

**AUTHOR QUERY FORM****Journal:** CELREP**Article Number:** 4257

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof. It is crucial that you NOT make direct edits to the PDF using the editing tools. Rather, please request corrections by using the tools in the Comment pane to annotate the PDF and call out the changes you would like to see. For further information on how to do this, please watch this short video: <https://authorcenter.dartmouthjournals.com/Article/PdfAnnotation>.

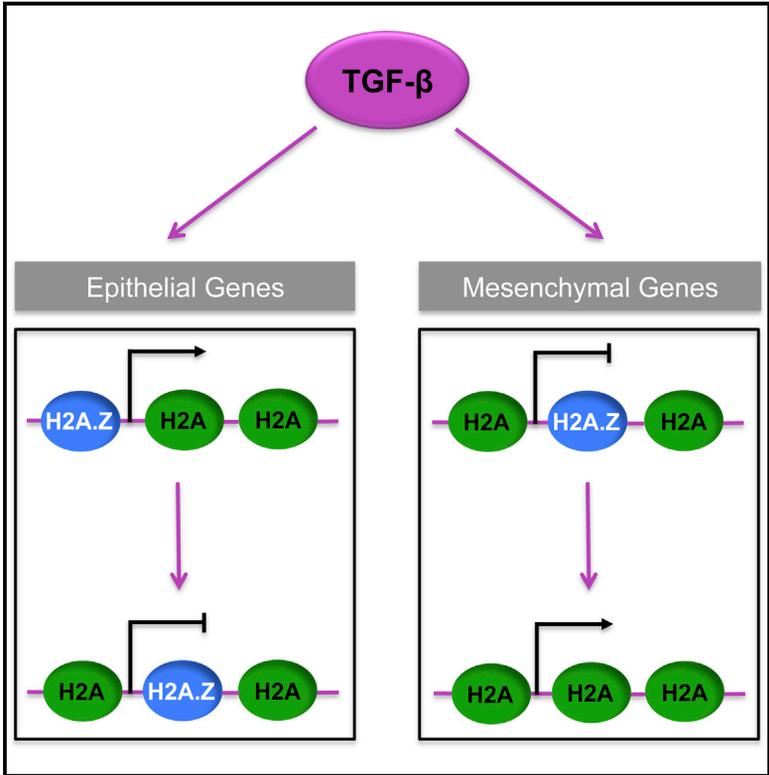
<b>Location in article</b>	<b>Query / Remark: Click on the Q link to find the query's location in text Please insert your reply or correction at the corresponding line in the proof</b>
<b>Q1</b>	Could you please provide the grant number for National Health and Medical Research Council, if any?

Thank you for your assistance.

Grant number is  
shown 1058941

## The Histone Variant H2A.Z Is a Master Regulator of the Epithelial-Mesenchymal Transition

### Graphical Abstract



### Authors

Renae Domaschenz,  
Sebastian Kurscheid, Maxim Nekrasov,  
Shuyi Han, David J. Tremethick

### Correspondence

david.tremethick@anu.edu.au

### In Brief

EMT is one of the most intensely studied differentiation-dedifferentiation processes. Domaschenz et al. now demonstrate that H2A.Z has the unique ability to simultaneously serve as either an activator or a repressor of epithelial or mesenchymal gene expression, respectively.

### Highlights

- H2A.Z acts as a central coordinator of TGF-β-induced EMT
- TGF-β induces the loss of H2A.Z from both epithelial and mesenchymal promoters
- H2A.Z loss up- or downstream of TSS represses epithelial or activates mesenchymal genes
- A role for H2A.Z in EMT can explain why H2A.Z is essential for early development

### Data and Software Availability

E-MTAB-5628  
E-MTAB-5637

# The Histone Variant H2A.Z Is a Master Regulator of the Epithelial-Mesenchymal Transition

Renae Domaschenz,<sup>1,2</sup> Sebastian Kurscheid,<sup>1,2</sup> Maxim Nekrasov,<sup>1,2</sup> Shuyi Han,<sup>1,2</sup> and David J. Tremethick<sup>1,3,\*</sup>

<sup>1</sup>The John Curtin School of Medical Research, The Australian National University, Canberra, ACT 2601, Australia

<sup>2</sup>These authors contributed equally

<sup>3</sup>Lead Contact

\*Correspondence: [david.tremethick@anu.edu.au](mailto:david.tremethick@anu.edu.au)  
<https://doi.org/10.1016/j.celrep.2017.09.086>

## SUMMARY

Epithelial-mesenchymal transition (EMT) is a profound example of cell plasticity that is crucial for embryonic development and cancer. Although it has long been suspected that chromatin-based mechanisms play a role in this process, no master regulator that can specifically regulate EMT has been identified to date. Here, we show that H2A.Z can coordinate EMT by serving as either an activator or repressor of epithelial or mesenchymal gene expression, respectively. Following induction of EMT by TGF- $\beta$ , we observed an unexpected loss of H2A.Z across both downregulated epithelial and upregulated mesenchymal promoters. Strikingly, the repression of epithelial gene expression was associated with reduction of H2A.Z upstream of the transcription start site (TSS), while the activation of mesenchymal gene expression was dependent on removal of H2A.Z downstream of the TSS. Therefore, the ability of H2A.Z to regulate EMT is dependent on its position, either upstream or downstream of the TSS.

## INTRODUCTION

Epithelial-mesenchymal transition (EMT) is essential for driving plasticity during early development, while the reverse process, mesenchymal-epithelial transition (MET), is responsible for the differentiation of numerous tissues and organs (Thiery et al., 2009). The dynamic nature of this dedifferentiation-re-differentiation process is highlighted by the requirement of up to three successive cycles of EMT/MET during the formation of some tissues in the mammalian developing mammalian embryo (Thiery et al., 2009).

The conversion of epithelial cells to mesenchymal cells involves drastic phenotypic changes that include the loss of cell polarity, the loss of cell-cell adhesion, and the gain of migratory and invasive properties (Thiery et al., 2009). Interest in the EMT process is rising due to the fundamental role that inter-conversion of epithelial cells to mesenchymal cells plays in epithelial tumor metastasis (De Craene and Berx, 2013; Kiesslich et al., 2013). Therefore, elucidating the molecular mechanisms underlying the EMT process not only is important for understanding development processes but will also have broad implications in understanding cancer progression.

Epigenetic/chromatin-based mechanisms are stable enough to confer robustness to steady-state gene expression yet retain the plasticity to be able to direct a new fate for a cell ("epigenetic-based mechanisms" broadly refers to changes in the composition or modification state of chromatin or DNA that regulate patterns of gene expression and chromosome stability). Therefore, it is highly likely that epigenetic modifiers will play a key role in the control of EMT. Supporting this hypothesis, several studies have shown that EMT involves major epigenetic reprogramming, including extensive changes to chromatin modifications (Cieřlik et al., 2013; Malouf et al., 2013; McDonald et al., 2011; Wang and Shang, 2013; Wu et al., 2012). Transforming growth factor  $\beta$  (TGF- $\beta$ ) is a pleiotropic transforming growth factor that can induce EMT in numerous lineages of epithelial cells (Zavadiil and Böttinger, 2005). It was previously observed that, upon TGF- $\beta$ -induced EMT, there is a global reduction in the heterochromatic mark, H3K9me2, with a corresponding bulk increase in histone modifications associated with active transcription (H3K4me3 and H3K36me3) (McDonald et al., 2011). These global histone modification changes were attributed to an increase in the expression of lysine-specific demethylase-1 (McDonald et al., 2011).

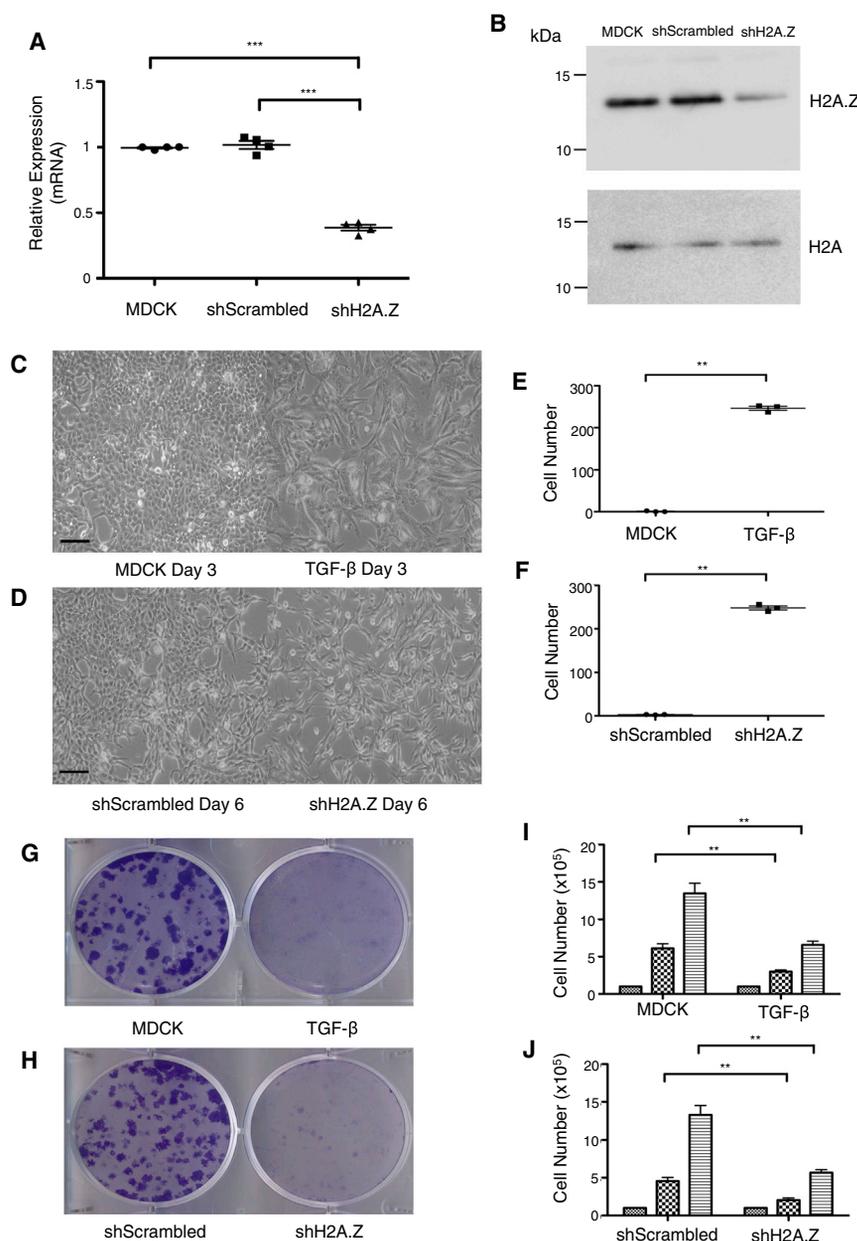
However, despite these widespread changes in histone methylation, these changes cannot account for the regulation of important EMT marker genes (McDonald et al., 2011; Wu et al., 2012). Similarly, histone deacetylases 1–3 have also been shown to play a role in the repression of certain epithelial genes (Wang and Shang, 2013), but neither histone methylation nor acetylation are specifically associated with TGF- $\beta$ -induced EMT.

H2A.Z is evolutionarily conserved and an essential metazoan histone variant of the H2A class (Faast et al., 2001). Mice deficient in H2A.Z die during early development, but the reason for this is unknown (Faast et al., 2001). Previously, we showed that the loss of H2A.Z in *Xenopus laevis* impaired the cell movement required for the formation of the mesoderm and neural crest (Ridgway et al., 2004). Given that mesoderm formation is critically dependent on EMT, we therefore wondered whether H2A.Z might be a chromatin regulator of EMT.

## RESULTS

### Knockdown of the Histone Variant H2A.Z Mimics TGF- $\beta$ -Induced EMT in Madin-Darby Canine Kidney Cells

To investigate whether H2A.Z has a role in EMT, its expression was inhibited in Madin-Darby Canine Kidney (MDCK) cells using



### Figure 1. Inhibition of H2A.Z Expression Mimics TGF-β in Inducing EMT

(A) qRT-PCR analyses of H2A.Z gene expression relative to the Ribosomal Protein Lateral Stalk Subunit P0 (RPLP0) gene in MDCK cells following 72 hr post-transduction with pLVTHM shH2A.Z or pLVTHM shScrambled vector. Error bars denote SD of the mean (n = 4; \*\*\*p < 0.0005).

(B) Western blot analysis of H2A.Z and H2A protein expression following H2A.Z knockdown.

(C) Representative phase-contrast photomicrograph images of mesenchymal-like changes in MDCK cells treated with 5 ng/mL TGF-β for 3 days. Scale bar, 200 μm.

(D) Representative phase-contrast photomicrograph images of mesenchymal-like changes in H2A.Z knockdown cells at 6 days post-transduction. Scale bar, 200 μm.

(E) Cell migration was determined using a Transwell migration assay for untreated and TGF-β-treated MDCK cells (SD of the mean, n = 3; \*\*p < 0.005).

(F) Cell migration was determined using a Transwell migration assay for shH2A.Z and shScrambled transduced cells (SD of the mean, n = 3; \*\*p < 0.005).

(G) Representative clonogenic formation assay of untreated or TGF-β-treated MDCK cells (n = 3).

(H) Representative clonogenic formation assay of shH2A.Z- and shScrambled-transduced cells.

(I) Proliferation of untreated and TGF-β-treated MDCK cells at days 0, 3, and 5 (SD of the mean, n = 3; \*\*p < 0.005).

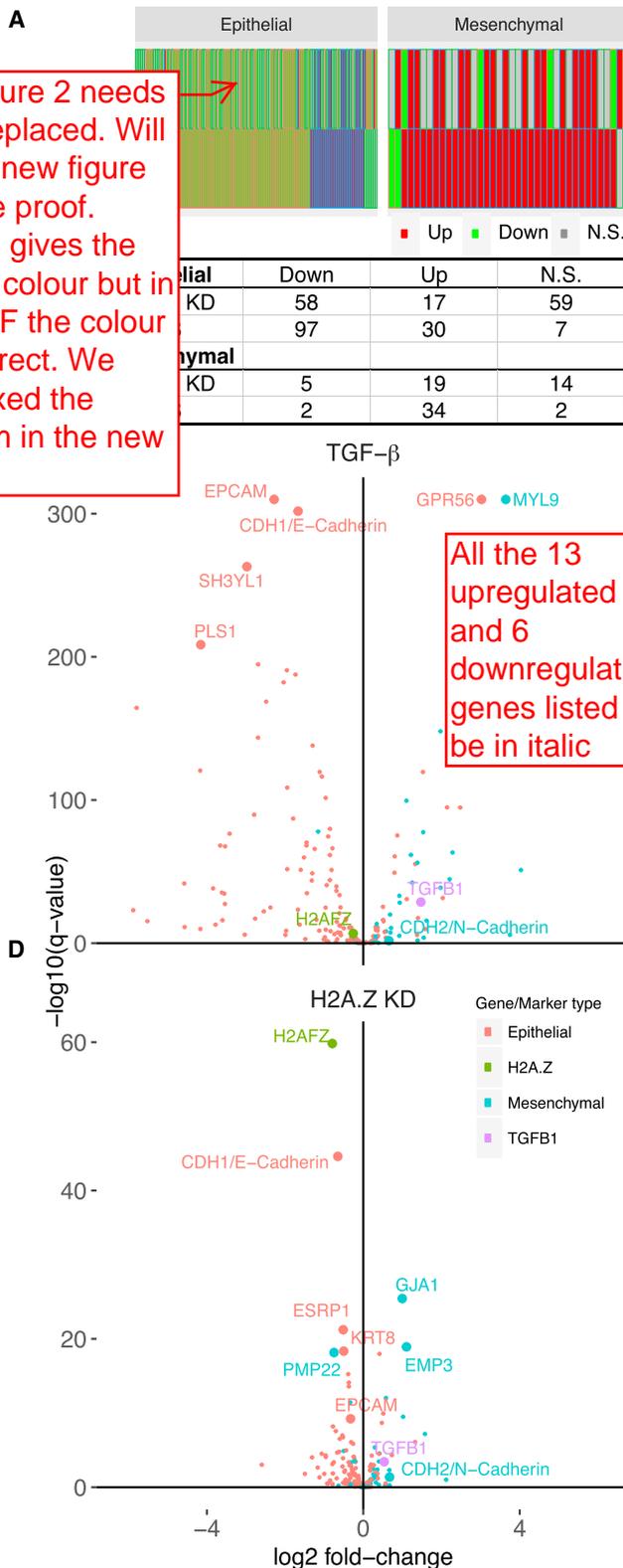
(J) Proliferation of shScrambled and shH2A.Z cells at days 0, 3, and 4 post-transduction (SD of the mean, n = 3; \*\*p < 0.005).

See also [Figures S1, S2, S3, and S6](#).

a short hairpin RNA (shRNA) approach ([Figures 1A and 1B](#)). This cell line has been extensively used as a model system for EMT because the cells fully convert from the epithelial state to the mesenchymal state in response to TGF-β ([Gregory et al., 2008](#)) ([Figure 1C](#)). MDCK cells were transduced either with the lentiviral vector pLVTHM shH2A.Z or with a pLVTHM shScrambled vector for 6 days. Strikingly, MDCK cells depleted in H2A.Z (~2.5 fold) underwent an EMT-like morphological change, whereas MDCK cells transduced with pLVTHM shScrambled vector did not ([Figures 1C and 1D](#)) ([Figure S1](#) shows untreated MDCK cells seeded at a lower cell density).

Quantifiable phenotypic features that distinguish mesenchymal cells from epithelial cells include: (1) their ability to migrate, (2) their inability to form tight and structured colonies,

and (3) a slower proliferation rate. To investigate whether H2A.Z knockdown MDCK cells are able to migrate, migration assays were performed using a Transwell migration chamber. Untreated cells, cells treated with TGF-β, or MDCK cells that were either transduced with pLVTHM shH2A.Z vector or pLVTHM shScrambled vector were seeded onto a 0.8-μm-pore filter of a Transwell insert and allowed to migrate. Quantification revealed that the rate of migration of H2A.Z knockdown MDCK cells was indistinguishable from that of cells that were treated with TGF-β ([Figures 1E and 1F](#)). On the other hand, no cell migration was observed for untreated MDCK cells or cells transduced with pLVTHM shScrambled vector ([Figures 1E and 1F](#)). Colony formation assays revealed that untreated MDCK cells or cells transduced with pLVTHM shScrambled vector formed tightly packed colonies, as expected ([Figures 1G and 1H](#)). In contrast, MDCK cells treated with TGF-β or transduced with pLVTHM shH2A.Z vector failed to form such structured colonies ([Figures 1G and 1H](#)). Finally, knockdown of H2A.Z reduced the rate of cell proliferation in a manner identical to that of MDCK cells treated



This figure 2 needs to be replaced. Will upload new figure with the proof. printing gives the correct colour but in this PDF the colour is incorrect. We have fixed the problem in the new figure

All the 13 upregulated genes and 6 downregulated genes listed should be in italic

with TGF- $\beta$  (Figures 1I and 1J). Based on these phenotypic characteristics, we conclude that loss of H2A.Z mimics TGF- $\beta$  in inducing complete EMT.

**Differential Gene Expression Analysis of EMT Marker Genes in H2A.Z Knockdown and TGF- $\beta$ -Treated MDCK Cells Reveals Highly Similar Transcription Profiles**

Does the loss of H2A.Z also mirror the gene expression changes induced by TGF- $\beta$ ? To investigate this, we first utilized a custom-made canine EMT real-time PCR array comprising 83 EMT-related genes plus five housekeeping genes (*B2M*, *GISB*, *HPRT1*, *LDHAL6B*, and *GAPDH*). Two biological replicates with two technical replicates were performed. Untreated MDCK cells, MDCK cells treated with TGF- $\beta$ , and MDCK cells transduced with either shScrambled or shH2A.Z were analyzed (Tables S1, S2, and S3). Upon treatment with TGF- $\beta$  for 3 days, the array revealed that all the important mesenchymal genes were significantly upregulated, which included *SPARC*, *TWIST1*, *ZEB2*, *IFN1*, *TGFB-1*, *CDH2*, *SERPINE1*, *SNAI2*, *WNT5B*, *ITGAV*, and *VACN*. Downregulated genes included *CDH1*, *MMP9*, *SPP1*, *PDGFRB*, and *KRT7*. Importantly, these gene expression changes observed in TGF- $\beta$ -induced EMT were mimicked in MDCK cells depleted of H2A.Z for 6 days. The housekeeping genes did not change significantly (Tables S1, S2, and S3).

Next, we performed a comprehensive differential gene expression analysis using high-throughput 75-bp paired-end sequencing of poly(A) selected RNA (in triplicate). To focus specifically on the EMT program, in this case, we utilized a generic human 218 gene EMT expression signature that was previously derived from tumors and cell lines of different origins (Tan et al., 2014). Of these 218 genes, 211 homologs were identified in the canine genome. Of these 211 genes, 189 were expressed in MDCK cells, either in the epithelial state or in the mesenchymal state. 169 of these were differentially regulated by TGF- $\beta$ . 34 and 19 genes were upregulated in response to TGF- $\beta$  or H2A.Z knockdown, respectively (Figures 2A and 2B and summarized in 2C). On the other hand, 97 and 58 genes were downregulated in response to TGF- $\beta$  or H2A.Z knockdown, respectively (Figures 2A and 2B and summarized in 2C).

As observed previously (Tan et al., 2014), EMT appears to require the repression of many more epithelial genes compared to the number of genes activated in the mesenchymal state.

mesenchymal marker genes in MDCK cells following treatment with TGF- $\beta$  or knockdown of H2A.Z for 6 days. Genes with a q value >0.1 are indicated in gray as “not significant” (N.S.). RNA-seq experiments were performed in triplicate. (B) The table summarizes the number of differentially expressed genes upon treatment with TGF- $\beta$  or knockdown of H2A.Z.

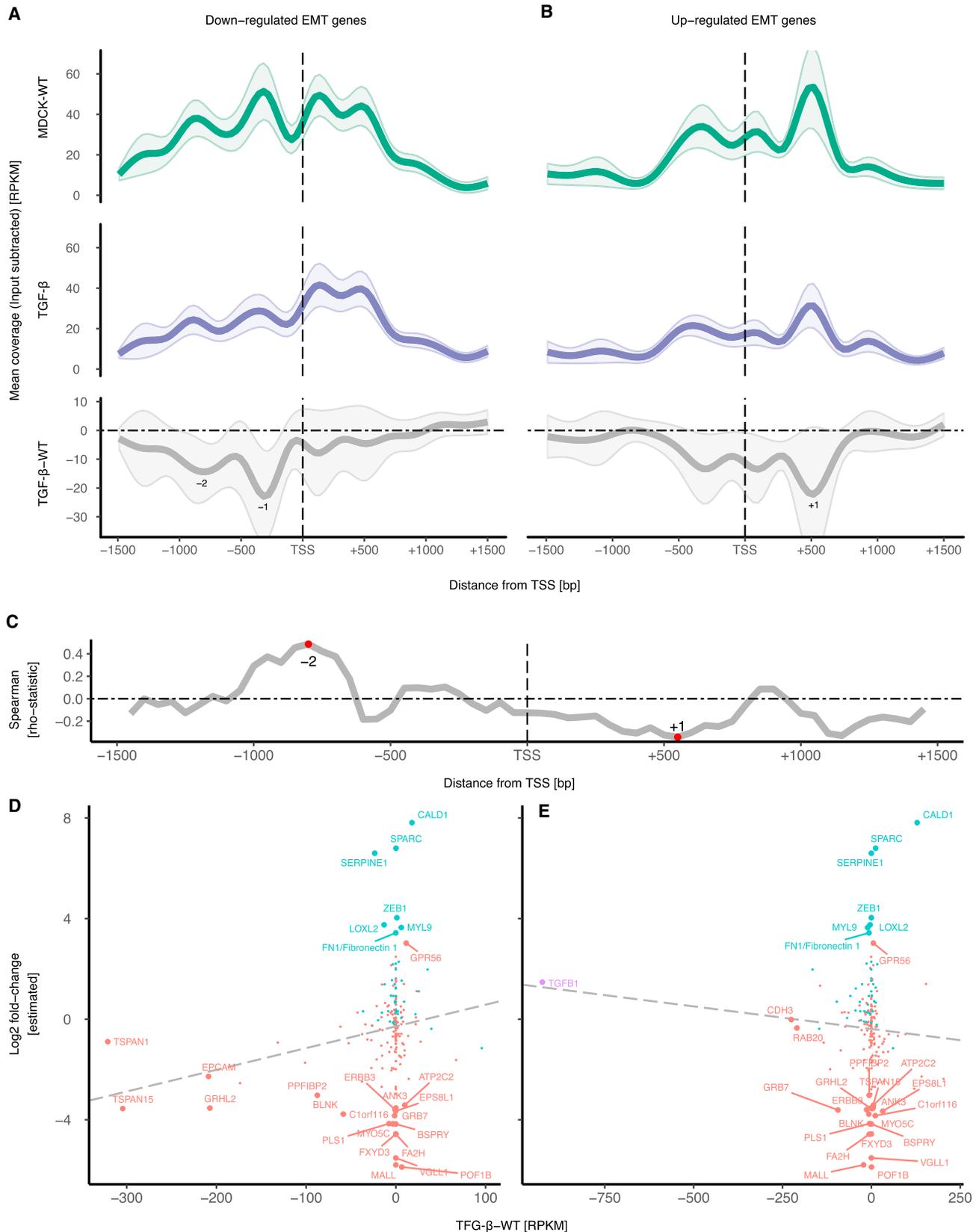
(C) Differentially expressed EMT genes in MDCK cells treated with TGF- $\beta$ . Log<sub>2</sub> fold changes in gene expression are plotted against the inverted log<sub>10</sub>-transformed q values (adjusted p values for multiple testing). The plots broadly illustrate that epithelial genes (red) are downregulated and that mesenchymal genes (blue) are upregulated following both treatments. The expression changes of *H2AFZ* (green) and *TGF $\beta$ 1* (purple) are also shown.

(D) Differentially expressed EMT genes in MDCK cells with H2A.Z knockdown. Generally, TGF- $\beta$  treatment has a stronger effect than H2A.Z knockdown on differential gene expression.

See also Figures S2, S5, and S7 and Tables S1, S2, S3, S4, S5, and S6.

**Figure 2. Knockdown of H2A.Z Mimics TGF- $\beta$  in Inducing the EMT Gene Expression Program**

(A) Heatmap showing the behavior (upregulated is indicated in red; down-regulated is indicated in green) of EMT genes, epithelial marker genes, and



(legend on next page)

Further, despite the generation of a generic EMT expression profile, it was observed that a number of these epithelial or mesenchymal genes respond in a non-canonical manner, dependent on the specific cellular context (Tan et al., 2014). This feature of the EMT gene expression program was also observed here (Figures 2A and 2B). The expression of 2 and 5 designated mesenchymal genes were downregulated rather than being upregulated in response to TGF- $\beta$  or H2A.Z knockdown, respectively (Figures 2A and 2B). Similarly, 30 and 17 designated epithelial genes were upregulated in response to TGF- $\beta$  or H2A.Z knockdown, respectively (Figures 2A and 2B).

To illustrate the magnitude and the statistical significance ( $q$  value) of the changes in gene expression ( $\log_2$  fold change) upon TGF- $\beta$  stimulation or H2A.Z knockdown, differential gene expression volcano plots were produced for all 211 epithelial and mesenchymal genes (Figures 2D and 2E; Tables S3, S4, S5, and S6). The repression or activation of key epithelial genes (*EPCAM* and *E-Cadherin*) and mesenchymal genes (*TGFB-1* and *N-Cadherin*) is highlighted for both TGF- $\beta$  treatment and H2A.Z knockdown (Figures 2D and 2E). This analysis also revealed that *H2AFZ* expression was slightly downregulated (1.2-fold) in response to TGF- $\beta$ . Not surprisingly, there is no significant change in the amount of H2A.Z protein following TGF- $\beta$  treatment (Figure S2). Therefore, it can be concluded that H2A.Z itself is not a major target of TGF- $\beta$ , despite the loss of H2A.Z-inducing EMT.

H2A.Z is encoded by two genes, *H2AFZ* (H2A.Z.1) and *H2AFV* (H2A.Z.2), which were recently shown to have different roles in malignant melanoma (Nock et al., 2012; Zerbino et al., 2016), despite differing by only 3 amino-acid residues. We found that *H2AFV* was expressed at approximately 30-fold less than *H2AFZ* in MDCK cells and that its expression was not affected by shH2A.Z treatment (data not shown). Therefore, we conclude that *H2AFV* has no role in EMT. Furthermore, based on both the real-time PCR array (Tables S1, S2, and S3) and the RNA-sequencing (RNA-seq) data (Figure 2), we conclude that the loss of H2A.Z mirrors the overall gene expression changes induced by TGF- $\beta$ , although for some genes, the magnitude of these expression changes is smaller, which may explain why shH2A.Z-induced EMT is a slower process compared to treatment with TGF- $\beta$  (Figures 1C and 1D).

### TGF- $\beta$ Induces a Loss of H2A.Z from the Promoters of EMT Marker Genes, which Correlates with Changes in Gene Expression

How does the loss of H2A.Z mimic TGF- $\beta$  in inducing the EMT gene expression program? It is conceivable that H2A.Z functions as a repressor of mesenchymal gene expression and/or as an activator of epithelial gene expression. To address this question, we performed H2A.Z chromatin immunoprecipitation sequencing (ChIP-seq) experiments (in duplicate) on MDCK cells that were induced by TGF- $\beta$ . The resulting libraries were sequenced on a paired-end reads. The 189 differentially expressed genes (including TGF- $\beta$ ) were grouped into genes that were either downregulated or upregulated upon TGF- $\beta$  treatment. Genes were grouped according to expression changes, rather than their classification (i.e., epithelial or mesenchymal) because, as noted earlier, not all previously designated mesenchymal and epithelial genes behaved as expected.

For each group of genes, a single line represents the normalized H2A.Z tag counts (input subtracted) at each base pair, which has been aligned with the annotated canine transcription start site (TSS) ( $\pm 1.5$  kb) for both untreated and TGF- $\beta$ -treated cells (Figures 3A and 3B). Most interestingly, while a clear short array of H2A.Z-containing nucleosomes is seen at the promoters of both down- and upregulated EMT genes in untreated MDCK cells (Figures 3A and 3B, green single line), the abundance of H2A.Z within these arrays decreases upon TGF- $\beta$  treatment (Figures 3A and 3B, purple single line). To analyze this decrease in H2A.Z in more detail, H2A.Z ChIP-seq difference profiles (the difference in H2A.Z ChIP signal between TGF- $\beta$ -treated and untreated cells) were produced (Figures 3A and 3B, gray single line). Indeed, there was an overall loss of H2A.Z across both downregulated epithelial and upregulated mesenchymal promoters (Figures 3A and 3B). Even more striking, the repression of epithelial gene expression was associated with the apparent reduction of the  $-1$  and  $-2$  nucleosomes, while the activation of mesenchymal gene expression was linked with the loss of the  $+1$  nucleosome (defined as those nucleosomes that display the highest occupancy immediately downstream or upstream of the annotated TSS, respectively).

To investigate this further, the correlation between EMT gene expression and H2A.Z occupation across  $\pm 1.5$  kb from the TSS

### Figure 3. H2A.Z Nucleosomes Positioned Upstream or Downstream of the TSS Have Opposing Functions in Gene Expression

- (A) Metagene plots of input-subtracted coverage (RPKM) were aligned between  $-1.5$  and  $+1.5$  kb from the TSS for downregulated genes in untreated (MDCK-WT [wild-type]; green) and TGF- $\beta$ -treated (TGF- $\beta$ ; purple) MDCK cells. The corresponding difference plot (untreated coverage subtracted from TGF- $\beta$ -treated coverage) is shown in gray, in which the shaded areas indicate the propagated SEM. H2A.Z ChIP-seq experiments were performed in duplicate.
- (B) Metagene plots of input-subtracted coverage (RPKM) were aligned between  $-1.5$  and  $+1.5$  kb from the TSS for upregulated genes for untreated (MDCK-WT; green) and TGF- $\beta$ -treated (TGF- $\beta$ ; purple) MDCK cells. The corresponding difference plot (untreated coverage subtracted from TGF- $\beta$ -treated coverage) is shown in gray, in which the shaded areas indicate the propagated SEM. H2A.Z ChIP-seq experiments were performed in duplicate.
- (C) Spearman correlation between changes in EMT gene expression and histone variant H2A.Z occupation in 100-bp sliding windows across  $\pm 1.5$  kb from the TSS. The strongest positive correlation upstream of the TSS approximates the  $-2$  nucleosome position from the TSS. A negative correlation between the H2A.Z signal and the change in gene expression is displayed downstream of the TSS, at the approximate  $+1$  nucleosome position.
- (D) In order to explore which specific genes drive the positive correlation, we plotted the changes in H2A.Z signal against estimated  $\log_2$  fold changes in gene expression between untreated and TGF- $\beta$ -treated MDCK cells. At the approximate  $-2$  nucleosome, the epithelial marker *EPCAM* is among the genes with the greatest loss of H2A.Z while simultaneously showing approximately  $-3.5$   $\log_2$  fold change in gene expression.
- (E) In order to explore which specific genes drive the negative correlation, we plotted the changes in H2A.Z signal against estimated  $\log_2$  fold changes in gene expression between untreated and TGF- $\beta$ -treated MDCK cells. At the  $+1$  position, the master regulator gene of EMT, *TGFB1*, displayed the most pronounced loss of H2A.Z, while simultaneously being upregulated ( $\sim 1.5$   $\log_2$  fold).

was determined (Figure 3C). This analysis clearly shows a positive correlation between the repression of gene expression and the loss of H2A.Z upstream of the TSS. Further, the strongest positive correlation occurs at the  $-2$  nucleosome ( $\sim 750$  bp upstream of the TSS; Figure 3C). On the other hand, there is a negative correlation between gene activation and the loss of H2A.Z upstream of the TSS, with one of the strongest negative correlations occurring at the  $+1$  nucleosome position ( $\sim 500$  bp upstream of the TSS; Figure 3C). We conclude that TGF- $\beta$  induces the loss of H2A.Z from EMT gene promoters, but this leads to different transcriptional outcomes for epithelial and mesenchymal expressed genes. The loss of H2A.Z from the  $-2$  nucleosome is correlated with epithelial gene repression, while the loss of H2A.Z from the  $+1$  nucleosome is associated with mesenchymal gene activation. These observations can explain how the knockdown of H2A.Z expression regulates EMT gene expression.

To explore which epithelial and mesenchymal genes regulate their expression by specifically losing H2A.Z at the  $-2$  and  $+1$  positions, respectively, we determined the relationship between the change in gene expression ( $\log_2$  fold change) and the difference in H2A.Z ChIP signal (reads per million mapped [RPKM]) in TGF- $\beta$ -treated and untreated MDCK cells for all 189 EMT genes (Figures 3D and 3E). For most EMT genes, there is no relationship between changes in gene expression and the loss of H2A.Z; however, some genes do display this relationship. Epithelial genes displaying the strongest loss of H2A.Z at the  $-2$  position, including *EPCAM*, *TSPAN15*, *FOX1A*, and *GRHL2*, also show a strong correlation with repression of transcription (Figure 3D). Similarly, a few mesenchymal genes lose H2A.Z at the  $+1$  nucleosome position, but the most dramatic loss occurs at the *TGFB1* gene promoter (Figure 3E).

Individual H2A.Z ChIP-seq profiles for *EPCAM* (Figure 4A) and *TGFB1* (Figure 4B) clearly showed the loss of the  $-2$  and  $+1$  nucleosomes, respectively. Most interestingly, repression of *EPCAM* is associated not only with the loss of the  $-2$  nucleosome but also with a gain of the  $+1$  nucleosome. Quantitative H2A.Z ChIP-PCR assays were performed on the *EPCAM* and *TGFB1* promoters to substantiate these observations (Figure S3). These assays confirmed the loss of H2A.Z from the  $+1$  position or the  $-2$  position from the *TGFB1* gene promoter or *EPCAM* promoter, respectively, when MDCK cells were induced with TGF- $\beta$ . Significantly, the loss of H2A.Z at these important promoter positions also occurred when MDCK cells were transduced with pLVTHM shH2A.Z (Figure S3). Notably, H2A.Z knockdown also resulted in a gain of H2A.Z at the  $+1$  nucleosome position in the *EPCAM* promoter, as seen with TGF- $\beta$  treatment (Figure S3). These findings confirmed that reducing H2A.Z expression mirrors the effects of TGF- $\beta$  treatment. Taken together, these data show that the loss of the  $+1$  H2A.Z-containing nucleosome is both necessary and sufficient for the activation of the *TGF $\beta$ 1* gene.

The induction of *TGFB1* gene expression by TGF- $\beta$ , and its subsequent excretion, creates an autocrine loop enabling MDCK cells to stimulate themselves to maintain the mesenchymal state. Therefore, the induction of TGF- $\beta$  expression caused by the loss of H2A.Z could, in itself, be the primary driver of EMT. To investigate this, conditioned media were collected from H2A.Z knockdown (and shScrambled) cells, which had

been cultured for 2 days, and then added to MDCK cells with and without the TGF- $\beta$  inhibitor SB505124 and subsequently cultured for 4 days (Figure S4). Conditioned media from only shH2A.Z-transduced MDCK cells induced EMT, which was inhibited by SB505124 (Figure S4). This demonstrated that the loss of H2A.Z produces the TGF- $\beta$  autocrine loop.

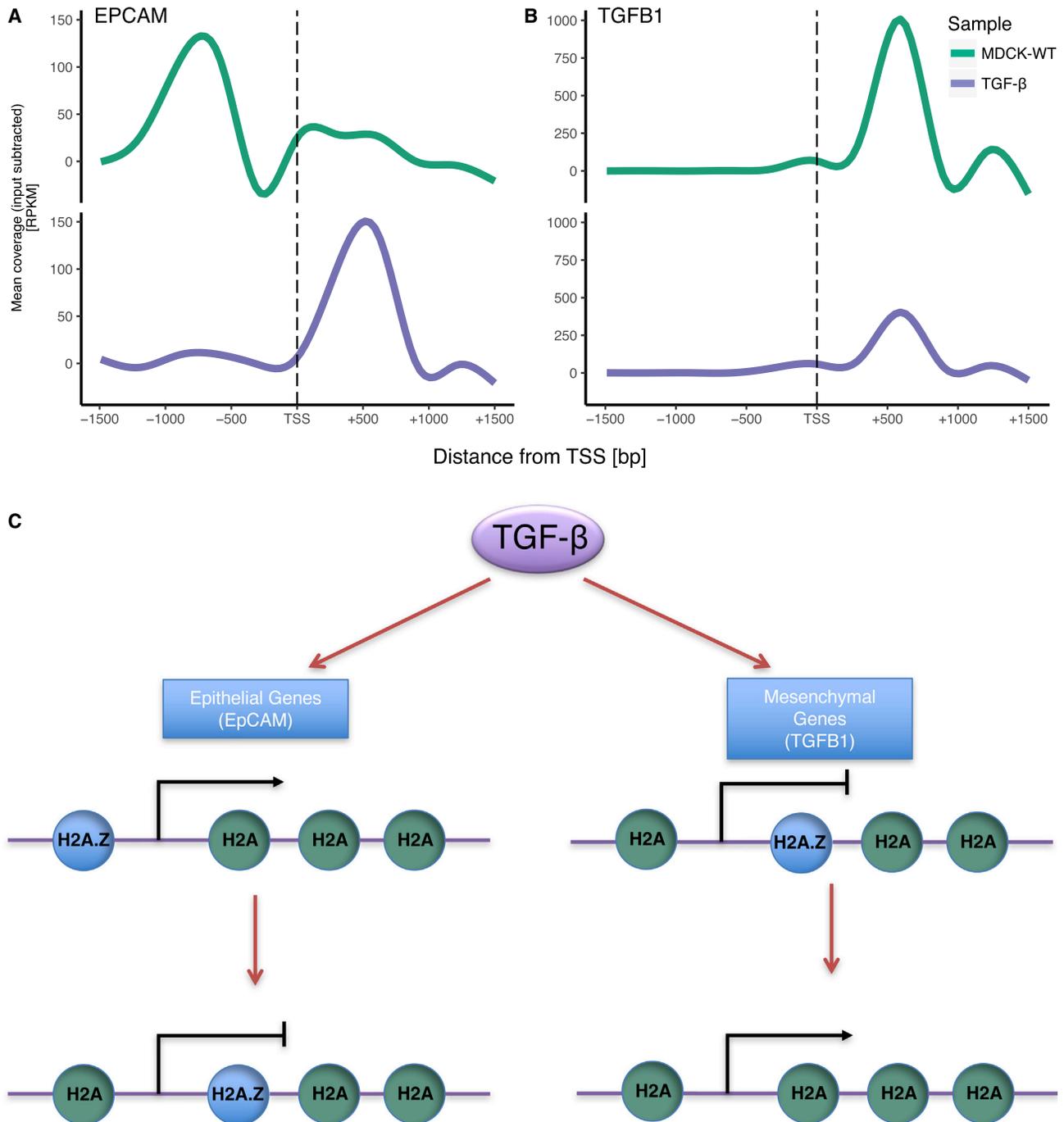
To determine whether this production of endogenous TGF- $\beta$  is sufficient to induce EMT when H2A.Z expression is inhibited, we added the TGF- $\beta$  inhibitor SB505124 immediately following the transduction of MDCK cells with the lentiviral vector pLVTHM shH2A.Z (Figure S5). The addition of SB505124 inhibited the phenotypic changes associated with EMT, showing that, indeed, the ability of shH2A.Z to induce EMT is primarily due to its capacity to activate *TGFB1* gene expression and the autocrine loop (Figure S5). However, despite inhibiting the action of TGF- $\beta$  by this inhibitor, the loss of H2A.Z in MDCK cells still downregulated *EPCAM* expression (Figure S5). On the other hand, the loss of H2A.Z in the presence of SB505124 did not affect the expression of *FN1* (Figure S5). While the *EPCAM* promoter is directly regulated by H2A.Z, the *FN1* promoter is not (Figure 3D). This suggests that the ability of shH2A.Z to promote EMT is a combination of inducing the TGF- $\beta$  autocrine loop and directly regulating key EMT genes.

To investigate whether H2A.Z can control EMT in other cellular contexts, we used MCF10A cells (a human mammary epithelial cell line). Using the same lentiviral system as that used in MDCK cells, the knockdown of H2A.Z, indeed, induced EMT-associated phenotypic and gene expression changes—most notably, the activation of TGF- $\beta$  expression (Figure S6). Taken together, these results strongly support H2A.Z as being a universal regulator of EMT.

Finally, to investigate which transcription factors might be involved in mediating TGF- $\beta$ -induced transcriptional changes in MDCK cells, we computationally identified the transcription factor binding sites present in the promoters of the 211 canine EMT genes analyzed by RNA-seq in Figure 2. To do this, we first searched for the annotated transcription factor binding sites in the promoters of the 218 genes that constitute the human EMT gene signature (Tan et al., 2014). We found 35 transcription factor binding motifs for 24 transcription factors in the promoters of the 218 human EMT genes. Of these, the binding motifs of 20 canine homologs were present in the promoters of the 211 canine EMT genes (Figure S7). Next, we determined how the expression of these transcription factors changed upon TGF- $\beta$  treatment (Table S7). Significant changes in the expression of some of these transcription factors occurred. The transcriptional activator AP1 family (FOS, JUN, and JUND) was downregulated, while USF, SRF, MEF2A, MAX, and ZEB-1 were upregulated, suggesting that these transcription factors may play an important role in the control of the EMT-associated transcriptional changes in MDCK cells.

## DISCUSSION

To date, no master epigenetic regulator of EMT has been identified. Our data indicate that H2A.Z can simultaneously regulate the expression of epithelial and mesenchymal genes. Specifically, TGF- $\beta$  induces the loss of H2A.Z from the promoter



**Figure 4. H2A.Z Coordinates EMT by Serving as Both an Activator and Repressor of Epithelial and Mesenchymal Gene Expression**

(A) Coverage plots of input-subtracted H2A.Z signal (RPKM) across the TSS of *EPCAM* for untreated MDCK cells (MDCK-WT; green) and TGF- $\beta$ -treated MDCK cells (TGF- $\beta$ ; blue).

(B) Coverage plots of input-subtracted H2A.Z signal (RPKM) across the TSS of *TGFB1* for untreated MDCK cells (MDCK-WT; green) and TGF- $\beta$ -treated MDCK cells (TGF- $\beta$ ; blue).

(C) TGF- $\beta$  regulates both *Epithelial Genes* and *Mesenchymal Genes* expression by inducing the loss of H2A.Z. The loss of H2A.Z upstream or downstream of the TSS correlates with gene repression. For *EPCAM*, gene repression is also associated with the formation of a +1 H2A.Z-containing nucleosome.

See also [Figures S3](#) and [S4](#).

regions of both epithelial and mesenchymal genes, but the functional outcomes of this loss are different; epithelial genes become repressed, while mesenchymal genes (most notably, *TGFB1*) are activated. This can be explained by the observation that H2A.Z nucleosomes positioned at different locations in a promoter can have different functions, i.e., the  $-2$  nucleosome is involved in gene activation, whereas the  $+1$  nucleosome is linked to gene repression. This TGF- $\beta$ -induced loss of H2A.Z from both epithelial and mesenchymal promoters is mimicked by the inhibition of H2A.Z expression (Figure S3).

It has been suggested that, in certain contexts, the  $+1$  H2A.Z-containing nucleosome may form a barrier for RNA polymerase II (Kumar and Wigge, 2010; Mao et al., 2014; Soboleva et al., 2014). For example, it was also recently shown that  $+1$  H2A.Z-containing nucleosomes are involved in the repression of memory-promoting genes in the hippocampus (Zovkic et al., 2014). We suggest that this differential function of a H2A.Z nucleosome, dependent on where it sits (upstream or downstream of the TSS), represents a previously unknown mechanism of epigenetic control (Figure 4C). These results also demonstrate that different promoters can have different types of H2A.Z-containing promoter organizations. For example, the active *EPCAM* promoter lacks a  $-1$  H2A.Z-containing nucleosome, while an active *TGFB1* promoter lacks H2A.Z all together. Notably, the *EPCAM* promoter also gains a  $+1$  H2A.Z nucleosome when it becomes repressed (Figures 3 and S3). Concerning the loss of H2A.Z from EMT promoters induced by TGF- $\beta$ , we cannot exclude the possibility that this loss is not complete, i.e., homotypic H2A.Z-H2A.Z nucleosomes may convert to unstable heterotypic H2A.Z-H2A nucleosomes rather than canonical H2A-H2A nucleosomes (Nekrasov et al., 2012). Taken together, we suggest that H2A.Z has an important role in maintaining the epithelial state, because its removal from certain promoters causes **de-differentiation** to the mesenchymal state. This is consistent with the finding that H2A.Z appears to regulate the expression of many more epithelial genes compared to the number of mesenchymal genes (Figure 3).

One key gene that H2A.Z regulates is the *TGFB1* gene itself. Knocking down H2A.Z expression induced the TGF- $\beta$  autocrine loop. However, in the presence of a TGF- $\beta$  inhibitor, the expression of *EPCAM* was still inhibited. This shows that shH2A.Z can regulate the expression of *EPCAM* independently of TGF- $\beta$ . Therefore, we suggest that the ability of shH2A.Z to promote EMT is a combination of inducing the TGF- $\beta$  autocrine loop as well as directly regulating important EMT genes.

TGF- $\beta$  induces EMT by activating major signaling pathways and transcription regulators to create intricate signaling networks (Zavadil and Böttinger, 2005). ATP-dependent remodeling complexes that deposit (SRCAP and P400) and potentially evict (INO80) H2A.Z (Billon and Côté, 2013) regulate the occupancy of H2A.Z at promoters. We suggest, based on the finding that H2A.Z is removed from EMT genes upon TGF- $\beta$  stimulation, that we have uncovered a new connection between the TGF- $\beta$  signaling pathway and these remodeling complexes. Ultimately, transcription factors might be responsible for the targeting of SRCAP, P400, and, possibly, INO80 to promoters (Gévry et al., 2007). Given that the *EPCAM* promoter has an AP1 binding site adjacent to the  $-2$  nucleosome position (at  $-600$  bp downstream

of the TSS; data not shown) and that AP1 is downregulated upon TGF- $\beta$  treatment, a speculative suggestion is that AP1 might recruit SRCAP and/or P400 to deposit H2A.Z on the *EPCAM* promoter but only in the epithelial state. A previous study concluded that AP1 has a critical role in EMT (Cieřlik et al., 2013). Alternatively, given the global reduction of the repressive H3K9me2 mark (McDonald et al., 2011), TGF- $\beta$  may increase the general dynamics of chromatin, thus facilitating the exchange of H2A.Z from the promoters of EMT genes to allow transcription factors (repressors or activators) to bind. Future experiments will explore these non-mutually exclusive possibilities.

The findings of this study will have important implications in understanding cancer progression, given that the conversion of epithelial cells to mesenchymal cells is required for tumor metastasis. However, the role of H2A.Z in cancer (Svotelis et al., 2010) may be even more complex than first envisaged; it was recently shown that the knockdown of H2A.Z expression in human liver cancer cell lines inhibits the expression of fibronectin (a mesenchymal marker gene) and activates E-cadherin expression (an epithelial marker gene) (Yang et al., 2016), which contrasts with our observations. Taken together, these findings suggest that the function of H2A.Z may change during cancer progression, according to cancer type. Finally, we previously showed that H2A.Z is essential for early mouse and *Xenopus* development, but until now, the reason remained unresolved. A central role for H2A.Z in the EMT process, as well as being required for chromosome segregation (Rangasamy et al., 2004), may provide answers to this long-standing question.

## EXPERIMENTAL PROCEDURES

### Cell Culture, shRNA Transfection, and Lentiviral Expression

The culture of HEK293T, MDCK, and MCF10A cells and the derivation and culture of stable knockdown pLVTHM shH2A.Z and pLVTHM shScramble cell lines are described in the Supplemental Experimental Procedures. For EMT induction, MDCK cells were treated with 5 ng/mL TGF- $\beta$ 1 (R&D Systems). Conditioned media were obtained from transduced cells at 75% confluence after 2 days in culture and added to MDCK cells at day 0. After 2 days of culture, equal numbers of cells were re-seeded and cultured in conditioned media in the presence or absence of 1 mM TGF- $\beta$ 1 receptor inhibitor SB505124 (Sigma) and analyzed at day 4.

### Proliferation Assay

Cells were seeded in triplicate in six-well plates at  $1 \times 10^5$  cells per well. Cells were counted using a hemocytometer.

### Real-Time qPCR

Equal amounts of RNA were reverse transcribed to cDNA using the Superscript III First-Strand Synthesis Kit (Invitrogen), as per manufacturer's instructions. qPCR was conducted with the 7900HT Fast Real-Time PCR System using SYBR Green I Master Mix (Applied Biosystems) and 0.15 mM primers. We utilized a customized canine Epithelial to Mesenchymal Transition (EMT) RT<sup>2</sup> Profiler PCR Array from SA Biosciences comprising 83 EMT-related genes and 5 housekeeping genes. For description and further details, see the Supplemental Experimental Procedures.

### Western Blotting

Cells were lysed in buffer containing 250 mM Tris-HCl (pH 6.8), 40% (v/v) glycerol, 5% (w/v) SDS, 0.005% (w/v) bromophenol blue, and 10%  $\beta$ -mercaptoethanol. Western blot analysis was performed using standard protocols, and the following antibodies were used as probes: Anti-H2A.Z (Rangasamy et al., 2003) and Anti-H2A (Abcam, 18255).

### Migration and Clonogenic Assay

For migration assay,  $1 \times 10^5$  cells were seeded into the Transwell migration chamber according to the manufacturer's instructions (Costar) with 10% fetal bovine serum (FBS) in media added as chemoattractant to the wells of the chamber for 4 hr. For clonogenic assay, equal numbers of cells were seeded in triplicate in six-well plates for 10 days. Detailed information is provided in the [Supplemental Experimental Procedures](#).

### RNA-Seq Library Preparation, Sequencing, Data Processing, and Analysis

Stranded mRNA sequencing (mRNAseq) libraries were constructed according to the Illumina TruSeq Stranded mRNA protocol with poly(A) enrichment (Illumina). Libraries were sequenced with  $2 \times 75$  bp paired-end reads on an Illumina NextSeq 500 instrument in high-output mode at the Biomolecular Research Facility of the Australian Cancer Research Foundation, Canberra, Australia. Detailed information about data analysis is provided in the [Supplemental Experimental Procedures](#) and all computer code, consisting of "snake-make" workflow files and R scripts, used in the preparation of this article is accessible at [https://github.com/JCSMR-Tremethick-Lab/H2AZ\\_EMT](https://github.com/JCSMR-Tremethick-Lab/H2AZ_EMT).

### ChIP and ChIP-Seq

ChIP assays and preparation of ChIP-seq libraries were carried out as described in [Nekrasov et al. \(2012\)](#). Libraries for ChIP-seq of DNA were prepared using the ChIP-Seq Sample Prep Kit (Illumina), according to the manufacturer's protocol. DNA from input nucleosomes and H2A.Z ChIP was sequenced on an Illumina NextSeq 500 instrument using 75-bp paired-end reads in high output mode. For further details, see the [Supplemental Experimental Procedures](#).

### ChIP-Seq Data Analysis

The complete workflow consisting of quality checks, adaptor trimming, read alignment, quality filtering, de-duplication, and processing with deepTools was implemented in "snakemake" and can be accessed at [https://github.com/JCSMR-Tremethick-Lab/H2AZ\\_EMT](https://github.com/JCSMR-Tremethick-Lab/H2AZ_EMT).

### Combined Analysis of RNA-Seq and ChIP-Seq Data

All steps of the analysis were performed using R/Bioconductor (3.3.1/3.4). A completely annotated version of the analysis code can be accessed at [https://github.com/skurscheid/MDCK\\_EMT\\_paper](https://github.com/skurscheid/MDCK_EMT_paper) in the form of an R Notebook.

### Transcription Factor Analysis

In order to determine the involvement of specific transcription factors (TFs) in the regulation of EMT genes in TGF- $\beta$ -treated MDCK cells, we first used the Ensembl Regulation ([Zerbino et al., 2016](#)) database to identify transcription factors with binding sites in the promoter of 218 EMT genes ([Tan et al., 2014](#)). Ensembl Regulation also provides identifiers for position-weighted matrices (PWMs) representing the conserved DNA motifs associated with transcription factor binding. The PWMs were then used to search the promoters of canine EMT genes for predicted binding sites. Canine transcription factors associated with EMT genes were then classified as "upregulated" if their beta values were  $>0$  and their q values were  $<0.05$ ; as "downregulated" if their beta values were  $<0$  ( $q < 0.05$ ); or as not significant if their q values were  $>0.05$ .

### DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA-seq data reported in this paper is ArrayExpress: E-MTAB-5628. The accession number for the ChIP-seq data reported in this paper is ArrayExpress: E-MTAB-5637.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and six tables and can be found with this article online at <https://doi.org/10.1016/j.celrep.2017.09.086>.

### AUTHOR CONTRIBUTIONS

R.D. helped design and perform the experiments and also helped with writing the manuscript. S.K. developed and performed the gene expression and ChIP-seq analysis, helped design and interpret the experiments, and also helped with writing the manuscript. M.N. performed the RNA and ChIP-seq experiments. S.H. initiated the project and performed the experiments. D.J.T. conceived the project, helped design the experiments, supervised the analysis, and wrote the manuscript.

### ACKNOWLEDGMENTS

We acknowledge the excellent high-throughput DNA sequencing service provided by our in-house Biomolecular Research Service headed by Stephanie Palmer. This work was supported by a National Health and Medical Research Council (NHMRC) project grant to D.J.T. (1058941).

Received: June 5, 2017

Revised: August 17, 2017

Accepted: September 25, 2017

Published: October 24, 2017

### REFERENCES

- Billon, P., and Côté, J. (2013). Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. *Biochim. Biophys. Acta* *1819*, 290–302.
- Ciešlik, M., Hoang, S.A., Baranova, N., Chodaparambil, S., Kumar, M., Allison, D.F., Xu, X., Wamsley, J.J., Gray, L., Jones, D.R., et al. (2013). Epigenetic coordination of signaling pathways during the epithelial-mesenchymal transition. *Epigenetics Chromatin* *6*, 28.
- De Craene, B., and Bercx, G. (2013). Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* *13*, 97–110.
- Faast, R., Thonglairoam, V., Schulz, T.C., Beall, J., Wells, J.R., Taylor, H., Matthaeei, K., Rathjen, P.D., Tremethick, D.J., and Lyons, I. (2001). Histone variant H2A.Z is required for early mammalian development. *Curr. Biol.* *11*, 1183–1187.
- Gévry, N., Chan, H.M., Laflamme, L., Livingston, D.M., and Gaudreau, L. (2007). p21 transcription is regulated by differential localization of histone H2A.Z. *Genes Dev.* *21*, 1869–1881.
- Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., and Goodall, G.J. (2008). The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* *10*, 593–601.
- Kiesslich, T., Pichler, M., and Neureiter, D. (2013). Epigenetic control of epithelial-mesenchymal-transition in human cancer. *Mol. Clin. Oncol.* *1*, 3–11.
- Kumar, S.V., and Wigge, P.A. (2010). H2A.Z-containing nucleosomes mediate the thermosensory response in *Arabidopsis*. *Cell* *140*, 136–147.
- Malouf, G.G., Taube, J.H., Lu, Y., Roysarkar, T., Panjarian, S., Estecio, M.R., Jelinek, J., Yamazaki, J., Raynal, N.J., Long, H., et al. (2013). Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. *Genome Biol.* *14*, R144.
- Mao, Z., Pan, L., Wang, W., Sun, J., Shan, S., Dong, Q., Liang, X., Dai, L., Ding, X., Chen, S., et al. (2014). Anp32e, a higher eukaryotic histone chaperone directs preferential recognition for H2A.Z. *Cell Res.* *24*, 389–399.
- McDonald, O.G., Wu, H., Timp, W., Doi, A., and Feinberg, A.P. (2011). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat. Struct. Mol. Biol.* *18*, 867–874.
- Nekrasov, M., Amrichova, J., Parker, B.J., Soboleva, T.A., Jack, C., Williams, R., Huttley, G.A., and Tremethick, D.J. (2012). Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics. *Nat. Struct. Mol. Biol.* *19*, 1076–1083.
- Nock, A., Ascano, J.M., Barrero, M.J., and Malik, S. (2012). Mediator-regulated transcription through the +1 nucleosome. *Mol. Cell* *48*, 837–848.

This is the grant number

cannot find titles for  
supp Tables

- Rangasamy, D., Berven, L., Ridgway, P., and Tremethick, D.J. (2003). Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. *EMBO J.* *22*, 1599–1607.
- Rangasamy, D., Greaves, I., and Tremethick, D.J. (2004). RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat. Struct. Mol. Biol.* *11*, 650–655.
- Ridgway, P., Brown, K.D., Rangasamy, D., Svensson, U., and Tremethick, D.J. (2004). Unique residues on the H2A.Z containing nucleosome surface are important for *Xenopus laevis* development. *J. Biol. Chem.* *279*, 43815–43820.
- Soboleva, T.A., Nekrasov, M., Ryan, D.P., and Tremethick, D.J. (2014). Histone variants at the transcription start-site. *Trends Genet.* *30*, 199–209.
- Svetelits, A., Gévry, N., Grondin, G., and Gaudreau, L. (2010). H2A.Z overexpression promotes cellular proliferation of breast cancer cells. *Cell Cycle* *9*, 364–370.
- Tan, T.Z., Miow, Q.H., Miki, Y., Noda, T., Mori, S., Huang, R.Y., and Thiery, J.P. (2014). Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol. Med.* *6*, 1279–1293.
- Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. *Cell* *139*, 871–890.
- Wang, Y., and Shang, Y. (2013). Epigenetic control of epithelial-to-mesenchymal transition and cancer metastasis. *Exp. Cell Res.* *319*, 160–169.
- Wu, C.Y., Tsai, Y.P., Wu, M.Z., Teng, S.C., and Wu, K.J. (2012). Epigenetic reprogramming and post-transcriptional regulation during the epithelial-mesenchymal transition. *Trends Genet.* *28*, 454–463.
- Yang, H.D., Kim, P.J., Eun, J.W., Shen, Q., Kim, H.S., Shin, W.C., Ahn, Y.M., Park, W.S., Lee, J.Y., and Nam, S.W. (2016). Oncogenic potential of histone-variant H2A.Z.1 and its regulatory role in cell cycle and epithelial-mesenchymal transition in liver cancer. *Oncotarget* *7*, 11412–11423.
- Zavadil, J., and Böttinger, E.P. (2005). TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* *24*, 5764–5774.
- Zerbino, D.R., Johnson, N., Juetteman, T., Sheppard, D., Wilder, S.P., Lavidas, I., Nuhn, M., Perry, E., Raffailac-Desfosses, Q., Sobral, D., et al. (2016). Ensembl regulation resources. *Database (Oxford)* *2016*, bav119.
- Zovkic, I.B., Paulukaitis, B.S., Day, J.J., Etikala, D.M., and Sweatt, J.D. (2014). Histone H2A.Z subunit exchange controls consolidation of recent and remote memory. *Nature* *515*, 582–586.