Comprehensive bioinformatic analysis reveals oncogenic role of H2A.Z isoforms in cervical cancer progression

Eric G. Salmerón-Bárcenas [1], Ana E. Zacapala-Gómez [2], Daniela Lozano-Amado [3], Leonardo J. Castro-Muñoz [4], Marco A. Leyva-Vázquez [2], Joaquín Manzo-Merino [4], Pedro A. Ávila-López [1]*

[1] Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Av. Instituto Politécnico Nacional 2508, Col. San Pedro Zacatenco, Delegación Gustavo A. Madero, Ciudad de México

[2] Laboratorio de Biomedicina Molecular, Facultad de Ciencias Químico-Biológicas, Universidad Autónoma de Guerrero, Chilpancingo, 39090, Gro, México

[3] Division of Infectious Diseases, Stanford University School of Medicine, United States

[4] Unidad de Investigación Biomédica en Cáncer, Instituto Nacional de Cancerología, México. Av. San Fernando No. 22, Col. Sección XVI, Tlalpan, 14080, Mexico City, Mexico

A R T I C L E I N F O	A B S T R A C T
	Objective(s): Cervical cancer ranks as the fourth most common neoplasia in women worldwide in which epigenetic alterations play an important role. Several studies have reported pro-oncogenic role of the histone variant H2A.Z in different types of cancer; however, the role of H2A.Z in cervical cancer remains poorly studied. This study aimed to determine the potential role of H2A.Z in cervical cancer through a bioinformatic approach. ***Materials and Methods:*** H2A.Z expression was analyzed in The Human Protein Atlas, The Cancer Genome Atlas, and Gene Expression Omnibus datasets. The promoter regions of H2AZ1 and H2AZ2 genes were downloaded from Expasy, and the prediction of transcription factor binding motifs was performed using CONSITE, Alibaba, and ALGGEN. ChIP-seq and RNA-seq data from HeLa-S3 cells were downloaded from ENCODE. The discovery motif was investigated using MEME-ChIP. The functional annotation was examined in Enrich. ***Results:*** The expression of H2A.Z is elevated in cervical cancer. Interestingly, DNA methylation, copy number, and transcription factors AP2α and ELK1 are involved in H2A.Z overexpression. Additionally, H2A.Z is enriched on promoter and enhancer regions of genes involved in pathways associated with cancer development. In these regions, H2A.Z enables the recruitment of transcription factors such as NRF1, NFYA, and RNA Pol II. Finally, H2A.Z allows the expression of genes associated with proliferation in patients with cervical cancer. ***Conclusion:*** Our findings suggest that H2A.Z overexpression and its presence in promoters and enhancers could be regulating the transcription of genes involved in cervical carcinogenesis.

Introduction

Cervical cancer (CC) is a public health problem, representing the fourth most common cancer among women worldwide with an incidence of 570,000 cases and 311,000 deaths reported in 2018 (1). Importantly, CC ranks second in incidence and mortality in developing countries thus representing a major public health threat (2). The 5-year survival rate is low (68.2%) mainly due to a deficient early diagnostic (3). The main risk factor for developing CC is persistent infection with high-risk Human Papillomavirus (HPV) (4). However, it is well known that infection with high-risk HPV is not sufficient for cervical carcinogenesis, since epigenetic alterations that contribute to the transformation process have been identified. Epigenetics consist of heritable changes in gene expression without modifying the DNA sequence (5). Epigenetic alterations in CC include DNA methylation, histone post-translational modifications,

and non-coding RNAs (6, 7). It has been demonstrated that the histone variants also play an important role in cancer onset (8, 9). For instance, the histone variant H2A.Z is involved in transcriptional control, DNA repair, and regulation of heterochromatin (9). Therefore, alterations in its expression have been associated with oncogenic processes in several types of cancer (10-15). H2A.Z is a highly conserved variant, sharing a 60% similarity with the canonical histone H2A (16). H2A.Z has two isoforms with non-redundant functions known as H2A.Z.1 and H2A.Z.2, which are encoded by H2AZ1 and H2AZ2 genes and regulated by independent promoters (17-19).

Several studies have revealed a role for H2A.Z in regulating processes leading to cancer. In pancreatic cancer, the reduction of the H2A.Z isoforms produces deregulation in the expression of genes associated with proliferation and chemoresistance (15). Specifically,

*Corresponding author: Pedro A. Ávila-López. Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Av. Instituto Politécnico Nacional 2508, Col. San Pedro Zacatenco, Delegación Gustavo A. Madero, Ciudad de México. Tel: +52 55 6098 2694; Email: pedroavila@cinvestav.mx

overexpression of H2A.Z.2 correlates with poor survival in patients with melanoma. In addition, BRD2 and E2F1 proteins interact with H2A.Z.2 at promoter regions to regulate genes involved in proliferation (10). On the other hand, in prostate cancer the acetylation on H2A.Z (K4, K7, and K11) allows greater chromatin accessibility, contributing to the activation of androgen receptor-associated enhancers and gene expression (12). Thus, H2A.Z is relevant to regulate different regulatory elements in cancer, allowing an oncogenic phenotype. However, the effects of H2A.Z.1 and H2A.Z.2 isoforms have not been studied in CC.

In this study, we integrated public data from different omic approaches to decipher the role of H2A.Z in CC. We showed that H2A.Z.1 and H2A.Z.2 isoforms are overexpressed in CC specimens compared with normal tissue. We also identified that both isoforms are associated with progression and nodal metastasis status. Importantly, we demonstrate that H2A.Z is enriched at promoters and enhancers in HeLa-S3 cells, allowing an increase in the expression of cancer-associated genes. Specifically, H2A.Z is associated with enrichment of transcription factors (TF), such as NRF1, NFYA, and RNA Pol II at promoter regions. In summary, our study suggests oncogenic role of H2A.Z in CC.

Materials and Methods
Expression analysis
The H2A.Z expression was analyzed in CC cases using the Human Protein Atlas database (20). The specific expression of H2AZ1 and H2AZ2 was analyzed in 306 CC samples and 13 normal cervical samples using the Gene Expression Profiling Interactive Analysis (GEPIA) webserver from The Cancer Genome Atlas (TCGA) project (21). In addition, GSE9750 and GSE7803 datasets using the GEO2R software from Gene Expression Omnibus (GEO) database (22) were also analyzed (Figure 1). The GSE9750 dataset (Platform: GPL96 [HG-U133A] Affymetrix Human Genome U133A Array) includes 28 CC samples and 24 normal cervical samples (23) whilst the GSE7803 dataset (Platform: GPL96 [HG-U133A] Affymetrix Human Genome U133A Array) includes 21 CC samples and 10 normal cervical samples (24). Expression was log2 transformed and Student's *t*-test was used to determine the differences between conditions, with a *P*-value<0.05 considered as statistically significant. Finally, the expression of H2AZ1 and H2AZ2 was analyzed according to clinical-pathological characteristics in CC cases using the Analyze, Integrate Discover (UALCAN) database (25), and according to copy number using the cBioportal database (26-27).
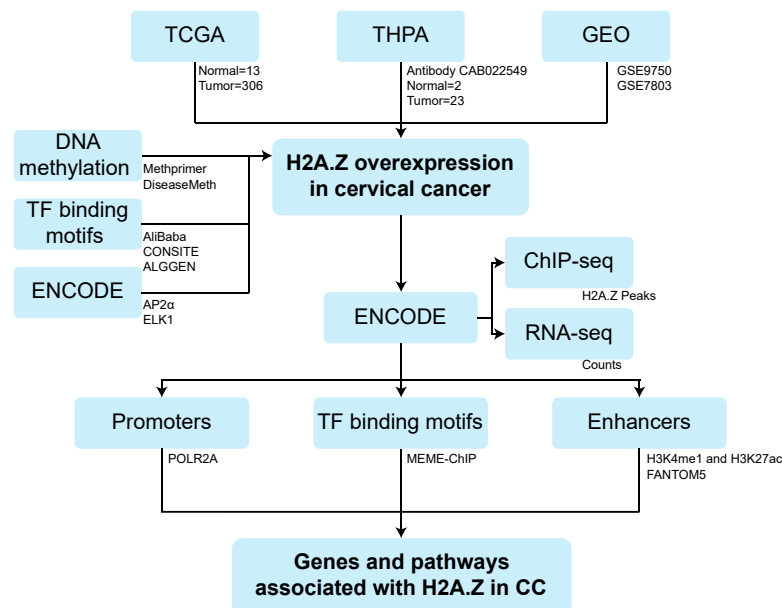
Prediction of transcription factors binding to H2A.Z promoters
The promoter sequences of H2AZ1 and H2AZ2 were downloaded from the Expasy portal (28). A region of -2000 to +2000 base pairs was selected in relation to the transcription start sites (TSS). Then, the AliBaba2.1 (http://gene-regulation.com/pub/programs/alibaba2/index.html), CONSITE (http://consite.genereg.net/) and ALGGEN (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3) software were used to identify potential TF binding motifs (Figure 1).

To confirm the presence of AP2α and ELK1 factors at H2AZ1 and H2AZ2 promoters in HeLa-S3 cells, the tool for interactive visual exploration of diverse genomic data, Integrative Genomics Viewer (IGV), was used (29).

Methylation analysis
The prediction of CpG islands at promoters of H2AZ1 and H2AZ2 was performed using Methprimer software (30). Then, methylation analysis was performed in CC cases using the DiseaseMeth2.0 database (31) (Figure 1).

Figure 1. Workflow of the study
Bioinformatics workflow for deciphering the role of H2A.Z in CC
TCGA: The Cancer Genome Atlas; THPA: The Human Protein Atlas; GEO: Gene Expression Omnibus; TFBM: Transcription Factor Binding Motifs

ChIP-seq data and analysis

To determine H2A.Z enrichment on the genome of CC cells, we analyzed chromatin immunoprecipitation with massively parallel DNA sequencing (ChIP-seq) data available in ENCODE (32-33) (Figure 1). We downloaded raw data corresponding to H2A.Z ChIP-seq (accession numbers: ENCFF000BBJ and ENCFF000BBK) and control ChIP-seq (accession numbers: ENCFF000BAU and ENCFF000BAO) from HeLa-S3 cell line.

Data analysis was performed on the Galaxy platform (34). Quality control of raw data was evaluated using FastQC and then filtered for quality with Trim Galore. Afterward, the reads were mapped to the reference genome hg19 with Bowtie2 using default parameters (35). The unmapped and duplicate reads were filtered using SAMtools (36) using the following parameter: MAPQ quality score (-q) 20. The H2A.Z peaks with respect to the control were determined with MACS2 (37), with the following parameters: --nomodel, broad regions --broad, extension size --extsize 300, and peak detection based on q-value=0.05 (Supplementary Table 1). Finally, to visualize the ChIP-seq signal, deepTools2 (38) was used with the parameter --binSize 25. In order to determine the annotation genomic region of the H2A.Z peaks, we used the ChIPseeker package (39).

To determine the association between H2A.Z and the enrichment of NRF1 and NFYA in HeLa-S3 cells, we downloaded the data processed by ENCODE corresponding to bigwig of NRF1 (accession number: ENCFF000XJF) and NFYA (accession number: ENCFF000XIR). The bigwig files were viewed in IGV genome browser.

RNA-seq data and analysis

To determine whether the gene expression levels are associated with enrichment of H2A.Z at promoters and enhancers, we analyzed RNA sequencing (RNA-seq) data available in ENCODE (Figure 1). Raw data corresponding to RNA-seq from HeLa-S3 cells (accession numbers: ENCFF000FOM/ENCFF000FOV and ENCFF000FOK/ENCFF000FOY) were downloaded.

Data analysis was performed on the Galaxy platform. The quality control of the raw data was evaluated using FastQC and then filtered for quality with Trim Galore. Reads were mapped to the reference genome hg19 with HISAT2 using default parameters (40). Expression levels were determined with featureCounts using RefSeq as gene annotation (41) (Supplementary Table 2). Genes showing zero counts were discarded.

To determine the differentially expressed genes between CC patients with low and high H2A.Z levels we used data from TCGA-CESC (42). Briefly, we sorted the patients according to the expression levels of both H2A.Z isoforms selecting the top 20 patients with low H2AZ1, high H2AZ1, low H2AZ2, and high H2AZ2. Differential expression between "low H2A.Z.1 *vs* high H2A.Z.1" and "low H2A.Z.2 *vs* high H2A.Z.2" was performed using TCGAbiolinks and DESeq2 packages from the R program (43-44). Cut-off or differential expression was an adjusted *P*-value<0.05 and a fold change of 1.5.

TF binding and Motif discovery

To determine the enrichment of TFs at the promoter regions of the genes regulated by H2A.Z, we used the ChIP Enrichment Analysis (ChEA) tool (45), implemented within Enrichr (46). The detection of transcription factor binding motifs at the enrichment regions of H2A.Z was carried out with the MEME-ChIP database (47-48) (Figure 1). A 500 base pairs window from the center of the H2A.Z peaks was selected. The parameters used were motifs between 7 and 25 base pairs in width (average width 13.4) from the Human and Mouse database (HOCOMOCO v11 FULL). MEME-ChIP was searched for motifs with an E-value<0.05.

The protein-protein interactions network between H2A.Z and TFs was performed in the STRING database (49). We used the H2A.Z.1, H2A.Z.2, and TFs obtained from ChEA and MEME-ChIP as input data.

Enhancer regions in HeLa-S3 cells

To determine the enhancer regions in HeLa-S3 cells, we downloaded the data processed by ENCODE corresponding to bed narrowPeak of the H3K4me1 (accession numbers: ENCFF250OWR and ENCFF086NLE), H3K27ac (accession numbers: ENCFF927JDY and ENCFF101ZZI), and POLR2AphosphoS2 (accession number: ENCFF001VJB) (Figure 1).

The enrichment of H2A.Z at validated enhancers by FANTOM5 was verified (50). To visualize the signal of the H3K4me1, H3K27ac, POLR2AphosphoS2, and H2A.Z deepTools2 was used.

To determine the genes near enhancer regions we used the Genomic Regions Enrichment of Annotations Tool (GREAT) with default parameters (51).

Functional enrichment analysis

To determine the biological processes and pathways regulated by H2A.Z we used the Enrichr database (46). We performed an analysis of biological process and pathways enrichment considering a *P*-value<0.05 as statistically significant.

To determine the processes regulated by low H2A.Z *vs* high H2A.Z in CC patients, a Gene Set Enrichment Analysis (GSEA) (52) was performed.

Statistical analysis

Visualization and statistical analysis were performed in the ggplot2 package of R (53). A *P*-value<0.05 was considered statistically significant.

Results

The expression of both H2A.Z.1 and H2A.Z.2 isoforms is elevated in cervical cancer.

To decipher the potential role of H2A.Z in CC we integrated public data from different omic approaches (Figure 1). First, we evaluated the expression of both H2A.Z.1 and H2A.Z.2 isoforms in CC patients by analyzing their expression in samples from the TCGA database. Increased levels of H2A.Z.1 and H2A.Z.2 were found in CC samples compared with normal tissue (Figures 2A, B). Similar results were obtained by analyzing two microarray expression profiling datasets (GSE9750 and GSE7803) from the public GEO database, where an increase of H2A.Z.1 and H2A.Z.2 in CC samples compared with normal cervical samples was found (one-way ANOVA; H2AZ1 *P*-value<0.05 and H2AZ2 *P*-value<0.05, respectively; Figures 2C, D). Additionally, we identified
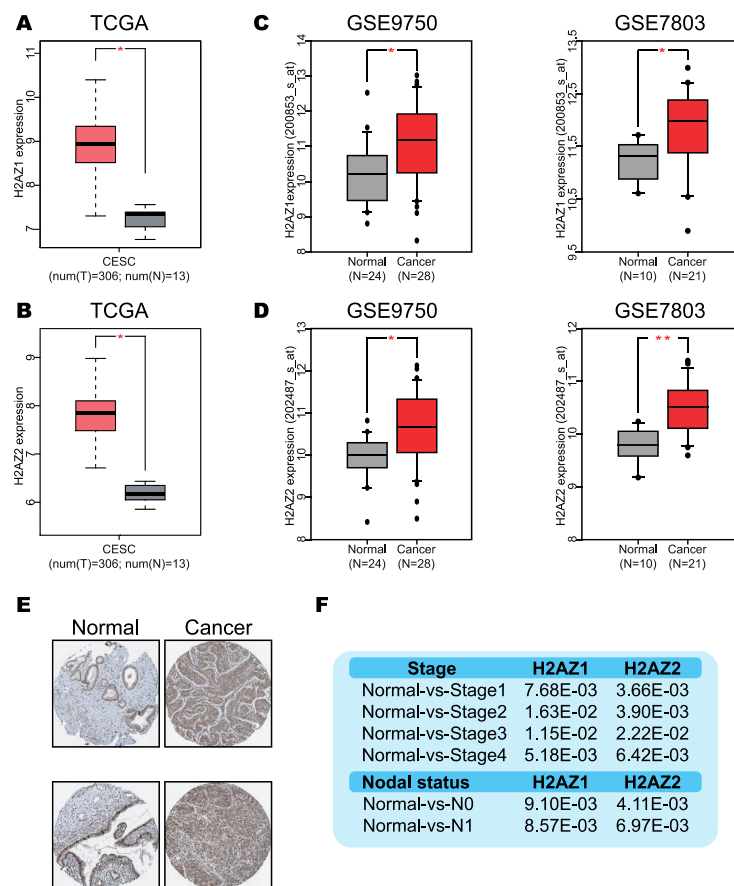
Figure 2. H2A.Z overexpression in cervical cancer

(A) H2AZ1 expression in CC samples and normal samples from TCGA dataset. One-way ANOVA, *P*-value<0.05. (B) H2AZ2 expression in CC samples and normal samples from TCGA dataset. One-way ANOVA, P-value<0.05. (C) H2AZ1 expression in CC samples and normal samples from GSE9750 and GSE7803 datasets. Student's t-test, *P*-value<0.05. (D) H2AZ2 expression in CC samples and normal samples from GSE9750 and GSE7803 datasets. Student's t-test, *P*-value<0.05-left and <0.01-right. (E) H2A.Z.1 y H2A.Z.2 expression in CC samples and normal samples from The Human Protein Atlas database. Two representative images are depicted. (F) H2AZ1 and H2AZ2 expression in CC samples according to FIGO stages and metastasis status from TCGA database. Student's t-test, *P*-value is shown for each group compared with normal samples

high levels of H2A.Z in tissues from CC patients compared with normal samples in The Human Protein Atlas database evaluated by immunohistochemistry (Figure 2E). These results reveal increased levels of H2A.Z in CC, suggesting a potential role in disease progression.

To determine whether the increased levels of H2A.Z are associated with CC progression, the expression levels relative to tumor grade in the TCGA database were analyzed using UALCAN. A significant increase was found in H2A.Z.1 and H2A.Z.2 isoforms in the different stages of CC comparing with normal samples (Figure 2F and supplementary Figure 1A). Interestingly, we identified a significant increase of both isoforms according to the presence of nodal metastasis, indicating potential role of H2A.Z in tumor invasiveness (Figure 2F and supplementary Figure 1B). Together, these data suggest that the increase in H2A.Z.1 and H2A.Z.2 contributes to the CC progression.

The increase of H2A.Z.1 and H2A.Z.2 in CC is regulated by TFs, DNA methylation, and copy number gain

To evaluate potential mechanisms involved in the overexpression of H2A.Z.1 and H2A.Z.2 isoforms in CC, we first identified the TF binding motifs present at promoters of both isoforms. By analyzing the promoter regions of both genes (-2.0 to +2.0 kb relative to TSS) with AliBaba2.1, CONSITE, and ALGGEN programs, binding motifs for YY1 and AP2α were identified at H2AZ1 promoter, and binding motifs for CREB, ELK1, E2F, and AP2α at H2AZ2 promoter (Figure 3A). In addition, we evaluated the expression levels of AP2α, which was found overexpressed in CC compared with normal tissue (Figure 3B), suggesting that this TF could induce the overexpression of both H2A.Z isoforms. Interestingly, a significant enrichment of AP2α at H2AZ2 promoter in the CC cell line, HeLa-S3, was identified (Figure 3C). Moreover, a significant enrichment of ELK1 at H2AZ1 and H2AZ2 promoters was also found (Figure 3C), suggesting that AP2α and ELK1 could be regulating the overexpression of H2AZ1 and H2AZ2 in CC.

Notably, H2AZ1 and H2AZ2 promoters harbor CpG islands (Figure 3D), suggesting a possible role of DNA methylation in H2A.Z regulation. To evaluate the methylation status of H2AZ1 and H2AZ2 promoters in CC, we used the DiseaseMeth2.0 database finding a significant reduction in H2AZ2 promoter methylation in CC compared with normal tissue (Student's *t*-test; H2AZ2 *P*-value=3.538e-07; Figure 3E right), suggesting
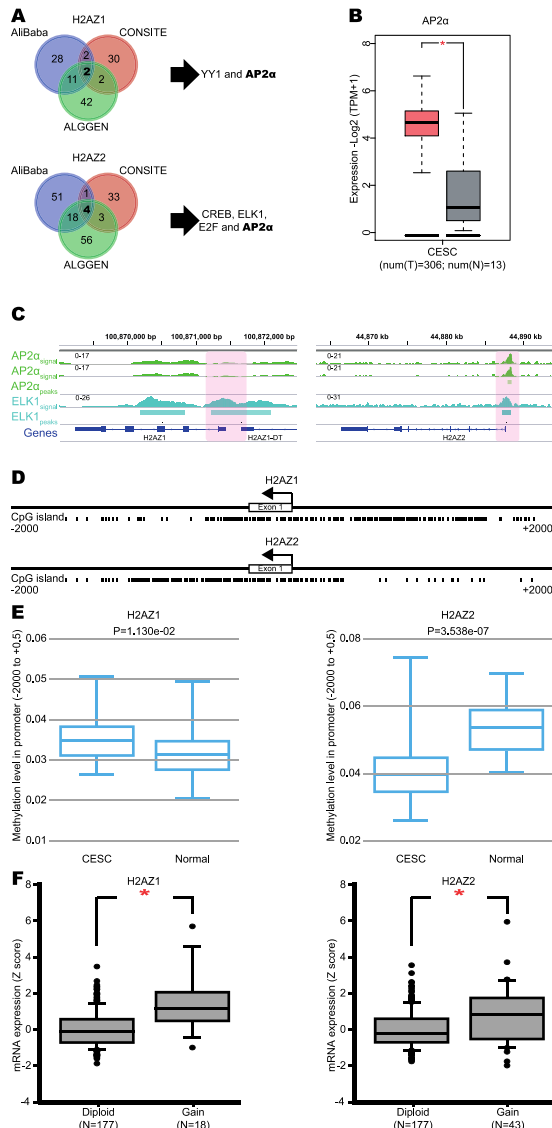
Figure 3. Transcription factors, DNA methylation, and copy number associate with increased H2A.Z.1 and H2A.Z.2
(A) Identification of TF binding sites at H2AZ1 and H2AZ2 promoters. (B) AP2α expression in CC samples and normal samples from the TCGA database. One-way ANOVA, *P*-value<0.05. (C) IGV tracks of the AP2α and ELK1 at H2AZ1 and H2AZ2 promoters in HeLa-S3. The red area defines the promoter region of both genes. (D) Identification of CpG islands at H2AZ1 and H2AZ2 promoters. (E) Methylation level at H2AZ1 and H2AZ2 promoters. Student's t-test, H2AZ1 P-value=1.130e-02 and H2AZ2 *P*-value=3.538e-07. (F) Expression of H2A.Z.1 and H2A.Z.2 isoforms in CC samples with copy number gain compared with diploid CC samples. Student's t-test, H2AZ1 *P*-value<0.05

that H2AZ2 hypomethylation could be associated with overexpression of the H2A.Z.2 isoform. Regarding the H2AZ1 gene, a significant increase was identified in the methylation grade of its promoter in CC compared with normal tissue (Student's *t*-test; H2AZ1 *P*-value=1.130e-02; Figure 3E left), indicating that methylation status could not affect the high expression of this isoform. Finally, using the cBioportal database, we identified a significant increase in copy number gain in patients with elevated levels of H2A.Z.1 and H2A.Z.2 isoforms in CC (Student's *t*-test; H2AZ1 *P*-value<0.05 and H2AZ2 *P*-value<0.05; Figure 3F). Taken together,
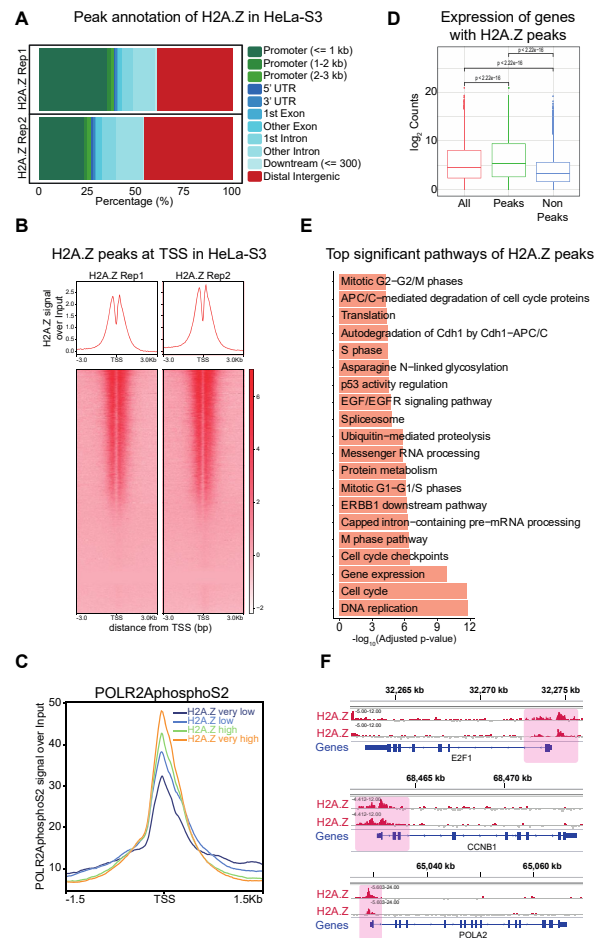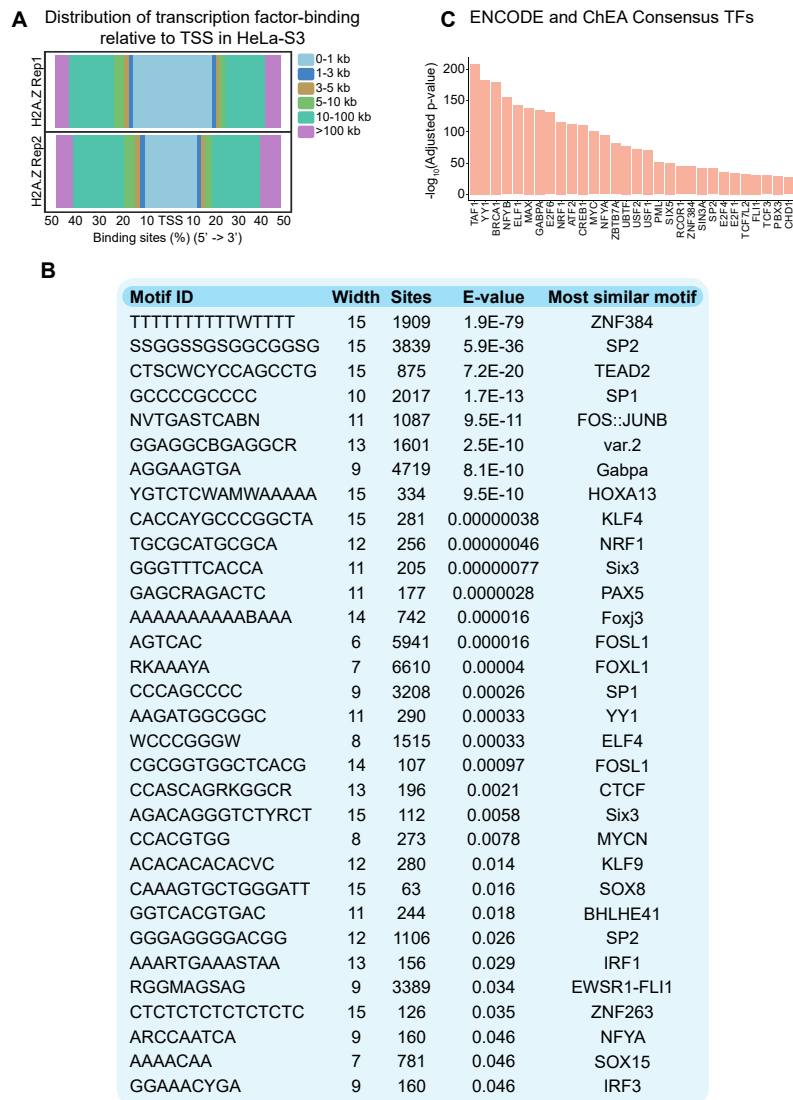

Figure 4. Genome occupancy of H2A.Z in HeLa-S3 cells
(A) Genomic region annotation of the H2A.Z peaks detected in the two replicates in HeLa-S3. (B) Heatmap and profile plots of H2A.Z enrichment at promoters. The signal intensity of H2A.Z is shown with respect to the input. (C) Profile plots of POLR2AphosphoS2 enrichment at promoters. The signal intensity of POLR2AphosphoS2 is shown with respect to the input in four groups of genes sorted with respect to the enrichment of H2A.Z. (D) Box plot of the expression levels of genes with H2A.Z enrichment and not H2A.Z-bound at promoter region in HeLa-S3. Genes with counts of zero were discarded. Wilcoxon test, *P*-value <2.22e-16. (E) Functional annotation of genes with H2A.Z enrichment at promoters. The top 20 of the most significant pathways are shown. We considered an adjusted *P*-value<0.05 as statistically significant. (F) IGV tracks of H2A.Z at the promoters of E2F1, CCNB1, and POLA2 in HeLa-S3. The red area defines the promoters

these results suggest that overexpression of H2A.Z.1 and H2A.Z.2 isoforms could be associated with genetic and epigenetic alterations in CC.

H2A.Z is enriched at promoters of genes associated with proliferation in HeLa-S3 cells

To understand how H2A.Z might promote CC progression, public data from H2A.Z ChIP-seq of the CC cell line HeLa-S3 were analyzed. The antibody used for ChIP-seq does not discriminate between the two H2A.Z isoforms, therefore the enrichment corresponds to both isoforms. It was identified that H2A.Z is distributed mainly at promoters (~33%) and distal intergenic regions (~42%) (Figure 4A). We found a clear enrichment of H2A.Z at promoter regions (-3.0 to +3.0 kb relative to TSS) (Figure 4B), associated with nucleosome-

A Distribution of transcription factor-binding relative to TSS in HeLa-S3

C ENCODE and ChEA Consensus TFs

B

| Motif ID	Width	Sites	E-value	Most similar motif
TTTTTTTTTTWTTTT	15	1909	1.9E-79	ZNF384
SSGGSSGSGGCGGSG	15	3839	5.9E-36	SP2
CTSCWCYCCAGCCTG	15	875	7.2E-20	TEAD2
GCCCCGCCCC	10	2017	1.7E-13	SP1
NVTGASTCABN	11	1087	9.5E-11	FOS::JUNB
GGAGGCBGAGGCR	13	1601	2.5E-10	var.2
AGGAAGTGA	9	4719	8.1E-10	Gabpa
YGTCTCWAMWAAAAA	15	334	9.5E-10	HOXA13
CACCAYGCCCGGCTA	15	281	0.00000038	KLF4
TGCGCATGCGCA	12	256	0.00000046	NRF1
GGGTTTCACCA	11	205	0.00000077	Six3
GAGCRAGACTC	11	177	0.0000028	PAX5
AAAAAAAAAABAAA	14	742	0.000016	Foxj3
AGTCAC	6	5941	0.000016	FOSL1
RKAAAYA	7	6610	0.00004	FOXL1
CCCAGCCCC	9	3208	0.00026	SP1
AAGATGGCGGC	11	290	0.00033	YY1
WCCCGGGW	8	1515	0.00033	ELF4
CGCGGTGGCTCACG	14	107	0.00097	FOSL1
CCASCAGRKGGCR	13	196	0.0021	CTCF
AGACAGGGTCTYRCT	15	112	0.0058	Six3
CCACGTGG	8	273	0.0078	MYCN
ACACACACACVC	12	280	0.014	KLF9
CAAAGTGCTGGGATT	15	63	0.016	SOX8
GGTCACGTGAC	11	244	0.018	BHLHE41
GGGAGGGGACGG	12	1106	0.026	SP2
AAAARTGAAASTAA	13	156	0.029	IRF1
RGGMAGSAG	9	3389	0.034	EWSR1-FLI1
CTCTCTCTCTCTCTC	15	126	0.035	ZNF263
ARCCAATCA	9	160	0.046	NFYA
AAAACAA	7	781	0.046	SOX15
GGAAACYGA	9	160	0.046	IRF3

Figure 5. Transcription factor binding to H2A.Z peaks in HeLa-S3 cells
(A) Distribution of transcription factor binding relative to TSS in HeLa-S3 identified with ChIPseeker. (B) Transcription factor binding motifs at H2A.Z peaks in HeLa-S3 identified with MEME-ChIP. We considered an adjusted *E*-value<0.05 as statistically significant. (C) Bar plot of the TFs associated with genes regulated by H2A.Z identified with ChEA and ENCODE databases. The top 30 of the most significant TFs are shown. An adjusted *P*-value<0.05 was considered statistically significant

free regions as reported (54). Previously, it has been suggested that the enrichment of H2A.Z at promoter regions allows gene expression (10). Interestingly, we found that H2A.Z enrichment is associated with a gradual increase in the presence of POLR2AphosphoS2 at promoters (Figure 4C) (S2 phosphorylation of POLR2A predominates during transcription elongation) (55), as well as high transcription rates compared with genes that lack the presence of H2A.Z at promoters (Wilcoxon test; *P*-value<2.22e-16) (Figure 4D). Pathway analysis showed that those genes enriched with H2A.Z regulate functions associated with a proliferative phenotype such as DNA replication, cell cycle, gene expression, among others (Figure 4E). Moreover, Figure 4F shows H2A.Z enrichment at the promoters of E2F1, CCNB1, and POLA2, genes involved in the proliferation of cancer cells. Together, these results suggest that H2A.Z is associated with recruitment of RNA Pol II at promoters of highly expressed genes associated with a proliferative phenotype in CC.

H2A.Z is associated with the recruitment of transcription factors

It has been suggested that H2A.Z can coordinate the accessibility of TFs to promoter regions (56). To determine which TFs are associated with H2A.Z peaks at promoter regions, the ChIPseeker annotation tool was used. We identified enrichment in the distribution of TF binding sites at promoters (1 kb relative to TSS) (Figure 5A). To evaluate the TF binding motifs present in these regions, we performed an analysis using MEME-ChIP. A significant enrichment of TF binding motifs, such as ZNF384, SP1/2, TEAD2, and FOS-JUNB, among others was found (Figure 5B). In addition, an ENCODE and ChEA Consensus analysis identified TFs enriched at promoters of H2A.Z-enriched genes (Figure 5C). Interestingly, these TFs establish a regulatory network with both H2A.Z isoforms (nodes: 51, edges: 206, PPI enrichment *P*-value<1.0e-16; supplementary Figure 2A), regulating important processes such as transcriptional
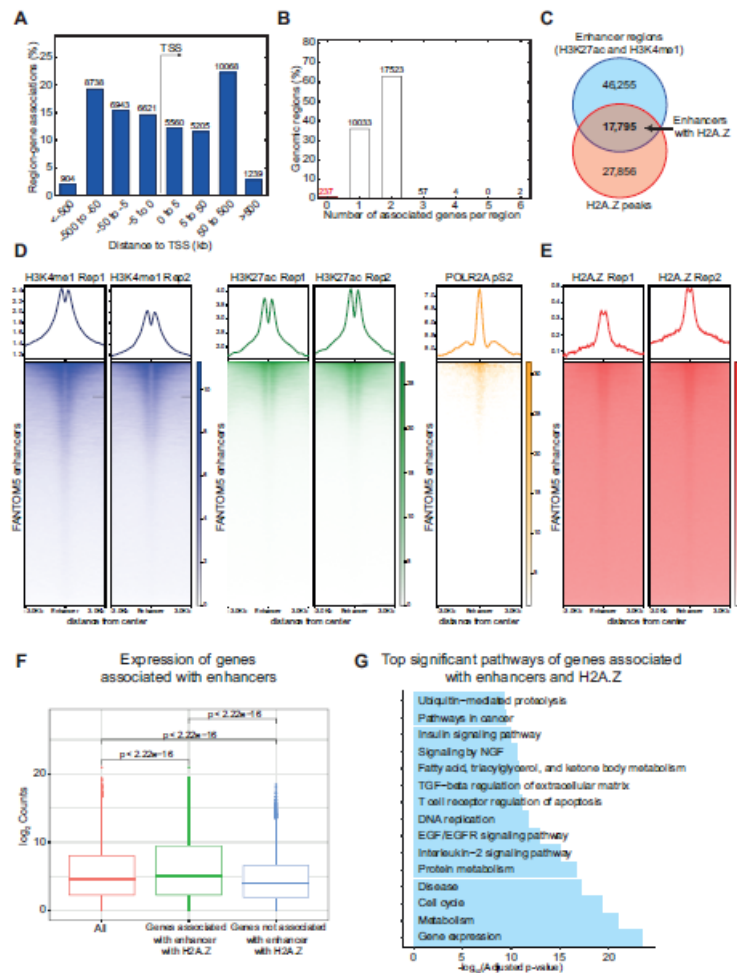
Figure 6. H2A.Z at enhancers in HeLa-S3 cells

(A) Orientation and distance to TSS of the H2A.Z peaks in HeLa-S3 identified with GREAT. The gene number of different regions are shown. (B) Number of associated genes with H2A.Z peaks in HeLa-S3 identified with GREAT. The gene number of different regions are shown. (C) Venn diagram for enhancers and H2A.Z peaks in HeLa-S3. The enhancers identified with H3K27ac and H3K4me1 in HeLa-S3 are represented in blue and the H2A.Z peaks in red. (D) Heatmap and profile plots of the enrichment of H3K4me1, H3K27ac, and POLR2AphosphoS2 at FANTOM5 enhancers. The signal intensity is shown with respect to the input. (E) Heatmap and profile plots of enrichment of H2A.Z at FANTOM5 enhancers. The signal intensity of H2A.Z is shown with respect to the input. (F) Box plot of expression levels of genes associated with H2A.Z enhancers in HeLa-S3. Genes with counts of zero were discarded. Wilcoxon test, *P*-value <2.22e-16. (G) Functional annotation of genes associated with H2A.Z enhancers in HeLa-S3. The top 15 of the most significant pathways are shown. We considered an adjusted *P*-value<0.05 as statistically significant
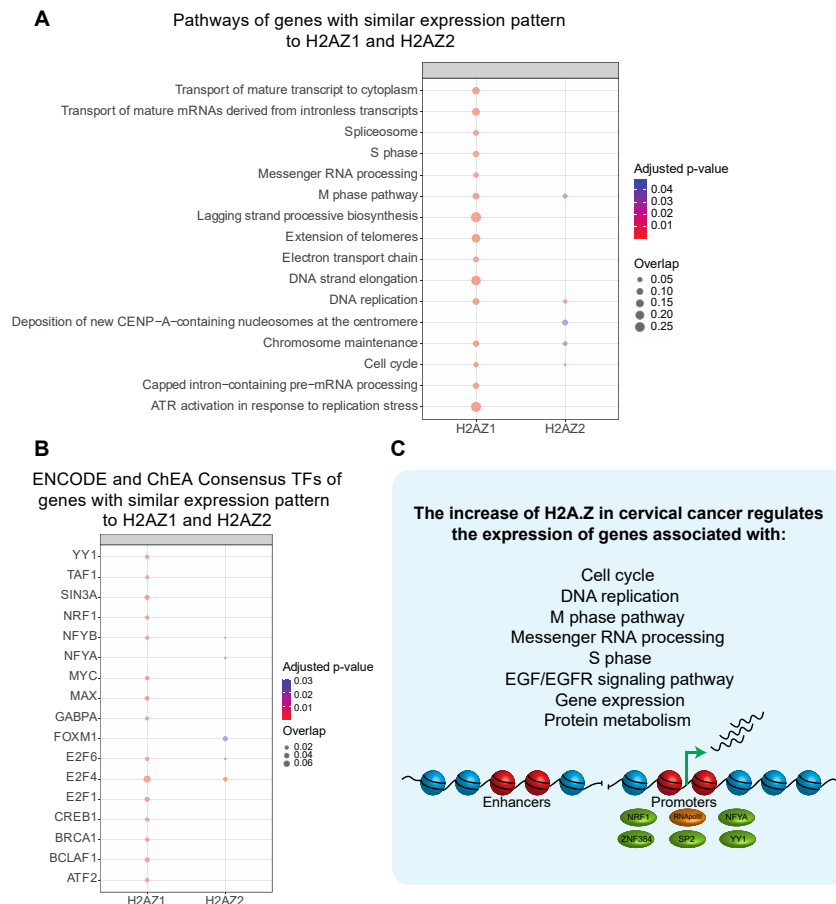
misregulation in cancer (Benjamini-Hochberg method, adjusted *P*-value=3.981e-9; supplementary Figure 2B). Notably, a high H2A.Z enrichment at promoters is associated with a gradual increase in the presence of NRF1 and NFYA at these regions (supplementary Figure 2C). Overall, these data suggest that both H2A.Z.1 and H2A.Z.2 isoforms can form a complex regulatory network with TFs to regulate the expression of cancer-associated genes, thus promoting progression events in CC.

H2A.Z is enriched at enhancers in HeLa-S3 cells

Recently, it was shown that H2A.Z promotes the expression of different oncogenes by activating enhancer regions (12). To further support whether H2A.Z is located at distal regions in HeLa-S3, we use the GREAT tool identifying an H2A.Z enrichment at distal regions (Figures 4A and 6A), suggesting its presence at enhancer regions. Also, most of H2A.Z peaks are associated with regulation of 1 or 2 genes (Figure 6B), as has been described for enhancers (57). To determine

the enhancer regions in HeLa-S3 cells, we analyzed enrichment peaks for the histone marks H3K4me1 and H3K27ac. We identified 46,255 enhancers (overlapping regions between both histone marks) in HeLa-S3 cells (Figure 6C). Interestingly, we identified 17,795 regions overlapping with H2A.Z peaks (Figure 6C), suggesting the presence of H2A.Z at HeLa-S3 enhancers. To confirm this result, validated enhancer regions by FANTOM5 were analyzed, which showed enrichment of H3K4me1, H3K27ac, and POLR2AphosphoS2 (Figure 6D), confirming their identity (57). As expected, H2A.Z is enriched in these validated enhancers (Figure 6E). These results suggest that H2A.Z could regulate the expression of genes associated with CC at enhancers level.

To explore whether H2A.Z promotes the expression of genes near detected enhancers, we evaluated the expression of genes associated with enhancers enriched with H2A.Z by RNA-seq, showing a significant increase in such genes (Wilcoxon test; *P*-value<2.22e-16; Figure

A

Pathways of genes with similar expression pattern
to H2AZ1 and H2AZ2

B

ENCODE and ChEA Consensus TFs of
genes with similar expression pattern
to H2AZ1 and H2AZ2

C

The increase of H2A.Z in cervical cancer regulates
the expression of genes associated with:

Cell cycle
DNA replication
M phase pathway
Messenger RNA processing
S phase
EGF/EGFR signaling pathway
Gene expression
Protein metabolism

Figure 7. H2A.Z.1 and H2A.Z.2 associate with genes related to proliferation in cervical cancer
(A) Functional annotation of genes with similar expression pattern to H2A.Z.1 and H2A.Z.2 in CC. The top of the most significant pathways is shown. We considered an adjusted *P*-value<0.05 as statistically significant. (B) TFs associated with genes with similar expression patterns to H2A.Z.1 and H2A.Z.2 in CC. The top of the most significant TFs is shown. We considered an adjusted *P*-value<0.05 as statistically significant. (C) Proposed model for the role of H2A.Z in CC. H2A.Z is enriched at promoters and enhancers in HeLa-S3. H2A.Z is associated with the recruitment of RNA Pol II, NRF1, NFYA, and other TFs at promoters to allow transcription. In enhancers, H2A.Z is associated with the transcription of near genes. Finally, H2A.Z.1 and H2A.Z.2 isoforms are associated with expression of cancer-associated genes in CC

6F). Pathway analysis showed that genes associated with H2A.Z enhancers regulate pathways associated with carcinogenesis such as gene expression, metabolism, cell cycle, disease, among others (Figure 6G). Taken together, these results suggest that H2A.Z enrichment at enhancer regions could promote CC progression by increasing the expression of genes associated with the transformation process.

H2A.Z.1 and H2A.Z.2 associate with genes related to proliferation in cervical cancer

To investigate the biological effects associated with the increase of H2A.Z.1 and H2A.Z.2 isoforms in CC patients, we evaluated genes with similar expression patterns to both isoforms. Interestingly, we found that genes associated with both isoforms regulate common processes such as DNA replication, cell cycle, chromosome maintenance, and M phase pathway (Figure 7A), which are involved in oncogenic processes (58, 59). Furthermore, these genes are also associated with regulation exerted by oncogenic TFs, such as E2F4, E2F6, NFYB, and MYC (Figure 7B) (60).

Additionally, we evaluated the differentially expressed genes between CC patients with low expression (top 20) and high expression (top 20) of both H2A.Z isoforms using TCGA data. We identified 765 genes associated with H2A.Z.1 overexpression (high H2A.Z.1) and 1,518 genes associated with low levels (low H2A.Z.1) in CC (adjusted *P*-value<0.05; FC of 1.5; supplementary Figure 3A and supplementary Table 3). In addition, we identified 1,446 genes associated with high H2A.Z.2 and 1,552 genes associated with low H2A.Z.2 in CC (adjusted *P*-value<0.05; FC of 1.5; supplementary Figure 3A and supplementary Table 3). GSEA analysis showed enrichment of processes associated with E2F targets, MYC targets, and DNA repair in high H2A.Z.1 and high H2A.Z.2 groups (NOM *P*-value<0.05; supplementary Figures 3B, C).

In summary, these results suggest oncogenic role of H2A.Z.1 and H2A.Z.2 isoforms in CC (Figure 7C), mainly, associated with transcriptional regulation of genes related to oncogenic processes exerted by H2A.Z at the level of promoters and enhancers (Figure 7C).

Discussion
CC is a global health problem for which new diagnostic

and therapeutic tools are needed to improve the quality of life of patients. For these reasons, epigenomic approaches may be required to understand the molecular mechanisms involved in cancer progression, which finally allows the identification of potential clinical tools (61).

Although the presence of HR-HPV infection is considered a requisite for CC development, several reports have shown that epigenetic alterations facilitate the carcinogenic process through transcriptional regulation of cancer-associated genes (61). In CC, alterations in the mechanisms of DNA methylation, histone post-translational modification, and non-coding RNAs have been reported (5), suggesting that different epigenetic mechanisms may be associated with cervical carcinogenesis such as histone variants. In this study, we found overexpression of both H2A.Z.1 and H2A.Z.2 isoforms in CC, increased levels of which are associated with progression and metastasis. Importantly, we show that H2A.Z is enriched at promoter and enhancer regions, associated with gene expression in HeLa-S3 cells. Nonetheless, association of HPV and regulation of H2A.Z remain to be clarified in CC, since there are no current data demonstrating such a relationship.

When analyzing expression data from CC patients, we identified an increase in the expression of H2A.Z.1 and H2A.Z.2 isoforms, which was associated with the stages of progression and nodal metastasis. These data agree with those reported by Yang *et al.*, where H2A.Z overexpression was associated with tumor stage, lymph node and metastasis in intrahepatic cholangiocarcinoma (13). Furthermore, Svotelis *et al.* also identified a correlation between high levels of H2A.Z and high-grade breast cancer (11), therefore our results suggest oncogenic role of H2A.Z in CC.

Few studies have addressed the molecular mechanisms responsible for H2A.Z overexpression. We identified an enrichment of AP2α and ELK1 at H2AZ1 and H2AZ2 promoters in HeLa-S3 cells that might facilitate their overexpression. Previous studies have demonstrated the oncogenic role exerted by AP2α and ELK1 in colorectal and bladder cancer (62-64). Specifically, AP2α and ELK1 form a regulatory network that facilitates the SIRPα expression in tumor-associated macrophages, which was associated with poor survival in colorectal cancer (64), suggesting that AP2α and ELK1 could promote the transcription of both H2A.Z isoforms. To date, the role of AP2α and ELK1 on gene expression in CC has not been demonstrated, thus its exploration in CC is required.

DNA methylation has been extensively studied in CC (65-66). We identified hypomethylation at the H2AZ2 promoter. Hypomethylation often occurs at promoter regions, allowing an increase in gene expression (66). On the other hand, we identified a copy number gain of H2AZ1 and H2AZ2 genes in CC. It is well known that an increment in copy number is associated with carcinogenesis (67), thus H2AZ gain can partially explain high transcription levels. Likewise, Vardabasso *et al.* identified an increase in the copy number of H2AZ1 and H2AZ2 in melanoma patients, which was associated with the overexpression of both isoforms (10). Together, our data suggest that the increase in H2A.Z.1 and H2A.Z.2 isoforms can be regulated by mechanisms associated

with TFs, as well as genetic and epigenetic alterations in CC.

Interestingly, we identified H2A.Z enrichment at promoters and enhancers of the CC cell line HeLa-S3. Our data show evidence to suggest that H2A.Z facilitates the recruitment of RNA Pol II at promoter and enhancer regions, inducing their activation. Specifically, at promoters, H2A.Z is associated with high RNA Pol II enrichment and expression of genes associated with the proliferative phenotype in HeLa-S3 cells. It has been previously reported that the presence of H2A.Z in promoters allows the expression of proliferation-related genes in bladder cancer (14). Specifically, H2A.Z is associated with high levels of H3K4me2/3 around the TSS. In addition, it coincides with recruitment of WDR5 and BPTF favoring gene expression (14). This suggests that H2A.Z promotes TF recruitment to promoters in HeLa-S3 cells.

Several studies have shown that H2A.Z allows TFs recruitment (56). Here we demonstrate that enrichment of H2A.Z at promoter regions is associated with the presence of TFs involved with transcriptional misregulation in cancer. Interestingly, a high H2A.Z enrichment is associated with high levels of NFYA and NRF1 at promoters which in turn have been associated with transcriptional alterations in CC (68-70). NFYA a trimetric transcription factor has a dual role as an activator and a repressor of transcription (71). It has been demonstrated that NFYA allows the expression of SOX2 in CC stem cells, being an important molecule for the maintenance of these cells (68). The regulatory network between H2A.Z and TFs suggested possible hub TFs and critical pathways for cervical carcinogenesis. However, an experimental approach is needed to verify the cooperation between H2A.Z and TFs such as SP2, ZNF384, YY1, NRF1, and NFYA in the transcriptional misregulation in cancer.

On the other hand, we identified that H2A.Z is associated with high expression of genes near enhancers in HeLa-S3, allowing the activation of pathways such as gene expression, metabolism, cell cycle, disease, etc. Previous studies have shown that the presence of H2A.Z at enhancers is associated with chromatin accessibility, DNA hypomethylation, RNA Pol II recruitment, and RAD21-dependent enhancer RNA transcription (72). In prostate cancer, the incorporation of H2A.Z is a requisite for activation of androgen receptor-associated enhancers. This enrichment allows the formation of the nucleosome-free region and the transcription of enhancer RNAs (12). Taken together, these results demonstrate that H2A.Z has a pro-oncogenic role in CC by regulating transcription involving promoters and enhancers.

With these results, we can propose an oncogenic mechanism associated with H2A.Z overexpression in CC, supported by H2A.Z binding to DNA regulatory elements, which in turn promotes a pro-oncogenic transcriptome, allowing the activation of genes associated with cell cycle, DNA replication, and gene expression in CC patients. Therefore, we can consider H2A.Z a potential therapeutic target for CC. For instance, in pancreatic cancer and intrahepatic cholangiocarcinoma, reduction of H2A.Z sensitizes cancer cells to chemotherapy (13, 15). Specifically, in pancreatic cancer, the reduction of

H2A.Z isoforms promotes sensitivity to gemcitabine chemotherapy as well as reduction of tumorigenic processess (15). Hence, it is necessary to investigate the role of H2A.Z in proliferation, DNA replication, migration, and invasion through loss-function assays in CC cell models to further support the oncogenic role of this histone variant.

Conclusion

Our bioinformatic analysis shows solid evidence to propose oncogenic role of H2A.Z in CC, by regulating the expression of cancer-associated genes at promoter and enhancer levels, as well as its association with TFs such as RNA Pol II, NFYA, and NRF1.

Acknowledgment

Authors' Contributions

EGSB and ALPA Study conception and design; EGSB and ALPA Data curation and bioinformatic analysis; EGSB, ALPA, DLA, and LJCM Writing original draft; JMM, MALV, and AEZG Writing, reviewing, and editing. EGSB, AEZG, DLA, LJCM, MALV, JMM, and ALPA Final approval of the version to be published.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

1. Arbyn M, Weiderpass E, Bruni L, de Sanjose S, Saraiya M, Ferlay J, *et al.* Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Health 2020; 8:e191-e203.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68:394-424.
3. Benard VB, Watson M, Saraiya M, Harewood R, Townsend JS, Stroup AM, *et al.* Cervical cancer survival in the United States by race and stage (2001-2009): Findings from the CONCORD-2 study. Cancer 2017; 123 Suppl 24:5119-5137.
4. Munoz N, Castellsague X, de Gonzalez AB, Gissmann L. Chapter 1: HPV in the etiology of human cancer. Vaccine 2006; 24 Suppl 3:S3/1-10.
5. Fang J, Zhang H, Jin S. Epigenetics and cervical cancer: from pathogenesis to therapy. Tumor Biology 2014; 35:5083-5093.
6. Darwiche N. Epigenetic mechanisms and the hallmarks of cancer: An intimate affair. Am J Cancer Res 2020; 10:1954-1978.
7. Fang J, Zhang H, Jin S. Epigenetics and cervical cancer: from pathogenesis to therapy. Tumour Biol 2014; 35:5083-5093.
8. Vardabasso C, Hasson D, Ratnakumar K, Chung CY, Duarte LF, Bernstein E. Histone variants: emerging players in cancer biology. Cell Mol Life Sci 2014; 71:379-404.
9. Giaimo BD, Ferrante F, Herchenröther A, Hake SB, Borggrefe T. The histone variant H2A.Z in gene regulation. Epigenetics & Chromatin 2019; 12:1-22.
10. Vardabasso C, Gaspar-Maia A, Hasson D, Punzeler S, Valle-Garcia D, Straub T, *et al.* Histone Variant H2A.Z.2 Mediates Proliferation and Drug Sensitivity of Malignant Melanoma. Mol Cell 2015; 59:75-88.
11. Svotelis A, Gevry N, Grondin G, Gaudreau L. H2A.Z overexpression promotes cellular proliferation of breast cancer cells. Cell Cycle 2010; 9:364-370.
12. Valdés-Mora F, Gould CM, Colino-Sanguino Y, Qu W, Song JZ, Taylor KM, et al. Acetylated histone variant H2A.Z is involved in the activation of neo-enhancers in prostate cancer. Nat Commun 2017; 8:1346.
13. Yang B, Tong R, Liu H, Wu J, Chen D, Xue Z, *et al.* H2A.Z regulates tumorigenesis, metastasis and sensitivity to cisplatin in intrahepatic cholangiocarcinoma. Int J Oncol 2018; 52:1235-1245.
14. Kim K, Punj V, Choi J, Heo K, Kim J-M, Laird PW, *et al.* Gene dysregulation by histone variant H2A.Z in bladder cancer. Epigenetics Chromatin 2013; 6:34.
15. Avila-Lopez PA, Guerrero G, Nunez-Martinez HN, Peralta-Alvarez CA, Hernandez-Montes G, Alvarez-Hilario LG, *et al.* H2A.Z overexpression suppresses senescence and chemosensitivity in pancreatic ductal adenocarcinoma. Oncogene 2021; 40:2065-2080.
16. Zlatanova J, Thakar A. H2A.Z: view from the top. Structure 2008; 16:166-179.
17. Matsuda R, Hori T, Kitamura H, Takeuchi K, Fukagawa T, Harata M. Identification and characterization of the two isoforms of the vertebrate H2A.Z histone variant. Nucleic Acids Res 2010; 38:4263-4273.
18. Henikoff S, Smith MM. Histone variants and epigenetics. Cold Spring Harb Perspect Biol 2015; 7:a019364.
19. Dryhurst D, Ishibashi T, Rose KL, Eirín-López JM, McDonald D, Silva-Moreno B, *et al.* Characterization of the histone H2A.Z-1 and H2A.Z-2 isoforms in vertebrates. BMC Biol 2009; 7:86.
20. Colwill K, Graslund S. A roadmap to generate renewable protein binders to the human proteome. Nat Methods 2011; 8:551-558.
21. Tang Z, Li C, Kang B, Gao G, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45:W98-W102.
22. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, *et al.* NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res 2013; 41:D991-995.
23. Scotto L, Narayan G, Nandula SV, Arias-Pulido H, Subramaniyam S, Schneider A, *et al.* Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. Genes Chromosomes Cancer 2008; 47:755-765.
24. Zhai Y, Kuick R, Nan B, Ota I, Weiss SJ, Trimble CL, *et al.* Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. Cancer Res 2007; 67:10163-10172.
25. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, *et al.* UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19:649-658.

26. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, *et al.* The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2:401-404.

27. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6:pl1.

28. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, *et al.* ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 2012; 40:W597-603.

29. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, *et al.* Integrative genomics viewer. Nat Biotechnol 2011; 29:24-26.

30. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics 2002; 18:1427-1431.

31. Xiong Y, Wei Y, Gu Y, Zhang S, Lyu J, Zhang B, *et al.* DiseaseMeth version 2.0: a major expansion and update of the human disease methylation database. Nucleic Acids Res 2017; 45:D888-D895.

32. An integrated encyclopedia of DNA elements in the human genome. Nature 2012; 489:57-74.

33. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, *et al.* The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res 2018; 46:D794-D801.

34. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 2018; 46:W537-W544.

35. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009; 10:R25.

36. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25:2078-2079.

37. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, *et al.* Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008; 9:R137.

38. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 2016; 44:W160-165.

39. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 2015; 31:2382-2383.

40. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015; 12:357-360.

41. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014; 30:923-930.

42. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013; 45:1113-1120.

43. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, *et al.* TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res 2016; 44:e71.

44. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15:550-570.

45. Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics 2010; 26:2438-2444.

46. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016; 44:W90-97.

47. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, *et al.* MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 2009; 37:W202-208.

48. Bailey TL. DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics 2011; 27:1653-1659.

49. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, *et al.* STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research 2018; 47:D607-D613.

50. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, *et al.* Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol 2015; 16:22-34.

51. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, *et al.* GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010; 28:495-501.

52. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102:15545-15550.

53. Wickham H. ggplot2-Elegant Graphics for Data Analysis. Springer International Publishing. Cham, Switzerland 2016.

54. Thakar A, Gupta P, Ishibashi T, Finn R, Silva-Moreno B, Uchiyama S, *et al.* H2A.Z and H3.3 histone variants affect nucleosome structure: biochemical and biophysical studies. Biochemistry 2009; 48:10852-10857.

55. Ni Z, Olsen JB, Guo X, Zhong G, Ruan ED, Marcon E, *et al.* Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B. Transcription 2011; 2:237-242.

56. Subramanian V, Fields PA, Boyer LA. H2A.Z: A molecular rheostat for transcriptional control. F1000Prime Rep 2015; 7:01.

57. Peng Y, Zhang Y. Enhancer and super-enhancer: Positive regulators in gene transcription. Animal Model Exp Med 2018; 1:169-179.

58. Yu S, Wang G, Shi Y, Xu H, Zheng Y, Chen Y. MCMs in Cancer: Prognostic Potential and Mechanisms. Anal Cell Pathol 2020; 2020:3750294.

59. Boyer AS, Walter D, Sorensen CS. DNA replication and cancer: From dysfunctional replication origin activities to therapeutic opportunities. Semin Cancer Biol 2016; 37-38:16-25.

60. Nebert DW. Transcription factors and cancer: an overview. Toxicology 2002; 181-182:131-141.

61. Kanwal R, Gupta K, Gupta S. Cancer epigenetics: an introduction. Methods Mol Biol 2015; 1238:3-25.

62. Stabach PR, Thiyagarajan MM, Woodfield GW, Weigel RJ. AP2α alters the transcriptional activity and stability of p53. Oncogene 2006; 25:2148-2159.

63. Kawahara T, Shareef HK, Aljarah AK, Ide H, Li Y, Kashiwagi E, *et al.* ELK1 is up-regulated by androgen in bladder cancer cells and promotes tumor progression. Oncotarget 2015; 6:29860-29876.

64. Wang X, Luo X, Chen C, Tang Y, Li L, Mo B, *et al.* The Ap-2alpha/Elk-1 axis regulates Sirpalpha-dependent tumor phagocytosis by tumor-associated macrophages in colorectal cancer. Signal Transduct Target Ther 2020; 5:35.

65. Yang HJ. Aberrant DNA methylation in cervical carcinogenesis. Chin J Cancer 2013; 32:42-48.

66. Feng C, Dong J, Chang W, Cui M, Xu T. The progress of methylation regulation in gene expression of cervical cancer. Int J Genomics 2018; 2018:8260652.

67. Shao X, Lv N, Liao J, Long J, Xue R, Ai N, *et al.* Copy number variation is highly correlated with differential gene expression: A pan-cancer study. BMC Med Genet 2019; 20:175.

68. Yang WT, Zhao ZX, Li B, Zheng PS. NF-YA transcriptionally activates the expression of SOX2 in cervical cancer stem cells. PLoS One 2019; 14:e0215494.

69. Fischer M, Quaas M, Wintsche A, Müller GA, Engeland K. Polo-like kinase 4 transcription is activated via CRE and NRF1 elements, repressed by DREAM through CDE/CHR sites and deregulated by HPV E7 protein. Nucleic Acids Research 2013; 42:163-180.

70. Johannsen E, Lambert PF. Epigenetics of human papillomaviruses. Virology 2013; 445:205-212.

71. Serra E, Zemzoumi K, di Silvio A, Mantovani R, Lardans V, Dissous C. Conservation and divergence of NF-Y transcriptional activation function. Nucleic Acids Res 1998; 26:3800-3805.

72. Brunelle M, Nordell Markovits A, Rodrigue S, Lupien M, Jacques PE, Gevry N. The histone variant H2A.Z is an important regulator of enhancer activity. Nucleic Acids Res 2015; 43:9742-9756.