

Expression signatures of HOX cluster genes in cervical cancer pathogenesis: Impact of human papillomavirus type 16 oncoprotein E7

Sweta Sharma Saha¹, Rahul Roy Chowdhury², Nidhu Ranjan Mondal², Sudipta Roy³ and Sharmila Sengupta¹

¹National Institute of Biomedical Genomics, Kolkata, West Bengal, India

²Department of Gynecology, Saroj Gupta Cancer Centre and Research Institute, Kolkata, West Bengal, India

³Sri Aurobindo Seva Kendra, Kolkata, West Bengal, India

Correspondence to: Sharmila Sengupta, **email:** ssg1@nibmg.ac.in, sharmilasg@gmail.com

Keywords: cervical cancer, human papillomavirus, HOX genes, E7, chromatin marks

Received: January 16, 2017

Accepted: March 14, 2017

Published: March 28, 2017

Copyright: Saha et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

The Homeobox (HOX) genes encode important transcription factors showing deregulated expression in several cancers. However, their role in cervical cancer pathogenesis, remains largely unexplored. Herein, we studied their association with Human Papillomavirus type 16 (HPV16) mediated cervical cancers. Our previously published gene expression microarray data revealed a significant alteration of 12 out of 39 HOX cluster members among cervical cancer cases, in comparison to the histopathologically normal controls. Of these, we validated seven (HOXA10, HOXA13, HOXB13, HOXC8, HOXC9, HOXC11 and HOXD10) by quantitative real-time PCR. We identified decreased HOXA10 expression as opposed to the increased expression of the rest. Such decrease was independent of the integration status of HPV16 genome, but correlated negatively with E7 expression in clinical samples, that was confirmed *in vitro*. HOXA10 and HOXB13 revealed association with Epithelial-Mesenchymal Transition (EMT). While HOXA10 expression correlated positively with E-Cadherin and negatively with Vimentin expression, HOXB13 showed the reverse trend. Chromatin immunoprecipitation study *in vitro* revealed the ability of E7 to increase HOX gene expression by epigenetic regulation, affecting the H3K4me3 and H3K27me3 status of their promoters, resulting from a loss of PRC2-LSD1 complex activity. Thus, besides identifying the deregulated expression of HOX cluster members in HPV16 positive cervical cancer and their association with EMT, our study highlighted the mechanism of HPV16 E7-mediated epigenetic regulation of HOX genes in such cancers, potentially serving as bedrock for functional studies in the future.

INTRODUCTION

Cervical cancer is the fourth leading cancer among women globally with HPV16 infection alone associating with ~50% of the cases [1-5]. HPV16 E7 is one of the driver molecules responsible for HPV-induced cellular transformation [6]. The ability of HPV16 E7 to interact with the host-encoded proteins has been well-documented [7-8]. Such functional interactions, specifically with transcription factors, cumulatively influence the transcriptional process and global

gene expression creating an environment conducive to HPV-mediated cervical cancer pathogenesis. However, in a recent report from our laboratory [9] we proposed a mechanism of E7-mediated increase of gene expression globally, by affecting lncRNA HOTAIR expression and function. Such functional inactivation of HOTAIR, in turn, could impair Polycomb Repressive Complex 2 (PRC2) and Lysine Specific Demethylase 1(LSD1) recruitment resulting in concomitant loss of global gene silencing H3K27me3 marks and gain of gene activating H3K4me3 marks, respectively. However, such proposition needs experimental validation.

HOX genes are a family of transcription factors that show deregulated expression in a variety of cancers and known as targets of PRC2-complex during embryonic development [10-11]. The class I HOX genes comprise of 39 members distributed across four clusters localized in four different chromosomes - HOXA at 7p15.3, HOXB at 17p21.3, HOXC at 12q13.3, and HOXD at 2q31[12-14]. Although many of the downstream targets of HOX genes are still not fully defined, it has been shown that HOX genes are integral to normal temporo-spatial limb [15] and organ [16] development along the anterior-posterior (A-P) axis.

Since HOX cluster members function as transcription factors, the common thought was that these show enhanced expression in cancers driving carcinogenesis by activating the downstream signalling cascades. However, several studies reported reduced expression of HOX cluster encoded molecules. Aberrant expression of HOX cluster genes is integral to a network of regulatory mechanisms involved in normal adult tissue homeostasis, and maintenance and activation of stem cell self-renewal process, crucial to malignant transformation [17-20]. The expression of genes belonging to the HOXC family, particularly HOXC4 to HOXC9, HOXC11, and HOXC13, shows predominant increase in most of the solid tumour types, except ovarian cancers. The two HOX genes most commonly activated in solid tumours are HOXA9 and HOXB13 [21]. Such activation of HOX cluster members are closely linked to the process of Epithelial-Mesenchymal Transition (EMT) [22-24]. However, studies on expression profiling of HOX cluster genes in cervical cancers are mostly based on cell lines [25-28], with only a couple of studies employing global gene expression profiling of cervical tissues [29-30]. However, none of these studies on cervical cancer cell lines or tissues, have taken into account the presence of HPV or expression of viral oncogenes. McLaughlin-Drubin et al., in 2011 [31], highlighted the increased expression of the PRC2-target HOX cluster genes in the presence of E7. This is concomitant with global loss of chromatin suppressive H3K27me3, despite the increased EZH2 expression, under the influence of E7 [31]. However, the impact of E7 on the gene activating H3K4me3 and gene silencing H3K27me3 marks, specifically at the HOX promoters, has not been studied thus far.

Therefore, in this study, we focussed on elucidating the expression profile of HOX cluster member genes associated with cervical cancer pathogenesis, and their probable association with the Epithelial Mesenchymal Transition (EMT). We also aimed at identifying the association of HOX cluster expression with viral factors like HPV16 integration status and E7 expression among the cervical cancer cases. Finally, we elucidated the role of HPV16 E7 in epigenetically affecting the expression of HOX cluster genes through concomitant functional abrogation of PRC2-LSD1 complex in cervical cancer cell lines. Towards this, we estimated the H3K4me3 and H3K27me3 status of the promoters of the HOX genes with altered expression.

RESULTS

HOX cluster genes show significant expression deregulation among cervical cancer samples

To identify the HOX cluster members with deregulated expression among HPV16 positive cervical cancer cases in comparison to the histopathologically normal control samples (HPV negative and HPV16 positive), microarray data generated on 20 HPV16 positive cervical cancer cases, 11 HPV16 positive histopathologically normal samples and 11 HPV negative histopathologically normal samples (Accession No: GSE67522) was used. The histopathologically normal samples were grouped together for comparison, as the HPV negative and HPV16 positive normal samples did not reveal any differentially expressed genes as described in our previous study [9].

Out of 39 HOX cluster genes, the microarray analysis revealed expression levels of 31 genes. Of these, 12 members of the HOX cluster portrayed significant differential expression among the cervical cancer cases, as compared to the histopathologically normal controls (Figure 1; Supplementary Table 1).

Validation of expression changes of seven deregulated HOX cluster members

We validated the expression profile of seven out of the twelve deregulated HOX cluster members, HOXA10, HOXA13, HOXB13, HOXC8, HOXC9, HOXC11 and HOXD10, using SYBR green based quantitative real-time PCR. For such analysis, we used a sample set of seventy cervical cancer cases (44 samples harbouring episomal HPV16 and 26 samples harbouring integrated HPV16 genomes) and thirty histopathologically normal cervical samples (11 HPV16 positive and 19 HPV negative). Such analysis, confirmed significant altered expression of all seven HOX cluster members among the cervical cancer cases, in comparison with the HPV negative control samples (Supplementary Table 2). Of these, upregulation of HOXD10 has been previously reported from our laboratory [9]. HOXA10 was the only member showing significantly reduced expression, as opposed to the rest revealing significant increase in expression (Figure 2). However, none of these HOX cluster members showed differential expression between HPV16 positive non-malignant samples and HPV negative control samples, which was concordant with our observations using microarray analysis. The cervical cancer cases harbouring episomal and integrated HPV16 also failed to show significant differences in the expression levels of all the seven HOX cluster genes. This was indicative of the fact that the expression of HOX cluster genes was independent of the physical status of the HPV16 genome.

Expression of HOXA10 and HOXB13 correlates with expression of EMT markers, E-Cadherin and Vimentin

Deregulated expression of several HOX genes in cervical cancer cases prompted us to interrogate the impact of such changes in cervical cancer pathogenesis. Existing reports [21-24], highlight the role of some HOX cluster members in driving metastasis, by the induction of Epithelial Mesenchymal Transition (EMT). We therefore determined the expression of EMT markers, E-Cadherin and Vimentin, by quantitative real-time PCR, among the same set of samples used for the validation of the expression levels of the HOX cluster members.

E-Cadherin expression was significantly decreased among the cervical cancer cases (9.85-fold; $p=0.024$)

as compared to the HPV negative control samples. The cervical cancer samples with episomal HPV16 genome showed a reduction by 11.79-fold ($p=0.03$), while the cervical cancer samples with integrated HPV16 genomes showed a reduction by 6.02-fold ($p=0.04$) (Figure 3A). However, such difference between the two cervical cancer subtypes was not statistically significant. Vimentin, on the other hand, showed a significantly increased expression among the cervical cancer cases (3.83-fold; $p=0.001$), as compared to the HPV negative control samples. The cervical cancer samples with episomal and integrated HPV16 genomes showed significantly increased expression of Vimentin by 3.66-fold ($p<0.001$) and 5.44-fold ($p=0.014$), respectively (Figure 3B). Both E-Cadherin and Vimentin expression failed to differ significantly between the cervical cancer cases with episomal and integrated HPV16 genomes.

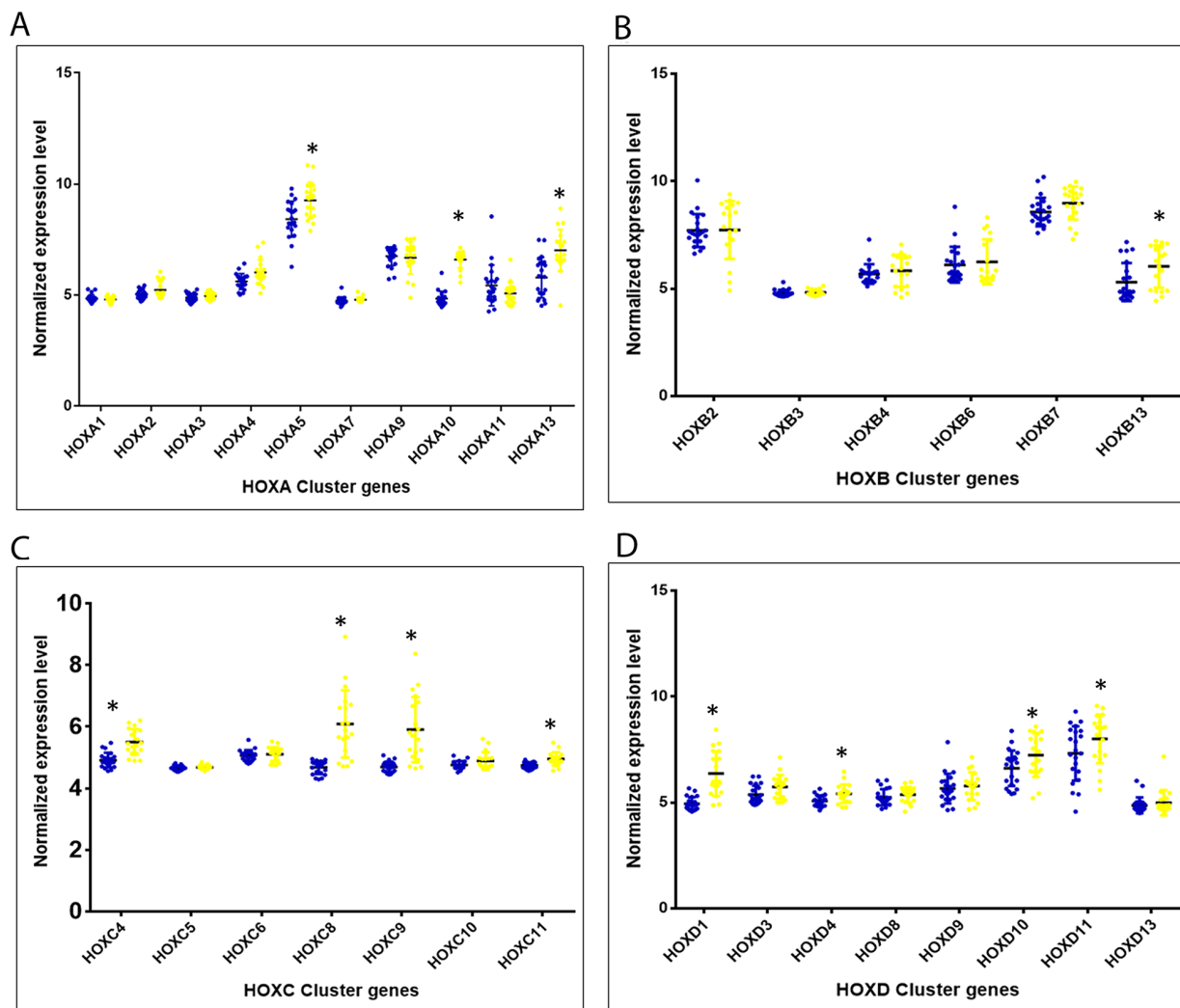


Figure 1: Altered expression of 4 HOX clusters (covered in microarray) among cervical cancers as compared to histopathologically normal samples. (A – D) represent genes belonging to the clusters, HOXA, HOXB, HOXC and HOXD, respectively. (Blue dots: Histopathologically normal controls samples; Yellow dots: HPV16 positive cervical cancer cases; * Statistically significantly up-regulated among cervical cancer cases with $p<0.05$).

Expression of both the EMT markers, E-Cadherin and Vimentin, showed correlation with the expression of the HOX cluster members, HOXA10 and HOXB13 (Supplementary Tables 3 and 4). While HOXA10 showed a positive correlation with E-Cadherin (Pearson's

Correlation Coefficient=0.645; p=0.0017) and a negative correlation with Vimentin (Pearson's Correlation Coefficient=-0.456; p=0.014), the reverse trend was observed for HOXB13. HOXB13 showed a negative correlation with E-Cadherin (Pearson's Correlation

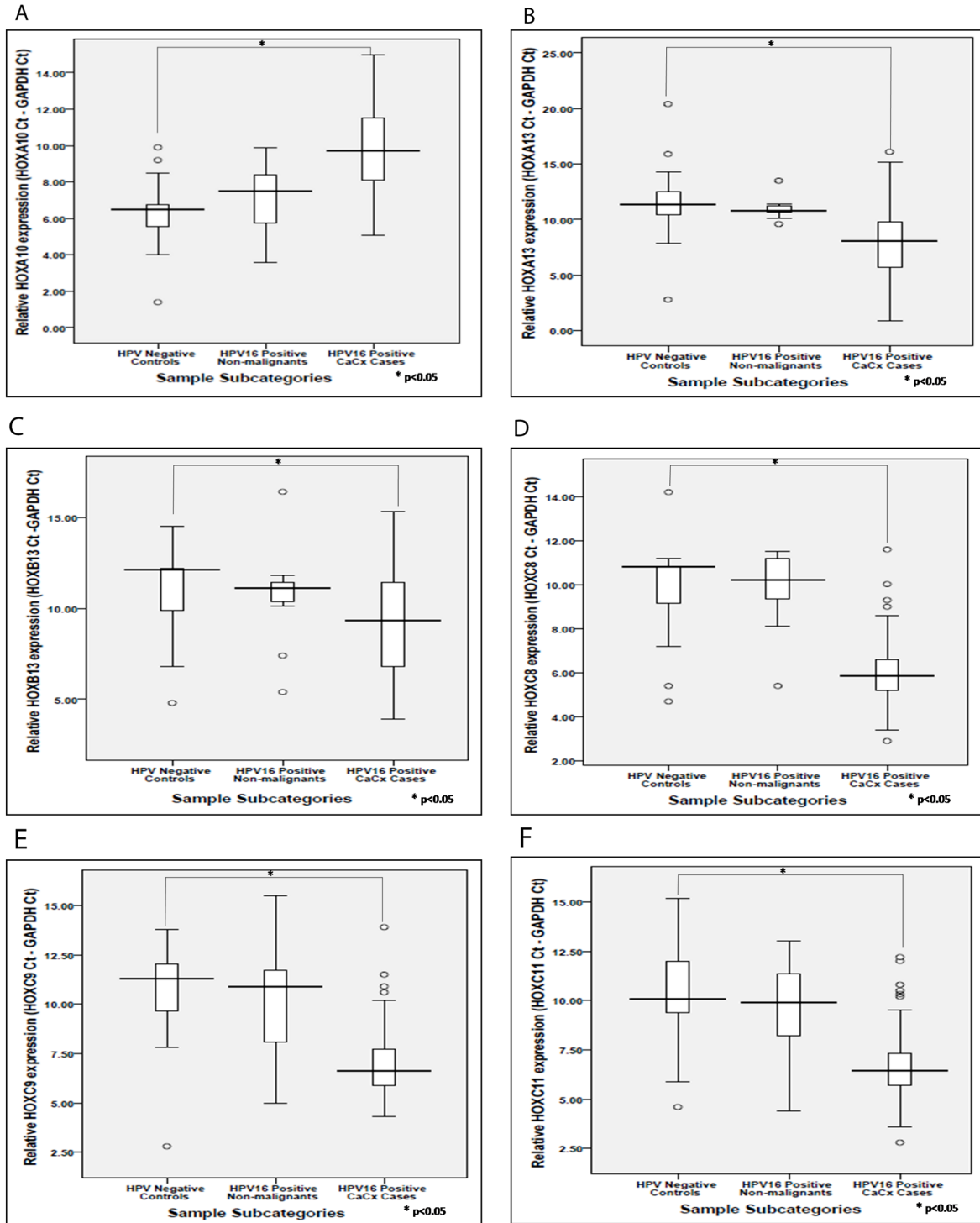


Figure 2: Box plots representing distribution of the expression levels of the HOX cluster genes based on real time PCR analysis. (A – F) corresponds to HOX cluster members such as HOXA10, HOXA13, HOXB13, HOXC8, HOXC9, HOXC11, respectively, among different categories of cervical samples.

Coefficient= -0.467; $p=0.0014$) and a positive correlation with Vimentin (Pearson's Correlation Coefficient=0.328; $p=0.035$). Thus the correlation plots of HOXA10 and HOXB13 with the EMT markers (Figure 3C–3F), suggest a role of HOXA10 and HOXB13 in driving metastasis in cervical cancer.

HOXA10 and HOXD10 expression correlates with HPV16 E7 expression

Existing reports show enhanced E-Cadherin expression in response to silencing of E7 in HPV16 positive keratinocytes [32]. Thus, the correlation between HOXA10 and HOXB13 and the EMT markers, E-Cadherin and Vimentin, was indicative of an interactive role of E7 and HOX cluster genes. In addition, HPV16 E7 impairs the activity of the Polycomb repressive complex, a known regulator of HOX cluster genes, resulting in decreased global H3K27me3 marks [31]. Such observations prompted us to determine if the expression of HOX cluster genes showed any correlation with HPV16 E7 expression among the case samples. For this, we first considered the microarray based data for the HOX genes showing deregulated expression (Supplementary Table 5). Subsequently, we validated the results employing qRT-PCR data on larger number of samples. Only, HOXA10 and HOXD10 expression showed a statistically significant negative correlation with HPV16 E7 expression (Pearson's correlation coefficients=-0.368 and -0.472, respectively; $p=0.032$ and 0.001 , respectively; Figure 4). We further confirmed the negative correlation between HOXA10 and HPV16 E7 expressions *in vitro*, where HOXA10 showed significantly reduced expression (450-fold, $p<0.001$) in C33A cells transfected with E7 expressing plasmid, as compared to untransfected C33A cells (Supplementary Figure 1).

The increased HOXD10 expression among the cervical cancer cases recorded in our previous report, correlated negatively with lncRNA HOTAIR expression [9]. HOTAIR is known to *trans* regulate HOXD10, through PRC2-LSD1 complex recruitment. However, the reduction of HOTAIR expression among the cervical cancers could be responsible for the loss of such gene silencing and hence HOXD10 upregulation, which demanded further validation.

HPV16 E7 mediated alteration of HOX gene expression and histone methylation (H3K4me3 and H3K27me3) status at the promoters of HOX cluster members

In our previous study, the increased expression of a large number of genes in cervical cancer cases, was explained through probable abrogation of lncRNA HOTAIR function in PRC2-LSD1 complex recruitment and repression of gene expression, under the influence of E7 [9]. Thus, we hypothesized that HPV16 E7 might have

an important role to play in the activation of HOX cluster genes through abrogating PRC2-LSD1 complex activity, despite the fact that some of the HOX cluster members failed to show any significant correlation with HPV16 E7. We therefore aimed to identify the expression levels and status of gene activating (H3K4me3) and gene silencing (H3K27me3) chromatin marks at the promoters of four HOX cluster genes (HOXA13, HOXB13, HOXC11 and HOXD10). We selected these genes in view of their significantly enhanced expression among cervical cancer cases (Supplementary Tables 1 and 2). We observed that all the four members showed increased expression in the HPV negative cervical cancer cell line, C33A, upon transfection with HPV16 E7 expression plasmid (Figure 5; Supplementary Table 6).

The enhanced expression of the four HOX cluster members was concomitant with significant alterations in the gene activating H3K4me3 marks and gene silencing H3K27me3 marks, at their promoter regions (Supplementary Table 7). The promoter regions of all four HOX cluster members, showed an enrichment of H3K4me3 marks with a reduction in H3K27me3 marks in three of the members, HOXA13, HOXB13 and HOXD10 (Figure 6), explaining the increased expression of HOX cluster members. HOXC11 showed a concomitant enrichment of both H3K4me3 and H3K27me3 marks, but a higher fold-enrichment for H3K4me3 marks as compared to H3K27me3 marks, indicating progression towards gene activation. Such observations clearly establish our model of E7-dependent epigenetic regulation of HOX genes, through abrogation of PRC2-LSD1 complex activity.

DISCUSSION

The role of HOX gene deregulation in cancer development remains unexplored not only in cervical cancers but in other cancers as well. Thus, we undertook this study to identify the HOX cluster members with altered expression, specifically in HPV16 positive cervical cancers. We observed significantly increased expression of several HOX cluster members, excepting HOXA10, among the cervical cancer cases, in comparison to the healthy controls. This observation impressed upon us, the possibility that HOX gene expression increase could in turn bring about the activation of several cancer-related pathways, during cervical cancer development. None of the HOX cluster genes with altered expression levels, showed any association with the integration status of the viral genome (episomal/integrated). However, HPV16 E7 seemed to play an important role in the activation of HOX cluster genes despite the fact that some of the HOX cluster members failed to show any significant correlation with HPV16 E7.

The HOX cluster members that revealed altered expression in cervical cancer cases in our study, also show deregulated expression in one or more of the other cancer types. Similar to our observations, HOXA13

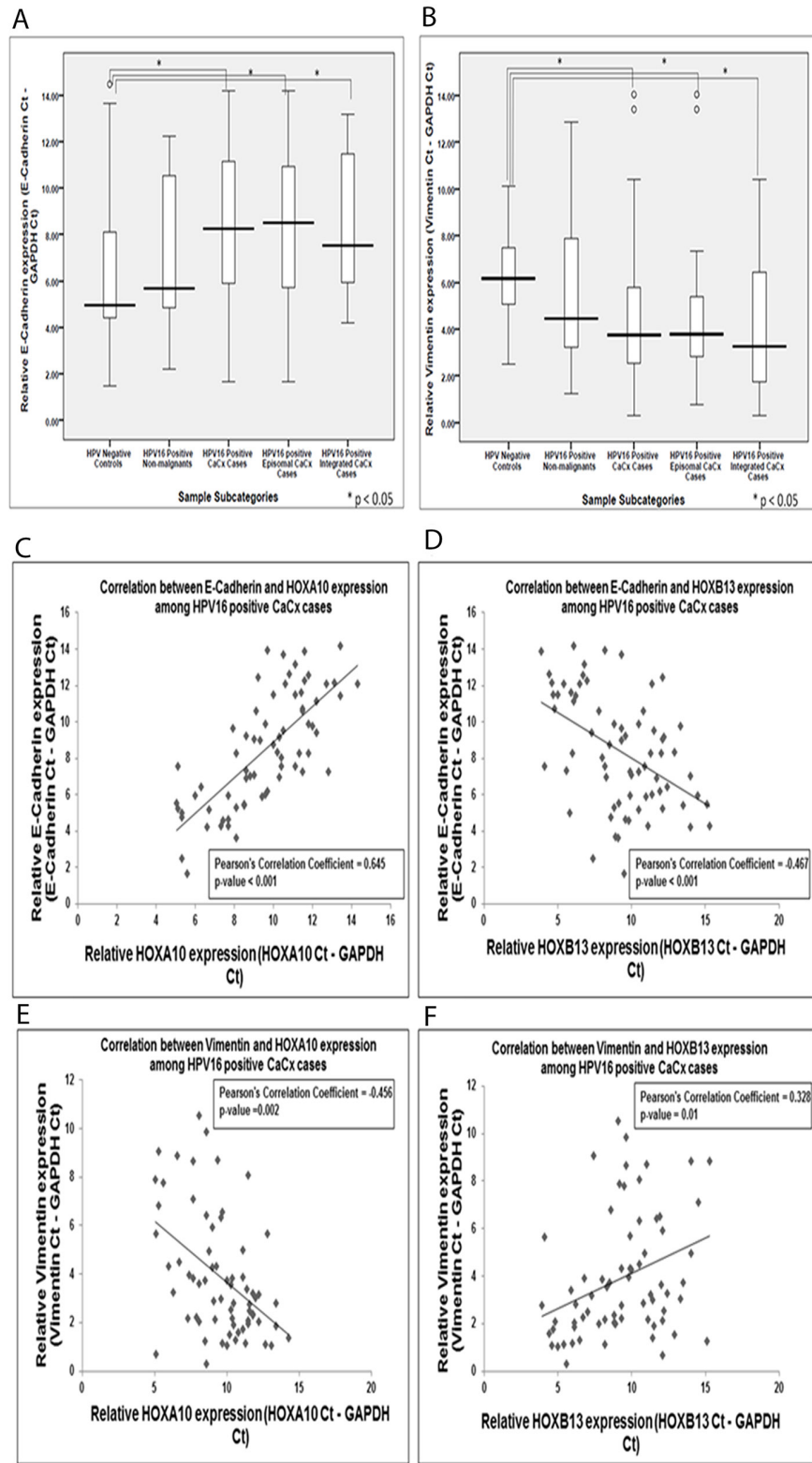


Figure 3: Box plots representing expression levels of the E-Cadherin and Vimentin among the different categories of cervical samples and their correlation with HOXA10 and HOXB13 expression, among cervical cancers. (A) Relative expression of E-Cadherin (E-Cadherin Ct – GAPDH Ct) **(B)** Relative expression of Vimentin (Vimentin Ct – GAPDH Ct) of samples **(C)** Correlation plot depicting significant positive correlation between HOXA10 and E-Cadherin expression **(D)** Correlation plot depicting significant negative correlation between HOXB13 and E-Cadherin expression **(E)** Correlation plot depicting significant positive correlation between HOXA10 and Vimentin expression **(F)** Correlation plot depicting significant negative correlation between HOXB13 and Vimentin expression.

shows enhanced expression in oesophageal squamous cell carcinoma, gastric cancer and hepatocellular carcinoma [33-35]. However, its mode of action in cancer development is still unclear. HOXB13 activation,

on the other hand, has been shown to be associated with metastasis through activation of MMP-9 [23, 36-37], in a variety of solid tumours. The existence of a similar mechanism in cervical cancer demands further

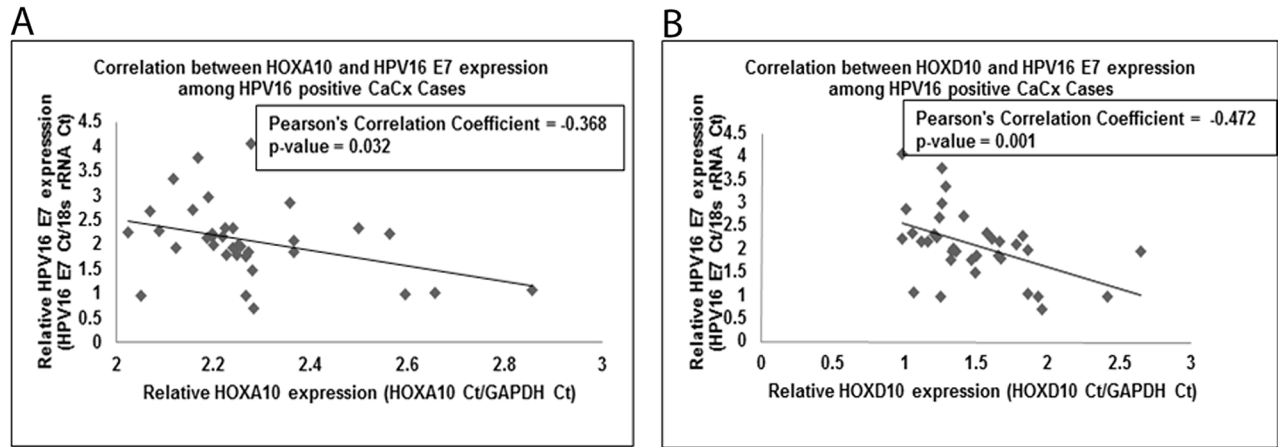


Figure 4: Linear regression analysis for correlation of (A) HOXA10 and (B) HOXD10, with HPV16 E7 expression among cervical cancers.

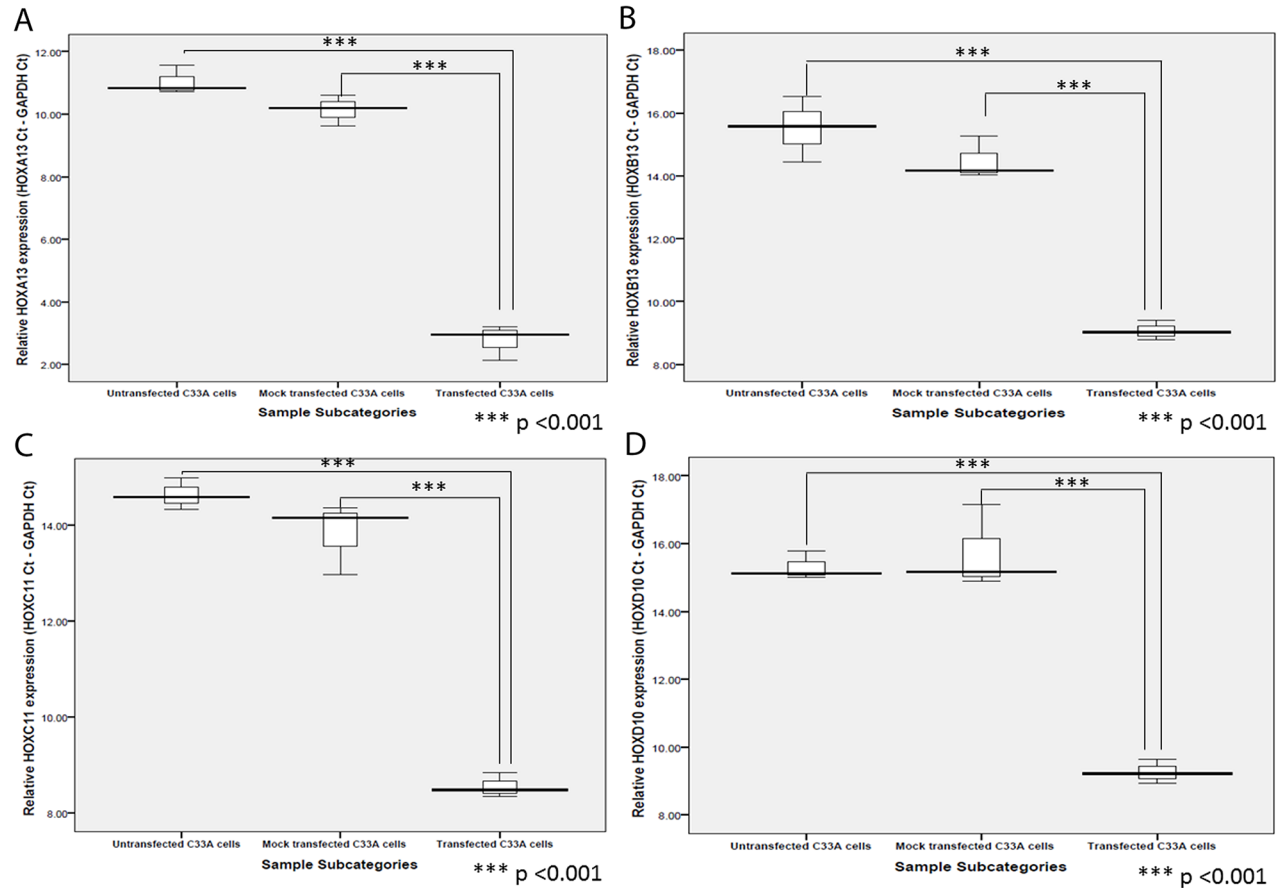


Figure 5: Box plots representing expression of (A) HOXA13, (B) HOXB13, (C) HOXC11, and (D) HOXD10, in C33A cells post-transfection of pcDNA3.1-HPV16 E7 vector in comparison to the untransfected and mock (empty vector) transfected C33A cells.

investigations. Studies on HOXC cluster members have also identified increased expression of HOXC8, HOXC9 and HOXC11, as observed by us in this study. HOXC8 correlates negatively with tumour growth [37] among pancreatic ductal adenocarcinomas, while showing increased expression in cervical cancers [38]. Activation of HOXC9 and HOXC11 in various cancers has been associated with cell proliferation and invasion [39-41]. A few studies have also highlighted that HOX cluster members drive metastasis, specifically by inducing Epithelial Mesenchymal Transition (EMT) [22-24]. In our study, the decreased HOXA10 expression and increased HOXB13 expression were associated with reduced E-Cadherin and enhanced Vimentin expression, proposing the need for further studies to establish HOXA10 and HOXB13 as biomarkers of metastasis in cervical cancer.

The increased expression of majority of the HOX cluster members was also in line with our hypothesis of HPV16 E7 driven abrogation of PRC2-LSD1 complex activity leading to loss of gene silencing. A significant correlation of HOXA10 and HOXD10 expression with HPV16 E7 expression highlighted the impact of E7 on HOX gene expression. One of the possible mechanism of HOXA10 reduction could be through HPV16 E7-dependent activation of DNMT1 leading to HOXA10 promoter methylation [22, 41]. On the other hand, in our earlier report, we interpreted the increased HOXD10 expression in cervical cancer cases to be the result of E7-dependent inactivation of HOTAIR mediated gene silencing through PRC2-LSD1 complex [9].

Although E7 expression did not show significant correlation with the expression of majority of the HOX

cluster members among the cervical cancer samples, E7 overexpression in HPV negative cervical cancer cell line resulted in the activation of HOX cluster genes. Such differences in clinical and *in vitro* observations could be due to the heterogeneity of cancer tissues as opposed to the homogenous pool of cells *in vitro*. The gain of H3K4me3 and loss of H3K27me3 marks at the promoters of HOX cluster genes could explain the activated expression of such genes, under the influence of HPV16 E7 expression. McLaughlin-Drubin et al., [31] had earlier shown that E7 expression reduces global H3K27me3 marks and increases the expression of PRC2-target HOX genes in Human Foreskin Keratinocytes. They interpreted this to be the result of transcriptional induction of KDM6A and KDM6B H3K27 specific demethylases, despite increased expression of EZH2 under the influence of E7 that results in the loss of EZH2 activity through Akt pathway mediated phosphorylation of EZH2 [31]. In this study, we clearly observed the concomitant and significant enrichment of H3K4me3 and reduction of H3K27me3 marks at the HOX cluster promoters, despite increased expression of the PRC2 complex member EZH2 as recorded in our previous study (9). Thus taken together, through this study we establish the mechanism of E7-mediated activation of HOX gene expression through functional abrogation of HOTAIR, ultimately suppressing the PRC2-LSD1 complex activity, a model that we proposed in our previous study [9].

In conclusion, along with identifying altered expression of HOX cluster genes and their association with metastasis in HPV16 related cervical cancers, our study also highlighted the master regulatory role of

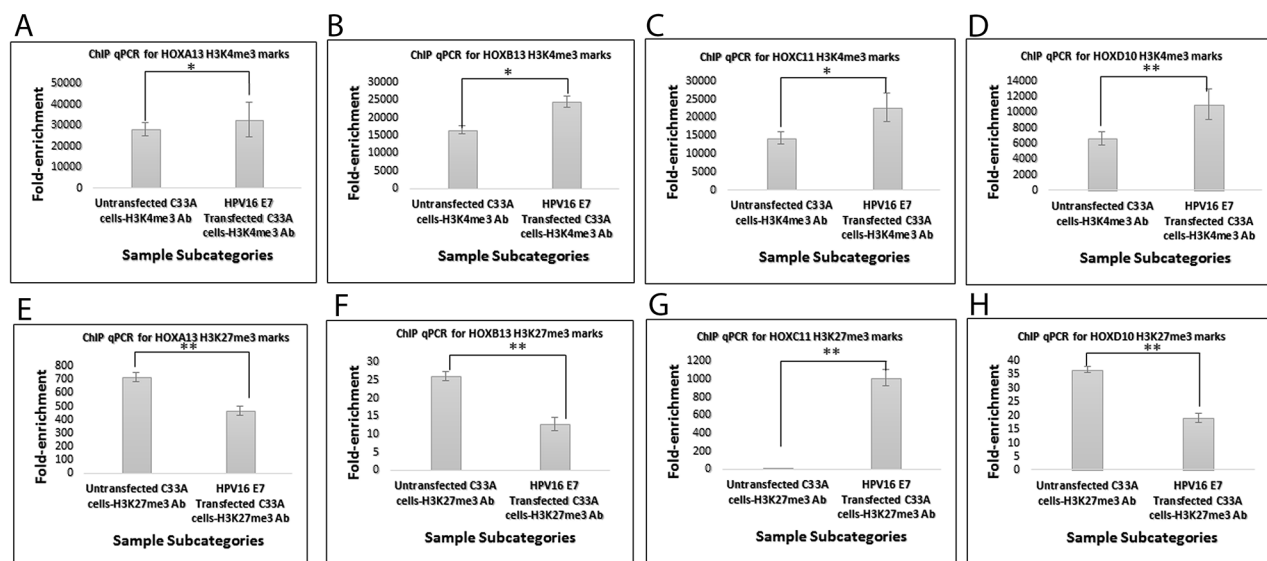


Figure 6: q-PCR based analysis of enrichment of H3K4me3 and H3K27me3 marks at the promoters of HOX cluster genes in HPV negative and E7 expressing C33A cells. Fold-enrichment of H3K4me3 marks at (A) HOXA13, (B) HOXB13, (C) HOXC11, (D) HOXD10 promoters, Fold-enrichment of H3K27me3 marks at (E) HOXA13, (F) HOXB13, (G) HOXC11, (H) HOXD10 promoters. (*p<0.05; **p<0.01; Abbreviations: Ab=Antibody).

HPV16 E7 in modulating the expression of HOX cluster genes. Such expression alteration predominantly involved enrichment of H3K4me3 and loss of H3K27me3 marks at the gene promoters. Thus global studies employing ChIP-Seq for identifying the status of H3K4me3 and H3K27me3 marks at the gene promoters, along with RNA-Seq, seem to be important to identify the concomitant changes in the transcriptome brought about by HPV16 E7. This would further identify and confirm the role of HPV16 E7 in the epigenetic regulation of host cellular transcription. Such studies would also help in identifying molecular targets for therapy of HPV-active cervical cancer, which have distinctly different molecular phenotypes from HPV-inactive cervical cancer, as has been identified in a recent report studying TCGA cervical cancer data [42].

MATERIALS AND METHODS

Clinical samples and HPV16 Status

The healthy control cervical biopsy samples were derived from married women aged 36 – 70 years (median: 52 years) without any prior history of cervical dysplasia/malignancy, from Calcutta Medical College and Hospital, Kolkata, West Bengal, India and were histopathologically confirmed as normal. The Cervical cancer samples used for this study were derived from married subjects aged 34 – 65 years (median: 52 years) attending a cancer referral hospital (Saroj Gupta Cancer Centre and Research Institute, South 24 Parganas, West Bengal, India). The cervical cancer biopsy samples were histopathologically confirmed as invasive squamous cell carcinomas (majority of which were diagnosed as moderately differentiated squamous cell carcinoma pathologically) categorised using the WHO classification and clinically diagnosed as stage III and above as per FIGO classification. Samples included in the different study groups did not differ significantly in terms of their median age.

The samples were collected with informed consent from the participants, approved by the ethical committee for human experimentation of the National Institute of Biomedical Genomics (NIBMG) and frozen post collection. We performed all experiments in accordance with the approved ethical guidelines and regulations. All such clinical samples were tested for their HPV/HPV16 status by PCR based detection as described earlier [9]. Confirmation of the integration or episomal status of HPV16 was done employing APOT-coupled Taqman assay as described previously [43].

Microarray expression analysis of HOX cluster members

The microarray data was generated using Illumina HumanHT-12 v4 Expression Bead Chip array based platform generated by our laboratory [9] and

submitted to the GEO repository (Accession Number: GSE67522, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67522>). This data was used to identify the distinct differences of HOX cluster members at the gene expression level, between the HPV16 positive cervical cancer cases and the histopathologically normal control samples.

RNA isolation, reverse transcription and quantitative real-time PCR

RNA isolation, reverse transcription and quantitative real-time PCRs were all performed, as described previously [9]. The primers used for the expression analysis using qRT-PCR are listed in Supplementary Table 8.

Cell culture and transfection

HPV negative cell line C33A and HPV16 positive cell line SiHa were cultured in DMEM supplemented with 10% FBS, 50 Units/ml of penicillin and 50 µg/ml of streptomycin at 37°C and 5% CO₂. C33A cells were transfected with 1 µg of plasmid pcDNA3.1-HPV16 E7 vector generated in our laboratory [9], using Lipofectamine 2000 reagent according to manufacturer's protocol. The cells were harvested and washed with 1X PBS, pH 7.4, trypsinized and collected by centrifugation at 300 xg for 10 minutes. The transfected cells were used for RNA and protein isolation. The expression of HPV16 E7 was confirmed by Western blot (Supplementary Figure 2; as shown in Ref. 9) The transfection experiments were carried out in three sets, each in triplicates.

Chromatin immunoprecipitation and quantitative real-time PCR

SiHa and C33A cells were cultured as described above. For the analysis of H3K4me3 and H3K27me3 marks, we used SiHa cells and C33A cells, untransfected or transfected (HPV16 E7 expressing). The cells were cross-linked using formaldehyde at 1% final concentration and incubated at room temperature for 10 mins. The cross-linking was quenched using glycine at a final concentration of 0.125M and incubated at room temperature for 5 mins, followed by washing twice with 1X PBS. The cells were harvested by scraping in 1X PBS (with Protease Inhibitor cocktail) and pelleting the cells by centrifugation at 300xg for 5 mins, at 4°C. The cells were re-suspended in 1ml of SDS lysis buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH 8.1 and protease inhibitors before use) and incubated on ice for 15 mins. The lysate was sonicated at 30% amplitude, 14 x 10 secs pulses, 20 secs pauses, followed by centrifugation at 14,000 rpm at 4°C. The lysate was pre-cleared with 80µl of 50% slurry of Protein A Agarose beads for 60 mins at 4°C, followed by centrifugation

at 4,000 rpm for 1 min at 4°C. Specific antibody (5µg) was added to the supernatant and precipitated overnight at 4°C with rotation. After overnight rotation, 60µl of blocked Protein A Agarose beads (1ml 50% bead slurry, 0.5mg BSA) were added to the antibody/protein complex, followed by rotation for 60 mins at 4°C. The beads were pelleted by centrifugation at 1,000 rpm for 1 min at 4°C, followed by sequential washing with 1X Low Salt Immune Complex wash buffer (0.1% SDS, 1% TritonX-100, 2mM EDTA, 20mM Tris-HCl pH 8.1, 150mM NaCl), 1X High Salt Immune Complex wash Buffer (0.1% SDS, 1% TritonX-100, 2mM EDTA, 20mM Tris-HCl pH 8.1, 500mM NaCl), 1X LiCl Immune Complex wash Buffer (0.25M LiCl, 1% NP-40, 1% Deoxycholic acid sodium salt, 1mM EDTA, 10mM Tris-HCl pH 8.1) and 1X TE (10mM Tris pH 8.1, 1mM EDTA). The beads were eluted with freshly prepared elution buffer (1% SDS, 0.1M NaHCO₃) and incubated at room temperature for 30 mins with agitation. The DNA was reverse cross-linked by adding 12µl of 5M NaCl per 100µl elute and incubated at 65°C overnight, followed by RNase A treatment and purified using PCR purification kit.

The purified DNA was then used for SYBR green based qPCR estimation in order to identify enrichment of H3K4me3 and H3K27me3 marks in the promoter regions of selected HOX cluster genes, which show altered expression in cancers. HPV16 positive SiHa cell line served as a positive control and HPV negative C33A cell line served as a negative control for HPV16 E7 immunoprecipitation. The primers used for qPCR enrichment study are listed in Supplementary Table 9. The amplification plots and dissociation curves for the qPCR reactions are provided in Supplementary Figures 3 and 4, respectively. The primers targeting RPL-30 and EVX-1 promoters were used as positive control for H3K4me3 and H3K27me3 enrichment (Supplementary Figures 5 and 6), respectively, while ZNF333-3' served as a negative control for H3K4me3 and H3K27me3 enrichment (Supplementary Figure 7). The H3K4me3 and H3K27me3 enrichment status in HPV16 positive cell line SiHa have been illustrated in Supplementary Figure 8.

Statistical analysis

Kolmogorov-Smirnov test was used to check for the normality of the Δ Ct values. Non-parametric test (Mann-Whitney U Test) was performed to check for statistical significance of the changes in expression levels, as the Δ Ct values did not follow normal distribution. Multiple sets of experiments using cell line model were compared using ANOVA for statistically significant differences in expression levels between the experimental categories. All statistical analyses were done using SPSS v16.0. The p values were reported after performing Bonferroni multiple testing corrections.

Author contributions

S. Sharma Saha and S. Sengupta conceived and designed the experiments. S. Sharma Saha performed the experiments. S. Sharma Saha and S. Sengupta analyzed the data. R. Roy Chowdhury, N.R. Mondal, and S. Roy, contributed reagents/materials/analysis tools and patient data and checked the draft of the manuscript. S. Sengupta supervised the study. S. Sharma Saha and S. Sengupta wrote the manuscript. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

We acknowledge the contributions of Saroj Gupta Cancer Centre and Research Institute (SGCC & RI, Thakurpukur, South 24 Parganas, West Bengal, India), Calcutta Medical College Hospital (Kolkata, West Bengal, India) and College of Medicine & JNM Hospital (Kalyani, Nadia, West Bengal, India), for providing us the clinical samples for the study; Council of Scientific and Industrial Research, India, for providing Sweta Sharma Saha (JRF and SRF) with the fellowship to work on this project; all the members of National Institute of Biomedical Genomics (Kalyani, West Bengal, India), for their support during the work.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

GRANT SUPPORT

This work was supported partially by Department of Biotechnology (Grant No.: BT/PR2012/MED/29/312/2011), Government of India, and partially by National Institute of Biomedical Genomics (Core Grant).

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN. *Int J Cancer*. 2015; 136: E359–E386.
2. Guidelines for cervical cancer screening programme, Government of India – World Health Organization Collaboration Programme. 2004-2005; (2006).
3. Zur Hausen H. Papillomavirus infections—a major cause of human cancers. *Biochim Biophys Acta*. 1996; 1228: F55–F78.
4. Bosch FX, Manos MM, Muñoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV. Prevalence of human papillomavirus DNA in cervical cancer: a worldwide perspective. *International Biological*

- Study on Cervical Cancer (IBSCC) Study Group. *J Natl Cancer Inst.* 1995; 87: 796–802.
5. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Muñoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol.* 1999; 189:12–19.
 6. Narisawa-Saito M, Kiyono T. Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: Roles of E6 and E7 proteins. *Cancer Sci.* 2007 98:1505–1511.
 7. Huang PS, Patrick DR, Edwards G, Goodhart PJ, Huber HE, Miles L, Garsky VM, Oliff A, Heimbrook DC. Protein domains governing interactions between E2F, the retinoblastoma gene product, and human papillomavirus type 16 E7 protein. *Mol Cell Biol.* 1993; 13: 953-960.
 8. Jones DL, Münger, K. Analysis of the p53-mediated G1 growth arrest pathway in cells expressing the human papillomavirus type 16 E7 oncoprotein. *J Virol.* 1997; 4: 2905-2912.
 9. Sharma S, Mandal P, Sadhukhan T, Roy Chowdhury R, Mondal NR, Chakravarty B, Chatterjee T, Roy S, Sengupta S. Bridging Links between Long Noncoding RNA HOTAIR and HPV Oncoprotein E7 in Cervical Cancer Pathogenesis. *Scientific Reports.* 2015; 5: 11724.
 10. Bracken AP, Pasini D, Capra M, Prosperini E, Colli E, Helin K. EZH2 is downstream of the pRb-E2F pathway, essential for proliferation and amplified in cancer. *EMBO J.* 2003; 22: 5323-5335.
 11. Gehring WJ, Hiromi Y. Homoeotic genes and the homeobox. *Ann Rev Genet.* 1986; 20: 147-173.
 12. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. *Cell.* 1992; 68: 283-302.
 13. Levine M, Hoey T. Homeobox proteins as sequence specific transcription factors. *Cell.* 1988; 55: 537-540.
 14. Apiou F, Flagiello D, Cillo C, Malfoy B, Poupon MF, Dutrillaux B. Fine mapping of human HOX gene clusters. *Cytogenet Cell Genet.* 1996; 73: 114-115.
 15. Dolle P, Izpisua-Belmonte JC, Brown J, Tickle C, Duboule D. Hox genes and the morphogenesis of the vertebrate limb. *Prog Clin Biol Res.* 1993; 383A: 11-20.
 16. Cardoso WV. Transcription factors and pattern formation in the developing lung. *Am J Physiol.* 1995; 269: L429–L442.
 17. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell.* 2006; 125: 301-313.
 18. Buske C, Feuring-Buske M, Abramovich C, Spiekermann K, Eaves CJ, Coulombel L, Sauvageau G, Hogge DE, Humphries RK. Deregulated expression of HOXB4 enhances the primitive growth activity of human hematopoietic cells. *Blood.* 2002; 100: 862-868.
 19. Veraksa A, Del Campo M, McGinnis W. Developmental patterning genes and their conserved functions: From model organisms to humans. *Mol Genet Metabol.* 2000; 69: 85–100.
 20. Sauvageau G, Thorsteinsdottir U, Eaves CJ, Lawrence HJ, Largman C, Lansdorp PM, Humphries RK. Overexpression of HOXB4 in hematopoietic cells causes the selective expansion of more primitive populations *in vitro* and *in vivo*. *Genes Dev.* 1995; 9: 1753-1765.
 21. Bhatlekar S, Fields JZ, Boman BM. HOX genes and their role in the development of human cancers. *J Mol Med (Berl).* 2014; 92: 811-823.
 22. Yoshida H, Broaddus R, Cheng W, Xie S, Naora H. Deregulation of the HOXA10 homeobox gene in endometrial carcinoma: role in epithelial-mesenchymal transition. *Cancer Res.* 2006; 66: 889-897.
 23. Yuan H, Kajiyama H, Ito S, Chen D, Shibata K, Hamaguchi M, Kikkawa F, Senga T. HOXB13 and ALX4 induce SLUG expression for the promotion of EMT and cell invasion in ovarian cancer cells. *Oncotarget.* 2015; 6: 13359-13370. doi: 10.18632/oncotarget.3673.
 24. Zhai LL, Wu Y, Cai CY, Tang ZG. Overexpression of homeobox B-13 correlates with angiogenesis, aberrant expression of EMT markers, aggressive characteristics and poor prognosis in pancreatic carcinoma. *Int J Clin Exp Pathol.* 2015; 8: 6919-6927.
 25. Hung YC, Ueda M, Terai Y, Kumagai K, Ueki K, Kanda K, Yamaguchi H, Akise D, Ueki M. Homeobox gene expression and mutation in cervical carcinoma cells. *Cancer Sci.* 2003; 94: 437-441.
 26. López R, Garrido E, Piña P, Hidalgo A, Lazos M, Ochoa R, Salcedo M. HOXB homeobox gene expression in cervical carcinoma. *Int J Gynecol Cancer.* 2006; 16: 329-335.
 27. López R, Garrido E, Vázquez G, Piña P, Pérez C, Alvarado I, Salcedo M. A subgroup of HOX Abd-B gene is differentially expressed in cervical cancer. *Int J Gynecol Cancer.* 2006; 16: 1289-1296.
 28. Gonzalez-Herrera A, Salgado-Bernabe M, Velazquez-Velazquez C, Salcedo-Vargas M, Andrade-Manzano A, Avila-Moreno F, Pina-Sanchez P. Increased expression of HOXB2 and HOXB13 proteins is associated with HPV infection and cervical cancer progression. *Asian Pac J Cancer Prev.* 2015; 16: 1349-1353.
 29. Santin AD, Zhan F, Bignotti E, Siegel ER, Cané S, Bellone S, Palmieri M, Anfossi S, Thomas M, Burnett A, Kay HH, Roman JJ, O'Brien TJ, et al. Gene expression profiles of primary HPV16- and HPV18-infected early stage cervical cancers and normal cervical epithelium: identification of novel candidate molecular markers for cervical cancer diagnosis and therapy. *Virology.* 2005;331: 269–291.
 30. Zhai Y, Kuick R, Nan B, Ota I, Weiss SJ, Trimble CL. Gene expression analysis of preinvasive and invasive cervical squamous cell carcinoma identifies HOXC10 as a key mediator of invasion. *Cancer Res.* 2007; 67: 10163-10172.

31. McLaughlin-Drubin ME, Crum CP, Münger K. Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proc Natl Acad Sci U S A*. 2011; 108: 2130-2135.
32. Caberg JH, Hubert PM, Begon DY, Herfs MF, Roncarati PJ, Boniver JJ, Delvenne PO. Silencing of E7 oncogene restores functional E-Cadherin expression in human papillomavirus 16-transformed keratinocytes. *Carcinogenesis*. 2008; 29: 1441-1447.
33. Gu ZD, Shen LY, Wang H, Chen XM, Li Y, Ning T, Chen KN. HOXA13 promotes cancer cell growth and predicts poor survival of patients with esophageal squamous cell carcinoma. *Cancer Res*. 2009; 69: 4969-4973.
34. Han Y, Tu WW, Wen YG, Li DP, Qiu GQ, Tang HM, Peng ZH, Zhou CZ. Identification and validation that up-expression of HOXA13 is a novel independent prognostic marker of a worse outcome in gastric cancer based on immunohistochemistry. *Med Oncol*. 2013; 30: 564.
35. Quagliata L, Matter MS, Piscuoglio S, Arabi L, Ruiz C, Procino A, Kovac M, Moretti F, Makowska Z, Boldanova T, Andersen JB, Hämmerle M, Tornillo L, et al. Long noncoding RNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. *Hepatology*. 2014; 59: 911-923.
36. Shah N, Jin K, Cruz LA, Park S, Sadik H, Cho S, Goswami CP, Nakshatri H, Gupta R, Chang HY, Zhang Z, Cimino-Mathews A, Cope L, et al. HOXB13 mediates tamoxifen resistance and invasiveness in human breast cancer by suppressing ER α and inducing IL-6 expression. *Cancer Res*. 2013; 73: 5449-5458.
37. Adwan H, Zhivkova-Galunska M, Georges R, Eyol E, Kleeff J, Giese NA, Friess H, Bergmann F, Berger MR. Expression of HOXC8 is inversely related to the progression and metastasis of pancreatic ductal adenocarcinoma. *Br J Cancer*. 2011; 105: 288-295.
38. Alami Y, Castronovo V, Belotti D, Flagiello D, Clausse N. HOXC5 and HOXC8 expression are selectively turned on in human cervical cancer cells compared to normal keratinocytes. *Biochem Biophys Res Commun*. 1999; 257: 738-45.
39. Omatu T. Overexpression of human homeobox gene in lung cancer A549 cells results in enhanced motile and invasive properties. [Article in Japanese]. *Hokkaido Igaky Zasshi*. 1999; 74: 367-376.
40. Vider BZ, Zimber A, Chastre E, Gespach C, Halperin M, Mashiah P, Yaniv A, Gazit A. Deregulated expression of homeobox containing genes, HOXB6, B8, C8, C9, and Cdx-1, in human colon cancer cell lines. *Biochem Biophys Res Commun*. 2000; 272: 513-518.
41. Burgers WA, Blanchon L, Pradhan S, de Launoit Y, Kouzarides T, Fuks F. Viral oncoproteins target the DNA methyltransferases. *Oncogene*. 2006; 26: 1650-1655.
42. Banister CE, Liu C, Pirisi L, Creek KE, Buckhaults PJ. Identification and characterization of HPV-independent cervical cancers. *Oncotarget*. 2017; 8: 13375-13386. doi: 10.18632/oncotarget.14533.
43. Das Ghosh D, Bhattacharjee B, Sen S, Premi L, Mukhopadhyay I, Roy Chowdhury R, Roy S, Sengupta S. Some novel insights on HPV16 related cervical cancer pathogenesis based on analyses of LCR methylation, viral load, E7 and E2/E4 expressions. *PLoS One*. 2012; 7: e44678.