLUSTER 1 n (%)	CLUSTER 2 n (%)	*p*-VALUE
Pathologic T	**T2a**	6 (3.17%)	1 (1.49%)	
NA=3	**T2b**	3 (1.59%)	2 (2.99%)	
	T2c	74 (39.15%)	12 (17.91%)	***0.00062****
	T3a	65 (34.39%)	19 (28.36%)	
	T3b	37 (19.58%)	31(46.27%)	
	T4	4 (2.12%)	2 (2.99%)	
Lymph nodes Metastasis	**N0**	130 (83.33%)	42 (71.19%)	***0.047****
NA=44	**N1**	26 (16.67%)	17 (28.81%)	
Gleason score	***6***	18 (9.42%)	2 (2.94%)	
	7	108 (56.54%)	23 (33.82%)	
	8	21 (10.99%)	9 (13.24%)	***0.00018****
	9	42 (21.99%)	34 (50.00%)	
	10	2 (1.05%)	0 (0%)	
Follow-up treatment success	**CRR**	109 (79%)	36 (76.60%)	
NA=74	**PRR**	5 (3.62%)	1 (2.13%)	*0.122*
	SD	12 (8.70%)	9 (19.15%)	
	PD	12 (8.70%)	1 (2.13%)	

CRR=complete remission/ response; *PRR*=partial remission/ response; *SD*=partial remission/ response; *PD*= progressive disease. NA=not analysed