Published in final edited form as:

J Cell Physiol. 2014 October ; 229(10): 1405–1416. doi:10.1002/jcp.24577.

Altered Histone Mark Deposition and DNA Methylation at Homeobox Genes in Human Oral Squamous Cell Carcinoma Cells

Katarzyna M. Marcinkiewicz and Lorraine J. Gudas

Department of Pharmacology, Weill Cornell Medical College and Weill Graduate School of Biomedical Sciences of Cornell University, 1300 York Avenue, New York, NY 10065

Abstract

We recently reported a role of Polycomb repressive complex 2 (PRC2) and PRC2 trimethylation of histone 3 lysine 27 (H3K27me3) in the regulation of homeobox (HOX) (Marcinkiewicz and Gudas, 2013) gene transcript levels in human oral keratinocytes (OKF6-TERT1R) and tongue squamous cell carcinoma (SCC) cells. Here, we assessed both the levels of various histone modifications at a subset of homeobox genes and genome wide DNA methylation patterns in OKF6-TERT1R and SCC-9 cells by using ERRBS (enhanced reduced representation bisulfite sequencing). We detected the H3K9me3 mark at HOXB7, HOXC10, HOXC13 and HOXD8 at levels higher in OKF6-TERT1R than in SCC-9 cells; at IRX1 and SIX2 the H3K9me3 levels were conversely higher in SCC-9 than in OKF6-TERT1R. The H3K79me3 mark was detectable only at IRX1 in OKF6-TERT1R and at IRX4 in SCC-9 cells. The levels of H3K4me3 and H3K36me3 marks correlate with the transcript levels of the assessed homeobox genes in both OKF6-TERT1R and SCC-9. We detected generally lower CpG methylation levels on DNA in SCC-9 cells at annotated genomic regions which were differentially methylated between OKF6-TERT1R and SCC-9 cells; however, some genomic regions, including the HOX gene clusters, showed DNA methylation at higher levels in SCC-9 than OKF6-TERT1R. Thus, both altered histone modification patterns and changes in DNA methylation are associated with dysregulation of homeobox gene expression in human oral cavity SCC cells, and this dysregulation potentially plays a role in the neoplastic phenotype of oral keratinocytes.

Keywords

epigenetic silencing; oral squamous cell carcinoma; bisulfite-seq; homeobox; HOX clusters

Introduction

Head and neck cancer is one of the 10 most prevalent cancers worldwide (Siegel et al., 2013). Around 90% of all head and neck cancers are squamous cell carcinomas (HNSCCs) (Curado and Hashibe, 2009). A subset of squamous cell carcinomas arising within oral

Address correspondence to: Lorraine J. Gudas, Department of Pharmacology, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065; Tel. 212-746-6250; Fax. 212-746-8858; ljgudas@med.cornell.edu.

The authors disclose no potential conflicts of interest.

cavity is often referred to as oral squamous cell carcinoma (OSCC). One in 66 men and one in 148 women in the US will develop oral cavity or pharyngeal cancer during their lifetime, and one in 263 men and one in 548 women will die of these diseases (American Cancer Society, 2013).

Many aspects of carcinogenesis are driven by epigenetic mechanisms, such as post-translational modifications of histone proteins and the conversion of cytosine to 5-methylcytosine within DNA (You and Jones, 2012). Trimethylation of H3 lysine 9 (H3K9me3) is an epigenetic modification of histone H3 commonly found in transcriptionally inactive regions of the genome (Hublitz et al., 2009; Kim and Kim, 2012). H3K9me3 can be found at satellite repeats, telomeres, and long terminal repeats, and H3K9me3 marks genomic regions that control imprinting (Blahnik et al., 2011; Mikkelsen et al., 2007; Pauler et al., 2009). The H3K9 trimethylation mark is added by members of the Suv39 family, such as EHMT2/G9a (ENSG00000204371, euchromatic histone-lysine N-methyltransferase 2), SUV39H1 (ENSG00000101945, suppressor of variegation 3–9 homolog 1), and SETDB2 (ENSG00000136169, SET domain, bifurcated 2), as well as the non-Suv39 methyltransferases PRDM2 (ENSG00000116731, PR domain containing 2, with ZNF domain) and ASH1L (ENSG00000116539, ash1 (absent, small, or homeotic)-like) (Völkel and Angrand, 2007). Trimethylation of histone 3 lysine 79 (H3K79me3) is present on both active and silent gene promoters (Barski et al., 2007), and plays an important role in heterochromatin formation (Jones et al., 2008). H3K79 methylation is mediated by Dot1L (Jones et al., 2008). Trimethylation of H3 lysine 4 and 36 is found on actively transcribed genes, with the H3K4me3 mark typically most abundant near proximal promoter regions (Barski et al., 2007; Bernstein et al., 2005; Heintzman et al., 2007) and the H3K36me3 mark typically associated with gene bodies (Bannister et al., 2005).

Methylation of cytosines in CpG-rich regions of the genome, known as CpG islands (CGIs), is typically found near transcription start sites or within gene promoter regions (Gardiner-Garden and Frommer, 1987; Illingworth and Bird, 2009). DNA methylation at promoters and gene bodies regulates transcriptional activity (Gardiner-Garden and Frommer, 1987; Illingworth and Bird, 2009; Jjingo et al., 2012; Kulis et al., 2013; Maunakea et al., 2010), and DNA methylation in repeat regions, such as centromeres, is vital for chromosomal stability (Cheng et al., 2008). Tumor suppressor silencing through gene promoter methylation and DNA hypomethylation of specific genes has been reported in HNSCC (Ha and Califano, 2006; Jithesh et al., 2013; Lleras et al., 2013; Shaw et al., 2006; Youssef et al., 2004).

We recently discovered a role for the epigenetic mark H3K27me3 in the regulation of homeobox gene transcript levels in human oral keratinocytes and tongue squamous cell carcinoma (SCC) cells (Marcinkiewicz and Gudas, 2013). We found that the levels of the H3K27me3 marks at chromatin near homeobox genes were inversely correlated with the transcript levels in non-tumorigenic, immortalized human oral keratinocytes (OKF6-TERT1R) and tumorigenic oral SCC-9 cells. Similarly, the association of the protein SUZ12, a core component of the Polycomb repressive complex 2 (PRC2), with chromatin near the homeobox genes was inversely correlated with the transcript levels in OKF6-TERT1R and SCC-9 cells (Marcinkiewicz and Gudas, 2013). SUZ12 depletion by shRNA

technology increased HOXB7, HOXC10, HOXC13, and HOXD8 transcript levels in OKF6-TERT1R cells, suggesting that the silencing of these HOX genes in non-tumorigenic, oral keratinocytes results from the deposition of the H3K27me3 mark by EZH2, a component of PRC2 complex. In contrast, SUZ 12 depletion did not increase the transcript levels of the homeobox genes IRX1, IRX4, SIX2 and TSHZ3 in tumorigenic SCC-9 cells (Marcinkiewicz and Gudas, 2013).

To understand the epigenetic mechanisms that become altered during the neoplastic transformation of normal epithelial cells of the oral cavity we utilized chromatin immunoprecipitation (ChIP) to assess the levels of the epigenetic marks H3K9me3, H3K79me3, H3K4me3, and H3K36me3 at homeobox genes differentially expressed in non-tumorigenic OKF6-TERT1R cells as compared to tumorigenic SCC-9 cells. We also employed enhanced reduced representation bisulfite sequencing (ERRBS) (Akalın et al., 2012a) to assess genome-wide DNA methylation patterns in OKF6-TERT1R and SCC-9 cells.

Materials and methods

Cell Culture and Chemicals

The immortalized, non-transformed human oral keratinocyte cell line OKF6-TERT1R (Dickson et al., 2000) was kindly provided by Dr. James G. Rheinwald, Harvard Medical School. These cells were cultured in Keratinocyte-SFM medium (# 10744019, Gibco, CA, USA), supplemented with 0.3 mM CaCl₂, 0.2 ng/ml EGF, penicillin/streptomycin/L-glutamine (#10378-016, Gibco, CA, USA) and bovine pituitary extract (BPE), as previously described (Dickson et al., 2000; Marcinkiewicz and Gudas, 2013). The human OSCC line SCC-9 was cultured in Dulbecco's Modified Eagle's Medium: Nutrient Mixture F-12 (DMEM/F-12; Gibco, CA, USA), supplemented with 10% fetal calf serum and 400 ng/ml hydrocortisone. For mRNA analyses and DNA extraction, cells were plated in 10 cm² tissue culture plates at a density of 2×10^6 cells/plate. For chromatin immunoprecipitation (ChIP), the cells were plated in 15 cm² tissue culture plates at a density of 3×10^6 cells/plate.

Total RNA Extraction, Deep Sequencing (RNAseq) and Data Analysis

For deep sequencing, total RNA was isolated from cells using an RNeasy mini kit (Qiagen Inc. Valencia, CA, USA), including an in-column DNase treatment. The RNA quality check, cDNA library preparation, and deep sequencing was carried out by the Weill Cornell Medical College Genomics Resources Core Facility and data analyses were performed as described (Marcinkiewicz and Gudas, 2013). Some of these data have been reported previously (Marcinkiewicz and Gudas, 2013).

Chromatin Immunoprecipitation (ChIP)

ChIP experiments were performed as described (Gillespie and Gudas, 2007), with two modifications: (1) the chromatin was sheared using a Branson 150 Sonifier (setting 3) for 3 20 s intervals, and (2) 20 µl of soluble chromatin sample was removed, diluted to a final volume of 100 µl with ChIP elution buffer (50 mM Tris-HCl (pH 8.0), 1% SDS, 1 mM EDTA), 4 µl of 5 M NaCl was added, and samples were incubated at 65°C overnight to

reverse the cross-links in the chromatin. The DNA was purified using a PCR purification kit (# 28106 Qiagen, Gaithersburg, MD, USA) according to the manufacturer's protocol, eluted using 20 µl of the elution buffer, and the DNA concentration was calculated based on 260 nm absorbance measurements. The remaining chromatin sample was stored at – 80°C. For each immunoprecipitation (IP), a volume of chromatin corresponding to 20 µg of DNA was used and diluted to a total volume of 500 µl with lysis buffer. Antibodies for ChIP analysis were: rabbit polyclonal anti-Histone H3 (trimethyl K4) antibody - ChIP Grade (ab 8580, Abcam, Cambridge, MA, USA); rabbit polyclonal to Histone H3 (trimethyl K9) - ChIP Grade (ab8898, Abcam, Cambridge, MA, USA); rabbit polyclonal to Histone H3 (trimethyl K79) - ChIP Grade (ab2621, Abcam, Cambridge, MA, USA); rabbit polyclonal to Histone H3 (trimethyl K36) - ChIP Grade (ab 9050, Abcam, Cambridge, MA, USA); and normal rabbit IgG (sc-2027, Santa Cruz Biotechnology, Inc. Dallas, TX, USA). 2 µg of each antibody was used for each immunoprecipitation sample. DNA purified from the input and the immunoprecipitated chromatin samples was used as template in quantitative real time PCR. The primers used have been described (Marcinkiewicz and Gudas, 2013).

Genomic DNA Extraction, Enhanced Reduced Representation Bisulfite Sequencing (ERRBS) and Data Analysis

Total genomic DNA was extracted from OKF6-TERT1R and SCC-9 cells using a DNeasy Blood & Tissue Kit (Qiagen, Germantown, MD, USA), including an RNase treatment step. ERRBS was performed at the Weill Cornell Medical College Epigenomics Core Facility as described (Akalin et al., 2012a). Data analysis was performed with guidance and help from Dr. Fabienne Campagne (Institute for Computational Biomedicine, Weill Cornell Medical College (WCMC)) and Dr. Nigel Mongan (School of Veterinary Medicine and Science, University of Nottingham, UK). Sequencing data, generated in .tar files, were uploaded into the GobyWeb application (Campagne lab, Institute for Computational Biomedicine, Weill Cornell Medical College of Cornell University) for alignment to the genome (Dorff et al., 2013). Differential methylation analysis was performed using GobyWeb (Dorff et al., 2013) and Cistrome-Galaxy (Liu et al., 2011). Cistrome is a web-based platform for data intensive biomedical research, using Galaxy as the backbone for the pipeline module (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010; Liu et al., 2011). Percent methylation levels for CpG dinucleotides are calculated by dividing the number of methylated Cs by the total coverage of that base. Methylation levels (ML) for genomic regions (promoters and gene bodies) between OKF6-TERT1R and SCC-9 cells were compared by taking the mean methylation percentage of CpG dinucleotides overlapping those regions. Regions with ML differences of 20 percentage points between the OKF6-TERT1R and SCC-9 cells were considered differentially methylated. Regions were filtered by significance, using a q value less than or equal to 0.01 in the Benjamini and Hochberg adjusted t-test (adjustment for False Discovery Rate.). Differential methylation analysis was also performed by Dr. Piali Mukherjee (Weill Cornell Medical College Epigenomics Core Facility) using MethylKit software (Akalin et al., 2012b), and identical results were obtained.

Statistical Analysis

All ChIP experiments include at least 3 independent biological replicates (n = 3). Results of quantitative PCR experiments were analyzed using the GraphPad Prism program. A two-

way ANOVA followed by Bonferroni post-test correction or an unpaired t test were applied, as indicated in the figure legends.

Sequencing Data Accession Numbers

RNA and bisulfite sequencing data were deposited in the GEO database with accession numbers GSE51415 and GSE51752, respectively.

Results

Transcriptionally silent HOX genes possess H3K9me3 chromatin marks in OKF6-TERT1R cells

We recently reported that numerous homeobox genes exhibit different transcript levels in non-tumorigenic, oral OKF6-TERT1R as compared to tumorigenic SCC-9 cells (Marcinkiewicz and Gudas, 2013). We hypothesized that the changes in homeobox gene transcript levels in the SCC-9 vs. OKF6-TERT1R cells are a consequence of changes in their epigenetic signatures.

To dissect the mechanisms by which HOX gene expression is regulated in non-tumorigenic oral keratinocytes and oral SCC cells, we used chromatin immunoprecipitation (ChIP) to evaluate the association of H3K9me3 marks with chromatin near the homeobox genes that are expressed differentially between the OKF6-TERT1R and SCC-9 cells (Fig. 1(a)). Homeobox gene transcripts expressed at higher levels in SCC-9 than in OKF6-TERT1R cells, specifically HOXB7, HOXC10, HOXC13, and HOXD8 (Fig. 1(a) A to D), exhibited lower levels of the H3K9me3 epigenetic marks in SCC-9 compared to OKF6-TERT1R (Fig. 1(b) A to D). We found that the H3K9me3 signal was 14.8 fold higher ($p < 0.0001$, Fig. 1(b) A) at the HOXB7 gene, 3.8 fold higher ($p < 0.0001$, Fig. 1(b) B) at the HOXC10 gene, 6.2 fold higher ($p < 0.0001$, Fig. 1(b) C) at the HOXC13 gene, and 7.7 fold higher ($p < 0.0001$, Fig. 1(b) D) at the HOXD8 gene in OKF6-TERT1R compared to SCC-9. A summary of these results is shown (Table 1).

Thus, H3K9me3 marks are present at DNA regions near the assessed HOX genes (HOXB7, HOXC10, HOXC13, and HOXD8) at higher levels in OKF6-TERT1R than in SCC-9 cells. A considerable body of experimental evidence suggests that the H3K9me3 mark is a hallmark of constitutive heterochromatin (Peters et al., 2002; Rea et al., 2000). Thus, these data indicate that these HOX genes are silenced in a manner that excludes rapid initiation of transcription in OKF6-TERT1R cells (Fig. 1(a) A to D).

Homeobox genes with lower transcript levels in SCC-9 than in OKF6-TERT1R cells possess H3K9me3 chromatin marks in SCC-9 cells

The genes IRX1, IRX4, SIX2, and TSHZ3, which are expressed at lower levels in SCC-9 than in OKF6-TERT1R (Fig. 1(a) E to H), exhibited variable levels of the H3K9me3 mark in SCC-9 compared to OKF6-TERT1R cells (Fig. 1(b) E to H). A summary of these results, including the fold changes in H3K9me3 levels between OKF6-TERT1R and SCC-9 cells at the IRX1, IRX4, SIX2, and TSHZ3 genes, is shown (Table 1). We conclude that the transcriptional silencing of these homeobox genes (Fig. 1(b) E to H) in SCC-9 cells is not

tightly associated with the presence of the H3K9me3 mark nearby. Therefore, other epigenetic mechanisms are likely to be involved in the transcriptional regulation of IRX1, IRX4, SIX2, and TSHZ3 in OKF6-TERT1R and SCC-9 cells.

The H3K79me3 mark is detectable only at IRX1 and IRX4

Our next set of experiments assessed the presence of the H3K79me3 epigenetic mark at the HOXB7, HOXC10, HOXC13, HOXD8, IRX1, IRX4, SIX2, and TSHZ3 genes in OKF6-TERT1R and SCC-9 cells. The H3K79me3 mark is associated with silent gene promoters (Barski et al., 2007) and plays an important role in heterochromatin formation (Jones et al., 2008); the H3K79me3 mark is also found at active gene promoters (Barski et al., 2007). The H3K79me3 signal was enriched over the negative control IgG signal at only two of the homeobox genes, IRX1 and IRX4. The H3K79me3 signal was 4.6 fold lower at IRX1 ($p=0.0418$; Fig. 2A) in SCC-9 as compared to OKF6-TERT1R cells and 13 fold lower at IRX4 ($p=0.0208$; Fig. 2B) in OKF6-TERT1R as compared to SCC-9 cells. Thus, the H3K79me3 mark is not consistently present at DNA regions near the assessed homeobox genes in OKF6-TERT1R and SCC-9 cells.

HOX genes with higher transcript levels in SCC-9 than in OKF6-TERT1R cells possess H3K4me3 and H3K36me3 marks in SCC-9 cells

Next, we evaluated the placement of two marks generally associated with active transcription, H3K4me3 and H3K36me3 (Bannister et al., 2005; Barski et al., 2007; Bernstein et al., 2002), near the same HOX genes expressed differentially between the OKF6-TERT1R and SCC-9 cells. Homeobox transcripts expressed at higher levels in SCC-9 than in OKF6-TERT1R cells, specifically HOXB7, HOXC10, HOXC13, and HOXD8 (Fig. 1(a) A to D), exhibited higher levels of the H3K4me3 epigenetic mark in SCC-9 compared to OKF6-TERT1 cells (Fig. 3(a) A to D), but only the HOXC13 gene was associated with higher levels of the H3K36me3 mark in SCC-9 compared to OKF6-TERT1 cells (Fig. 3(b) C). The H3K4me3 levels were 25 fold higher ($p < 0.001$, Fig. 3(a) A) at the HOXB7 gene, 2.6 fold higher at the HOXC10 gene ($p < 0.01$, Fig. 3(a) B), and 3.0 fold higher at the HOXD8 gene ($p < 0.0001$, Fig. 3(a) D) in SCC-9 as compared to OKF6-TERT1R cells. Similarly, at the HOXC13 gene the H3K4me3 levels were 3.3 fold ($p < 0.0001$, Fig. 3(a) C) higher and the H3K36me3 levels were 2.3 fold ($p < 0.05$, and; Fig. 3(a) C) higher in SCC-9 as compared to OKF6-TERT1R cells. A summary of these results is also shown (Table 1).

The H3K36me3 levels were 2.7 fold ($p < 0.0001$, Fig. 3(b) A) higher at the HOXB7 gene, 9.6 fold ($p < 0.001$, Fig. 3(b) B) higher at the HOXC10 gene, and 17.2 fold ($p < 0.0001$, Fig. 3(b) D) higher at the HOXD8 gene in OKF6-TERT1R compared to SCC-9 cells. Thus, the H3K4me3 mark is present at high levels at DNA regions near the assessed HOX genes in SCC-9 cells, but the H3K36me3 mark is not consistently localized at the assessed HOX genes in OKF6-TERT1R and SCC-9 cells. These data indicate that the H3K36me3 mark is likely to have functions in addition to promoting active transcription of the HOX genes in SCC-9 cells.

Homeobox genes with lower transcript levels in SCC-9 than in OKF6-TERT1R cells possess H3K4me3 and H3K36me3 chromatin marks in OKF6-TERT1R cells

The genes IRX1, IRX4, SIX2, and TSHZ3, expressed at lower levels in SCC-9 than in OKF6-TERT1R cells, generally exhibited higher levels of the H3K4me3 and H3K36me3 marks in OKF6-TERT1R than in SCC-9 cells (Figs. 3(a) E to H and 3(b) E to H). The H3K4me3 levels were 16.5 fold ($p < 0.0001$; Fig. 3(a) E) higher at the IRX1 gene, 6.0 fold ($p < 0.0001$; Fig. 3(a) G) higher at the SIX2 gene, and 1.5 fold (Fig. 3(a) H) higher at the TSHZ3 gene in OKF6-TERT1R as compared to SCC-9 cells. The H3K36me3 levels were 5.9 fold ($p < 0.0001$; Fig. 3(b) E) higher at the IRX1 gene, 4.5 fold higher ($p < 0.001$, Fig. 3(b) F) at the IRX4 gene, 5.9 fold ($p < 0.0001$; Fig. 3(b) G) higher at the SIX2 gene, and 3.1 fold ($p < 0.01$, Fig. 3(b) H) higher at the TSHZ3 gene in OKF6-TERT1R compared to SCC-9 cells. The one exception was the H3K4me3 level at the IRX4 gene (Fig. 3(a) F). The H3K4me3 signal at IRX4 was enriched over that of the negative control IgG (Fig. 3(a) J), but was essentially identical in OKF6-TERT1R and SCC-9 (Fig. 3(a) F). A summary of these results is shown (Table 1). We conclude that the transcriptional silencing of the homeobox genes IRX1, IRX4, SIX2, and TSHZ3 in SCC-9 cells is associated with a major reduction in the H3K4me3 and H3K36me3 marks nearby.

DNA methylation and gene expression in OKF6-TERT1R and SCC-9 cells

In our investigation of epigenetic mechanisms altered in tumorigenic versus non-tumorigenic human oral keratinocytes we first measured one type of modification, namely histone 3 lysine methylation, by ChIP. We next used enhanced reduced representation bisulfite sequencing (ERRBS) (Akalin et al., 2012a), a next generation sequencing based technique, to assess the levels of cytosine methylation in genomic DNA. Roughly thirty percent of the cytosine residues at CpGs were methylated in both OKF6-TERT1R and SCC-9 cells (33.1% and 30.6 %, respectively), indicating no gross differences in the overall levels of DNA methylation between the cells in our model system.

Annotated genomic regions differentially methylated between OKF6-TERT1R and SCC-9 cells have lower methylation levels in SCC-9 than in OKF6-TERT1R

To understand how DNA methylation influences the phenotype of non-tumorigenic and tumorigenic oral keratinocytes, we next focused on methylation across annotated genomic regions: gene bodies and gene promoters (defined as a 2000 bp sequence immediately upstream of the first TSS), as opposed to the single base level. A graphical representation of CpG methylation levels along gene bodies and proximal promoters with methylation levels different by at least 20 percentage points between OKF6-TERT1R and SCC-9 cells is shown (Fig. 4). Four times more genes had higher CpG methylation levels along gene bodies in OKF6-TERT1R than in SCC-9 cells than vice versa (4671 vs. 1152 genes, Fig. 4A). Additionally, 50% more proximal promoter regions had methylation levels higher in OKF6-TERT1R cells than in SCC-9 cells than vice versa (7733 vs. 5085 regions, Fig. 4B). Thus, we detected *lower global methylation levels in SCC-9 cells at the annotated genomic regions* which were differentially methylated between OKF6-TERT1R and SCC-9 cells.

Some promoters frequently methylated in human OSCC samples have higher methylation levels in SCC-9 than in OKF6-TERT1R

Next, we reviewed the literature to identify genes known to undergo promoter methylation during carcinogenesis, and we compiled gene body and proximal promoter region ERRBS data for these genes (Table 2). Many of the genes with promoter regions frequently methylated in human OSCC samples versus normal oral mucosa also show higher methylation levels in their proximal promoter regions in SCC-9 compared to OKF6-TERT1R cells; CDKN2A, 76.7% vs. 56.4%; DAPK1, 10.8% vs. 4.4%; IRF8, 38.2% vs. 19.7%; IRX1, 66.1% vs. 2.7%; MGMT1, 28.1% vs. 23.1%; p53, 96.0% vs. 86.0%; p73, 11.0% vs. 6.5%; and RAR β , 20.0% vs. 11.1% (Table 2). Other genes have higher methylation levels along gene bodies in SCC-9 than in OKF6-TERT1R cells; CDKN2A, 69.6% vs. 32.8%; EBF3, 61.7% vs. 28.1%; HOXA9, 70.5% vs. 12.1%; IRX1, 72.9% vs. 45.9%; and SERPINB5, 83.9% vs. 61.2% (Table 2). In contrast, some genes show higher methylation levels along gene bodies in OKF6-TERT1R compared to SCC-9 cells: AIM2, 40.3% vs. 7.7%; DCC, 31.6% vs. 2.9%; and MGMT, 39.4% vs. 6.2% (Table 2). These data suggest that some of the differences in transcript levels between OKF6-TERT1R and SCC-9 may result from different DNA methylation patterns.

Methylation of genomic regions of homeobox genes is unlikely to be the major silencing mechanism in OKF6-TERT1R and SCC-9 cells

Methylation levels at individual CpGs within genomic regions of homeobox genes for which we assessed histone methylation patterns are shown (Fig. 5(a) and (b)). Most of these homeobox genes are characterized by very similar DNA methylation levels in OKF6-TERT1R and SCC-9 cells. Thus, we conclude that DNA methylation is unlikely to constitute a major transcription regulatory mechanism for these genes. The only exception to this general conclusion is the IRX1 gene (Fig. 5(b) A). IRX1 exhibited much higher overall CpG methylation levels along both the gene body and proximal promoter region in SCC-9 compared to OKF6-TERT1R (72.9% vs. 45.9%; and 63.9% vs. 3.8%, respectively; Table 2); thus, high promoter methylation of IRX1 may contribute to the transcriptional silencing of this gene in SCC-9 (Figs. 1(a)E and 5(b)A).

Methylation of entire homeobox (HOX) gene clusters may regulate specific HOX gene expression in SCC-9

An overview of the gene body and proximal promoter region methylation data of *all homeobox genes with different transcript levels between the OKF6-TERT1R and SCC-9 cells* is shown (Table 3). Interestingly, HOX genes show higher DNA methylation levels in SCC-9 than in OKF6-TERT1R. HOXB3, HOXB7, HOXD4, HOXC4, and HOXD10 have higher DNA methylation levels along their gene bodies in SCC-9 than in OKF6-TERT1R (HOXB3, 69.2% vs. 5.2%; HOXB7, 20.6% vs. 2.5%; HOXD4, 54.5% vs. 11.3%; HOXC4, 46.2% vs. 9.0%; and HOXD10, 59.9% vs. 11.8%; Table 3). These data are consistent with reports in the literature that more actively transcribed genes have DNA methylation in their gene bodies (Hahn et al., 2011; Hellman and Chess, 2007; Jjingo et al., 2012; Kulis et al., 2013; Maunakea et al., 2010; Nguyen et al., 2001; Shenker and Flanagan, 2012). Additionally, HOX genes B3, B7, D4, and C4 have higher methylation levels along their

proximal promoter regions in SCC-9 than in OKF6-TERT1R (HOXB3, 78.4% vs. 13.4%; HOXB7, 77.0% vs. 9.1%; HOXD4, 72.6% vs. 22.7%; HOXC4 50.2% vs. 9.4%; Table 3). The DNA methylation levels at individual CpGs within the genomic regions of entire HOX gene clusters are also shown (Fig. 5(c)).

Discussion

Homeobox genes transcriptionally silent in OKF6-TERT1R and SCC-9 cells bear the H3K9me3 mark of constitutive heterochromatin

A considerable body of experimental evidence suggests that the H3K9me3 mark is a marker for constitutive heterochromatin (Pauler et al., 2009; Peters et al., 2002; Rea et al., 2000). The H3K9me3 mark is present at high levels on histone H3 near the HOXB7, HOXC10, HOXC13, and HOXD8 in OKF6-TERT1R and near IRX1 and SIX2 in SCC-9 cells (Fig. 1(b) A to E and G). We previously reported high levels of the H3K27me3 mark at the same genes (Marcinkiewicz and Gudas, 2013). The H3K9me3 and the H3K27me3 epigenetic marks are usually found in distinct genomic locations, the first being generally targeted to gene-poor regions, and the latter generally associated with CpG-rich sequences, often within gene promoters (Barski et al., 2007; Kim and Kim, 2012). We propose that the concomitant presence of the two silencing marks, H3K9me3 and H3K27me3, at the homeobox genes we assessed suggests that these homeobox genes are silenced permanently (resembling chromatin compaction within centromeres or chromosome X inactivation (Almouzni and Probst, 2011; Lessing and Lee, 2013)) rather than in a facultative manner (resembling silencing of developmental genes in stem cells (Bernstein et al., 2006; Kashyap et al., 2011; Landeira et al., 2010)).

SUZ12 has also been implicated in regulation of H3K9 methylation (de la Cruz et al., 2007), and shRNA mediated SUZ12 depletion was sufficient to increase transcript levels of the assessed HOX genes in OKF6-TERT1R cells (Marcinkiewicz and Gudas, 2013). Thus, it is plausible that SUZ12 present near the HOXB7, HOXC10, HOXC13, and HOXD8 genes in the OKF6-TERT1R cells plays a role in H3K9me3 deposition. We postulate that the loss of SUZ12 association with the HOX genes in SCC-9 results in both the loss of the H3K27me3 and H3K9me3 marks and the transcriptional activation of these HOX genes in SCC-9 cells.

The H3K79me3 mark is present at the iroquois homeobox genes

Seemingly contradictory reports describe H3K79me3 as both a mark of eu- and heterochromatin (Barski et al., 2007; Jones et al., 2008; Nguyen et al., 2011). Careful analysis of the data, however, offers an explanation. Barski *et al.* found H3K79me3 to be located at a higher percentage of transcriptionally active compared to silent promoters; however, the levels of H3K79me3 were generally higher at silent promoters than at the actively transcribed ones (Barski et al., 2007). We detected the H3K79me3 mark at only two of the assessed genes: IRX1 and IRX4 (Fig. 2), suggesting that this mark is not consistently involved in the regulation of transcription of the homeobox genes in human oral keratinocytes.

IRX4 has recently been suggested to be a putative tumor suppressor in prostate cancer (Nguyen et al., 2012). The loss of IRX4 expression in the SCC-9 as compared to the OKF6-TERT1R cells implies that IRX4 could play a similar tumor suppressor role in OSCC.

The H3K4me3 mark correlates *positively* with high transcript levels at the assessed homeobox genes in OKF6-TERT1R and SCC-9

Trimethylation of H3 lysine 4 is widely recognized as a mark of active gene promoters (Bernstein et al., 2002; Bernstein et al., 2007; Santos-Rosa et al., 2002). The balance between H3K27 and H3K4 trimethylation near the gene proximal promoter regions appears to be a common mechanism of transcriptional activation and silencing regulation (Barski et al., 2007; Bartke et al., 2010; Schmitges et al., 2011). PRC2 activity is inhibited by H3K4me3 (Schmitges et al., 2011). In line with the above cited work, we have found that the H3K4me3 marks correlate positively with high transcript levels of the assessed genes in both OKF6-TERT1R and SCC-9 cells (Fig. 3(a)).

The H3K36me3 mark is likely to have functions in addition to promoting active transcription of the HOX genes in SCC-9 cells

We evaluated the deposition of the H3K36me3 mark at the selected homeobox genes (Fig. 3(b)). The H3K36me3 mark has been reported to demarcate transcriptionally active genes (Bannister et al., 2005; Barski et al., 2007), and to impact negatively H3K27me3 deposition, as assessed by *in vitro* methylation of lysine 27 of H3 with a pre-installed K36 trimethyl mark (Yuan et al., 2011) or a trimethyllysine analog at position K36 (Schmitges et al., 2011). However, the H3K36me3 mark may constitute a docking site for Tudor domains of mammalian PCL proteins, and hence PCL-containing PRC2 complexes (Ballaré et al., 2012; Brien et al., 2012; Cai et al., 2013).

Interestingly, HOXB7, HOXC10, and HOXD8, three genes with transcript levels higher in SCC-9 than in OKF6-TERT1R, exhibited a higher level of H3K36me3 signal in OKF6-TERT1R than in SCC-9 cells, which correlates with the high levels of associated SUZ12 and H3K27me3 (Fig. 3(b) A, B, and D). This presence of both “activating” K36 and “silencing” K27 trimethylation marks could indicate that the H3K36me3 mark at HOXB7, HOXC10, and HOXD8 is indeed recognized by one or more of PCL proteins that recruit the PRC2 complex and initiate silencing of these genes. It would be interesting to test whether the H3K36me3 mark is indeed necessary for recruitment of the PRC2 complex, subsequent H3K27 methylation, and transcriptional silencing of HOXB7, HOXC10, and HOXD8 genes in OKF6-TERT1R cells.

In summary, the H3 methylation marks that we describe here do not consistently correlate with the transcript levels of the assessed homeobox genes (HOXB7, HOXC10, HOXC13, HOXD8, IRX1, IRX4, SIX2, and TSHZ3). In contrast, SUZ12 levels and the levels of the H3K27me3 mark did correlate (Marcinkiewicz and Gudas, 2013). This strengthens our earlier postulate that SUZ12/PRC2 plays a particularly important role in widespread gene transcriptional regulation in tumorigenic oral keratinocytes (Marcinkiewicz and Gudas, 2013). Concomitantly, we identified gene-specific patterns of histone methylation for HOXC13, IRX1, IRX4, and TSHZ3. These results constitute a novel contribution to our

understanding of alterations in transcriptional regulation during neoplastic transformation of oral keratinocytes and encourage future investigation into the role of epigenetics in OSCC.

Functions of DNA methylation are highly context-specific and proximal promoter regions are differentially methylated between OKF6-TERT1R and SCC-9

In recent years, rapid developments in next generation sequencing, as well as more sophisticated bioinformatics algorithms to analyze the data (for a comprehensive review (Bock, 2012)), have allowed us to determine that the functions of DNA methylation are highly context-specific and complex (Baylin and Jones, 2011; Jones, 2012; You and Jones, 2012). We observed lower DNA methylation levels in SCC-9 cells at annotated genomic regions which are differentially methylated between OKF6-TERT1R and SCC-9, consistent with the theory that hypermethylation of some specific promoter regions accompanies a more widespread DNA hypomethylation in cancer cells (De Smet and Loriot, 2013; Ehrlich and Lacey, 2013; Toraño et al., 2012; You and Jones, 2012). Of course, these data are just associations. Additional experiments would be needed to test whether the promoter methylation level is the primary regulator of transcription in OKF6-TERT1R and SCC-9 cells, either genome-wide or at particular genes.

Genes frequently methylated in HNSCC are methylated in SCC-9 cells

We next assembled a list of genes frequently methylated in human HNSCC patients (Bennett et al., 2008; Chen et al., 2013; Guerrero-Preston et al., 2011; Ha and Califano, 2006; Hennessey et al., 2011; Huang et al., 2009; Jithesh et al., 2013; Koike et al., 2013; Kozaki et al., 2008; Lee et al., 2013; Liu et al., 2012; López et al., 2003; Maruya et al., 2004; Nagata et al., 2012; Pickering et al., 2013; Poage et al., 2012; Poage et al., 2011; Ruesga et al., 2007; Shaw et al., 2006; Steinmann et al., 2009; Su et al., 2010; Towle et al., 2013; Viswanathan et al., 2003; Wiklund et al., 2011; Youssef et al., 2004) and evaluated methylation levels along gene bodies and proximal promoter regions of these genes in OKF6-TERT1R and SCC-9 cells. We found that the data from our model are consistent with those found from studies of HNSCC patient data (Bennett et al., 2008; Lleras et al., 2013). For example, decreased methylation of AIM2 and MAGEA3 in HNSCC patient tumor samples vs. adjacent mucosa has been reported by Lleras *et al* (Lleras et al., 2013). Promoter hypomethylation of AIM2 in OSCC samples as compared to normal mucosa has also been described by other groups (Jithesh et al., 2013; Poage et al., 2010). Increased methylation of ZNF671 in HNSCC patient tumor samples vs. adjacent mucosa has also been reported by Lleras *et al* (Lleras et al., 2013). EBF3 has been reported to be silenced by aberrant promoter methylation in HNSCC patient tumor samples by Bennett *et al* (Bennett et al., 2008).

Our data are also in accord with the hypothesis proposed by Bennett *et al* (Bennett et al., 2008) that EBF3 and IRX1 are tumor suppressor genes silenced by promoter methylation in HNSCC tumors, and are similar to results from HNSCC patients in which promoter methylation of HOXA9 is reported to be higher in tumor samples as compared to normal mucosa (Guerrero-Preston et al., 2011; Jithesh et al., 2013; Poage et al., 2010).

Our results contrast with those of Hennessey *et al*. (Hennessey et al., 2011), in which promoter methylation of HNSCC cell lines was reported to be different from methylation in

primary human tumors. This report used human tissue samples along with OKF6 cells, similar to the OKF6-TERT1R cells used here, and several different HNSCC cell lines, though not SCC-9 cells. Importantly, several of the genes which Hennessey *et al.* (Hennessey et al., 2011) found to be differentially methylated between primary tumors and normal mucosa samples (namely HOXA9, HOXD12, HOXD9, IRF8, GREM1, ZNF671, SERPINA12, and TRIM58; Tables 2 and 3) are differentially methylated at either promoter or gene body levels in OKF6-TERT1R and SCC-9 cells. Thus, this could explain the discrepancy between Hennessey *et al.* (Hennessey et al., 2011) and our data, as they examined only promoter methylation.

DNA methylation of IRX1

Only one of the eight transcripts examined (HOXB7, HOXC10, HOXC13, HOXD8, IRX1, IRX4, SIX2, and TSHZ3) is likely to be regulated by DNA methylation: IRX1 (Table 2 and Figure 5(b) A). The entire proximal promoter region and gene body of IRX1 meet the criteria of a CpG island (CGI). CpGs in the proximity of the IRX1 TSS are almost entirely unmethylated in OKF6-TERT1R cells but are methylated at much higher level in SCC-9 cells. CpGs within the IRX1 body are methylated at variable levels in OKF6-TERT1R and SCC-9 cells, and the overall methylation level of the IRX1 gene body is higher in SCC-9 than in OKF6-TERT1R (Table 2). However, a subset of CpGs further downstream of the IRX1 gene body is methylated at a higher level in the OKF6-TERT1R cells than in SCC-9 (Fig. 5(b) A). Although a consensus view has not yet been reached, data indicate that gene body methylation is permissive for transcription (Jjingo et al., 2012). The methylation patterns of IRX1 in OKF6-TERT1R and SCC-9 cells are consistent with both a silencing role for promoter DNA methylation and a permissive role for gene body methylation for transcription. The silencing histone methylation marks present at IRX1, H3K27me3 ((Marcinkiewicz and Gudas, 2013)) and H3K9me3 (Fig. 1(b) E)) in SCC-9 and the activating histone marks H3K4me3 and H3K36me3 (Figs. 3(a)E and 3(b)E) absent from IRX1 in SCC-9 cells also could play a role in lowering IRX1 transcript levels in SCC-9 cells.

DNA methylation patterns of HOX gene clusters

Half of the homeobox genes that show transcript levels higher in SCC-9 than in OKF6-TERT1R cells are HOX genes that are located on chromosomes in gene-dense clusters ((Marcinkiewicz and Gudas, 2013), Fig. 5(c) and Table 3). Despite their close genomic localization, HOX genes from one cluster are not always expressed at the same time or in the same tissues (Diez del Corral and Storey, 2004; Kashyap et al., 2011; Lemons and McGinnis, 2006; Noordermeer et al., 2011). It is possible that the entire HOX clusters remain methylated in SCC-9 (Fig. 5(c)) to prevent initiation of transcription of a particular HOX gene without the proper signal, and thus, to facilitate transcriptional regulation by other epigenetic mechanisms. Branciamore *et al.* (Branciamore et al., 2010) analyzed the density of CpG nucleotides within the coding regions of genes and found that about 10% of protein-coding genes showed much less depletion of CpG nucleotides than the average for the genome; these CpG rich genes include HOX genes, other homeobox containing genes, and genes encoding other transcription factors. Branciamore *et al.* (Branciamore et al., 2010) concluded that entire HOX gene clusters overlap with what the group termed CGI clusters-

genomic regions particularly rich not only in CpGs, but in entire CpG islands. They discussed the evolutionary advantage that the high intragenic CpG content could offer and proposed that the high CpG content within the exons of genes encoding transcription factors facilitates epigenetic, tissue-specific transcriptional control of these genes (Branciamore *et al.*, 2010). Branciamore *et al.* also observed that although most promoter associated CGIs are not methylated in normal tissue, intragenic CGIs are methylated in a tissue specific manner (Branciamore *et al.*, 2010). Tissue specific, intragenic DNA methylation has also been observed by others (Hawkins *et al.*, 2010; Lister *et al.*, 2009; Lister *et al.*, 2011; Tsumagari *et al.*, 2013) and is consistent with our data.

The high DNA methylation levels in either or both gene bodies and promoter regions of HOX genes in SCC-9 cells (Table 3) could also reflect aberrant methylation patterns present in neoplastically transformed cells. This explanation is consistent with published reports that describe hypermethylation of HOX genes in HNSCC patient tumor samples. Increased methylation of HOXD9, HOXD10, and HOXD12 associated CpG islands in HNSCC patient tumor samples vs. adjacent mucosa has been reported by Lleras *et al.* (Lleras *et al.*, 2013). HOXA9 and HOXA11 were identified as hypermethylated at promoter regions in multiple studies (Guerrero-Preston *et al.*, 2011; Jithesh *et al.*, 2013; Poage *et al.*, 2010).

In conclusion, our data suggest that differential gene expression between normal oral keratinocytes and tumorigenic SCC cells is a consequence of dysregulated epigenetic mechanisms: altered histone modification patterns, and to a lesser extent, DNA methylation.

Acknowledgments

This research was supported by NIH grant RO1DE010389 to LJG and by Weill Cornell funds. KM was supported by a Kirschstein-NRSA predoctoral fellowship F31-DE021316, and for a portion of this project on training grant 5T32GM073546-04. We thank Dr. Fabienne Campagne for guidance and help with deep sequencing experiments design and analysis. We thank Dr. Nigel Mongan for his advice and help with next generation sequencing data analysis, and Ms. Abigail Horstmann and Dr. Kwame Osei-Sarfo for critically reading the manuscript and for their insightful suggestions.

References

- Akalin A, Garrett-Bakelman FE, Kormaksson M, Busuttill J, Zhang L, Khrebtukova I, Milne TA, Huang Y, Biswas D, Hess JL, *et al.* Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet.* 2012a; 8:e1002781. [PubMed: 22737091]
- Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012b; 13:R87. [PubMed: 23034086]
- Almouzni G, Probst AV. Heterochromatin maintenance and establishment: lessons from the mouse pericentromere. *Nucleus.* 2011; 2:332–338. [PubMed: 21941119]
- American Cancer Society. 2013. <http://www.cancer.org/index>
- Ballaré C, Lange M, Lapinaite A, Martin GM, Morey L, Pascual G, Liefke R, Simon B, Shi Y, Gozani O, *et al.* Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat Struct Mol Biol.* 2012; 19:1257–1265. [PubMed: 23104054]
- Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem.* 2005; 280:17732–17736. [PubMed: 15760899]

- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007; 129:823–837. [PubMed: 17512414]
- Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell*. 2010; 143:470–484. [PubMed: 21029866]
- Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer*. 2011; 11:726–734. [PubMed: 21941284]
- Bennett KL, Karpenko M, Lin MT, Claus R, Arab K, Dyckhoff G, Plinkert P, Herpel E, Smiraglia D, Plass C. Frequently methylated tumor suppressor genes in head and neck squamous cell carcinoma. *Cancer Res*. 2008; 68:4494–4499. [PubMed: 18559491]
- Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A*. 2002; 99:8695–8700. [PubMed: 12060701]
- Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, Gingeras TR, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005; 120:169–181. [PubMed: 15680324]
- Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell*. 2007; 128:669–681. [PubMed: 17320505]
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006; 125:315–326. [PubMed: 16630819]
- Blahnik KR, Dou L, Echipare L, Iyengar S, O'Geen H, Sanchez E, Zhao Y, Marra MA, Hirst M, Costello JF, et al. Characterization of the contradictory chromatin signatures at the 3' exons of zinc finger genes. *PLoS One*. 2011; 6:e17121. [PubMed: 21347206]
- Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol*. 2010; Chapter 19(Unit 19.10.11–21)
- Bock C. Analysing and interpreting DNA methylation data. *Nat Rev Genet*. 2012; 13:705–719. [PubMed: 22986265]
- Branciamore S, Chen ZX, Riggs AD, Rodin SN. CpG island clusters and pro-epigenetic selection for CpGs in protein-coding exons of HOX and other transcription factors. *Proc Natl Acad Sci U S A*. 2010; 107:15485–15490. [PubMed: 20716685]
- Brien GL, Gambero G, O'Connell DJ, Jerman E, Turner SA, Egan CM, Dunne EJ, Jurgens MC, Wynne K, Piao L, et al. Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat Struct Mol Biol*. 2012; 19:1273–1281. [PubMed: 23160351]
- Cai L, Rothbart SB, Lu R, Xu B, Chen WY, Tripathy A, Rockowitz S, Zheng D, Patel DJ, Allis CD, et al. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol Cell*. 2013; 49:571–582. [PubMed: 23273982]
- Chen JH, Yeh KT, Yang YM, Chang JG, Lee HE, Hung SY. High expressions of histone methylation- and phosphorylation-related proteins are associated with prognosis of oral squamous cell carcinoma in male population of Taiwan. *Med Oncol*. 2013; 30:513. [PubMed: 23504335]
- Cheng YW, Pincas H, Bacolod MD, Schemmann G, Giardina SF, Huang J, Barral S, Idrees K, Khan SA, Zeng Z, et al. CpG island methylator phenotype associates with low-degree chromosomal abnormalities in colorectal cancer. *Clin Cancer Res*. 2008; 14:6005–6013. [PubMed: 18829479]
- Curado MP, Hashibe M. Recent changes in the epidemiology of head and neck cancer. *Curr Opin Oncol*. 2009; 21:194–200. [PubMed: 19363341]
- de la Cruz CC, Kirmizis A, Simon MD, Isono K, Koseki H, Panning B. The polycomb group protein SUZ12 regulates histone H3 lysine 9 methylation and HP1 alpha distribution. *Chromosome Res*. 2007; 15:299–314. [PubMed: 17406994]
- De Smet C, Lorient A. DNA hypomethylation and activation of germline-specific genes in cancer. *Adv Exp Med Biol*. 2013; 754:149–166. [PubMed: 22956500]

- Dickson MA, Hahn WC, Ino Y, Ronfard V, Wu JY, Weinberg RA, Louis DN, Li FP, Rheinwald JG. Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span become immortal yet retain normal growth and differentiation characteristics. *Mol Cell Biol.* 2000; 20:1436–1447. [PubMed: 10648628]
- Diez del Corral R, Storey KG. Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis. *Bioessays.* 2004; 26:857–869. [PubMed: 15273988]
- Dorff KC, Chambwe N, Zeno Z, Simi M, Shaknovich R, Campagne F. GobyWeb: Simplified Management and Analysis of Gene Expression and DNA Methylation Sequencing Data. *PLoS One.* 2013; 8:e69666. [PubMed: 23936070]
- Ehrlich M, Lacey M. DNA hypomethylation and hemimethylation in cancer. *Adv Exp Med Biol.* 2013; 754:31–56. [PubMed: 22956495]
- Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. *J Mol Biol.* 1987; 196:261–282. [PubMed: 3656447]
- Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, et al. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* 2005; 15:1451–1455. [PubMed: 16169926]
- Gillespie RF, Gudas LJ. Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. *J Mol Biol.* 2007; 372:298–316. [PubMed: 17663992]
- Goecks J, Nekrutenko A, Taylor J, Team G. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 2010; 11:R86. [PubMed: 20738864]
- Guerrero-Preston R, Soudry E, Acero J, Orera M, Moreno-López L, Macía-Colón G, Jaffe A, Berdasco M, Ili-Gangas C, Brebi-Mieville P, et al. NID2 and HOXA9 promoter hypermethylation as biomarkers for prevention and early detection in oral cavity squamous cell carcinoma tissues and saliva. *Cancer Prev Res (Phila).* 2011; 4:1061–1072. [PubMed: 21558411]
- Ha PK, Califano JA. Promoter methylation and inactivation of tumour-suppressor genes in oral squamous-cell carcinoma. *Lancet Oncol.* 2006; 7:77–82. [PubMed: 16389187]
- Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLoS One.* 2011; 6:e18844. [PubMed: 21526191]
- Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, Pelizzola M, Edsall LE, Kuan S, Luu Y, Klugman S, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell.* 2010; 6:479–491. [PubMed: 20452322]
- Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet.* 2007; 39:311–318. [PubMed: 17277777]
- Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. *Science.* 2007; 315:1141–1143. [PubMed: 17322062]
- Hennessey PT, Ochs MF, Mydlarz WW, Hsueh W, Cope L, Yu W, Califano JA. Promoter methylation in head and neck squamous cell carcinoma cell lines is significantly different than methylation in primary tumors and xenografts. *PLoS One.* 2011; 6:e20584. [PubMed: 21637785]
- Huang KH, Huang SF, Chen IH, Liao CT, Wang HM, Hsieh LL. Methylation of RASSF1A, RASSF2A, and HIN-1 is associated with poor outcome after radiotherapy, but not surgery, in oral squamous cell carcinoma. *Clin Cancer Res.* 2009; 15:4174–4180. [PubMed: 19509163]
- Hublitz P, Albert M, Peters AH. Mechanisms of transcriptional repression by histone lysine methylation. *Int J Dev Biol.* 2009; 53:335–354. [PubMed: 19412890]
- Illingworth RS, Bird AP. CpG islands--‘a rough guide’. *FEBS Lett.* 2009; 583:1713–1720. [PubMed: 19376112]
- Jithesh PV, Risk JM, Schache AG, Dhanda J, Lane B, Liloglou T, Shaw RJ. The epigenetic landscape of oral squamous cell carcinoma. *Br J Cancer.* 2013; 108:370–379. [PubMed: 23287992]

- Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. *Oncotarget*. 2012; 3:462–474. [PubMed: 22577155]
- Jones B, Su H, Bhat A, Lei H, Bajko J, Hevi S, Baltus GA, Kadam S, Zhai H, Valdez R, et al. The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet*. 2008; 4:e1000190. [PubMed: 18787701]
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012; 13:484–492. [PubMed: 22641018]
- Kashyap V, Gudas LJ, Brenet F, Funk P, Viale A, Scandura JM. Epigenomic reorganization of the clustered Hox genes in embryonic stem cells induced by retinoic acid. *J Biol Chem*. 2011; 286:3250–3260. [PubMed: 21087926]
- Kim J, Kim H. Recruitment and biological consequences of histone modification of H3K27me3 and H3K9me3. *ILAR J*. 2012; 53:232–239. [PubMed: 23744963]
- Koike K, Kasamatsu A, Iyoda M, Saito Y, Kouzu Y, Koike H, Sakamoto Y, Ogawara K, Tanzawa H, Uzawa K. High prevalence of epigenetic inactivation of the human four and a half LIM domains 1 gene in human oral cancer. *Int J Oncol*. 2013; 42:141–150. [PubMed: 23123766]
- Kozaki K, Imoto I, Mogi S, Omura K, Inazawa J. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. *Cancer Res*. 2008; 68:2094–2105. [PubMed: 18381414]
- Kulis M, Queirós AC, Beekman R, Martín-Subero JI. Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. *Biochim Biophys Acta*. 2013
- Landeira D, Sauer S, Poot R, Dvorkina M, Mazzarella L, Jørgensen HF, Pereira CF, Leleu M, Piccolo FM, Spivakov M, et al. Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. *Nat Cell Biol*. 2010; 12:618–624. [PubMed: 20473294]
- Lee CH, Wong TS, Chan JY, Lu SC, Lin P, Cheng AJ, Chen YJ, Chang JS, Hsiao SH, Leu YW, et al. Epigenetic Regulation of the X-Linked Tumor Suppressors BEX1 and LDOC1 in Oral Squamous Cell Carcinoma. *J Pathol*. 2013
- Lemons D, McGinnis W. Genomic evolution of Hox gene clusters. *Science*. 2006; 313:1918–1922. [PubMed: 17008523]
- Lessing D, Lee JT. X chromosome inactivation and epigenetic responses to cellular reprogramming. *Annu Rev Genomics Hum Genet*. 2013; 14:85–110. [PubMed: 23662665]
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462:315–322. [PubMed: 19829295]
- Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*. 2011; 471:68–73. [PubMed: 21289626]
- Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, Shin H, Wong SS, Ma J, Lei Y, et al. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol*. 2011; 12:R83. [PubMed: 21859476]
- Liu Y, Zhou ZT, He QB, Jiang WW. DAPK promoter hypermethylation in tissues and body fluids of oral precancer patients. *Med Oncol*. 2012; 29:729–733. [PubMed: 21516484]
- Lleras R, Smith RV, Adrien LR, Schlecht NF, Burk RD, Harris T, Childs G, Prystowsky MB, Belbin TJ. Unique DNA Methylation Loci Distinguish Anatomic Site and HPV Status in Head and Neck Squamous Cell Carcinoma. *Clin Cancer Res*. 2013
- López M, Aguirre JM, Cuevas N, Anzola M, Videgain J, Aguirregaviria J, Martínez de Pancorbo M. Gene promoter hypermethylation in oral rinses of leukoplakia patients--a diagnostic and/or prognostic tool? *Eur J Cancer*. 2003; 39:2306–2309. [PubMed: 14556921]
- Marcinkiewicz KM, Gudas LJ. Altered epigenetic regulation of homeobox genes in human oral squamous cell carcinoma cells. *Exp Cell Res*. 2013
- Maruya S, Issa JP, Weber RS, Rosenthal DI, Haviland JC, Lotan R, El-Naggar AK. Differential methylation status of tumor-associated genes in head and neck squamous carcinoma: incidence and potential implications. *Clin Cancer Res*. 2004; 10:3825–3830. [PubMed: 15173091]

- Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*. 2010; 466:253–257. [PubMed: 20613842]
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. 2007; 448:553–560. [PubMed: 17603471]
- Nagata S, Hamada T, Yamada N, Yokoyama S, Kitamoto S, Kanmura Y, Nomura M, Kamikawa Y, Yonezawa S, Sugihara K. Aberrant DNA methylation of tumor-related genes in oral rinse: a noninvasive method for detection of oral squamous cell carcinoma. *Cancer*. 2012; 118:4298–4308. [PubMed: 22252571]
- Nguyen AT, Xiao B, Nepl RL, Kallin EM, Li J, Chen T, Wang DZ, Xiao X, Zhang Y. DOT1L regulates dystrophin expression and is critical for cardiac function. *Genes Dev*. 2011; 25:263–274. [PubMed: 21289070]
- Nguyen C, Liang G, Nguyen TT, Tsao-Wei D, Groshen S, Lübbert M, Zhou JH, Benedict WF, Jones PA. Susceptibility of nonpromoter CpG islands to de novo methylation in normal and neoplastic cells. *J Natl Cancer Inst*. 2001; 93:1465–1472. [PubMed: 11584062]
- Nguyen HH, Takata R, Akamatsu S, Shigemizu D, Tsunoda T, Furihata M, Takahashi A, Kubo M, Kamatani N, Ogawa O, et al. IRX4 at 5p15 suppresses prostate cancer growth through the interaction with vitamin D receptor, conferring prostate cancer susceptibility. *Hum Mol Genet*. 2012; 21:2076–2085. [PubMed: 22323358]
- Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. The dynamic architecture of Hox gene clusters. *Science*. 2011; 334:222–225. [PubMed: 21998387]
- Pauler FM, Sloane MA, Huang R, Regha K, Koerner MV, Tamir I, Sommer A, Aszodi A, Jenuwein T, Barlow DP. H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res*. 2009; 19:221–233. [PubMed: 19047520]
- Peters AH, Mermoud JE, O'Carroll D, Pagani M, Schweizer D, Brockdorff N, Jenuwein T. Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nat Genet*. 2002; 30:77–80. [PubMed: 11740497]
- Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, Zhao M, Ortega Alves MV, Chang K, Drummond J, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. *Cancer Discov*. 2013; 3:770–781. [PubMed: 23619168]
- Poage GM, Butler RA, Houseman EA, McClean MD, Nelson HH, Christensen BC, Marsit CJ, Kelsey KT. Identification of an epigenetic profile classifier that is associated with survival in head and neck cancer. *Cancer Res*. 2012; 72:2728–2737. [PubMed: 22507853]
- Poage GM, Christensen BC, Houseman EA, McClean MD, Wiencke JK, Posner MR, Clark JR, Nelson HH, Marsit CJ, Kelsey KT. Genetic and epigenetic somatic alterations in head and neck squamous cell carcinomas are globally coordinated but not locally targeted. *PLoS One*. 2010; 5:e9651. [PubMed: 20300172]
- Poage GM, Houseman EA, Christensen BC, Butler RA, Avissar-Whiting M, McClean MD, Waterboer T, Pawlita M, Marsit CJ, Kelsey KT. Global hypomethylation identifies Loci targeted for hypermethylation in head and neck cancer. *Clin Cancer Res*. 2011; 17:3579–3589. [PubMed: 21505061]
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*. 2000; 406:593–599. [PubMed: 10949293]
- Ruesga MT, Acha-Sagredo A, Rodríguez MJ, Aguirregaviria JI, Videgain J, Rodríguez C, de Pancorbo MeL, Aguirre JM. p16(INK4a) promoter hypermethylation in oral scrapings of oral squamous cell carcinoma risk patients. *Cancer Lett*. 2007; 250:140–145. [PubMed: 17113222]
- Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T. Active genes are tri-methylated at K4 of histone H3. *Nature*. 2002; 419:407–411. [PubMed: 12353038]

- Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell*. 2011; 42:330–341. [PubMed: 21549310]
- Shaw RJ, Liloglou T, Rogers SN, Brown JS, Vaughan ED, Lowe D, Field JK, Risk JM. Promoter methylation of P16, RARbeta, E-cadherin, cyclin A1 and cytoglobin in oral cancer: quantitative evaluation using pyrosequencing. *Br J Cancer*. 2006; 94:561–568. [PubMed: 16449996]
- Shenker N, Flanagan JM. Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *Br J Cancer*. 2012; 106:248–253. [PubMed: 22166804]
- Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin*. 2013; 63:11–30. [PubMed: 23335087]
- Steinmann K, Sandner A, Schagdarsurengin U, Dammann RH. Frequent promoter hypermethylation of tumor-related genes in head and neck squamous cell carcinoma. *Oncol Rep*. 2009; 22:1519–1526. [PubMed: 19885608]
- Su PF, Huang WL, Wu HT, Wu CH, Liu TY, Kao SY. p16(INK4A) promoter hypermethylation is associated with invasiveness and prognosis of oral squamous cell carcinoma in an age-dependent manner. *Oral Oncol*. 2010; 46:734–739. [PubMed: 20729138]
- Toraño EG, Petrus S, Fernandez AF, Fraga MF. Global DNA hypomethylation in cancer: review of validated methods and clinical significance. *Clin Chem Lab Med*. 2012; 50:1733–1742. [PubMed: 23089701]
- Towle R, Truong D, Hogg K, Robinson WP, Poh CF, Garnis C. Global analysis of DNA methylation changes during progression of oral cancer. *Oral Oncol*. 2013
- Tsumagari K, Baribault C, Terragni J, Chandra S, Renshaw C, Sun Z, Song L, Crawford GE, Pradhan S, Lacey M, et al. DNA methylation and differentiation: HOX genes in muscle cells. *Epigenetics Chromatin*. 2013; 6:25. [PubMed: 23916067]
- Viswanathan M, Tsuchida N, Shanmugam G. Promoter hypermethylation profile of tumor-associated genes p16, p15, hMLH1, MGMT and E-cadherin in oral squamous cell carcinoma. *Int J Cancer*. 2003; 105:41–46. [PubMed: 12672028]
- Völkel P, Angrand PO. The control of histone lysine methylation in epigenetic regulation. *Biochimie*. 2007; 89:1–20. [PubMed: 16919862]
- Wiklund ED, Gao S, Hulf T, Sibbritt T, Nair S, Costea DE, Villadsen SB, Bakholdt V, Bramsen JB, Sørensen JA, et al. MicroRNA alterations and associated aberrant DNA methylation patterns across multiple sample types in oral squamous cell carcinoma. *PLoS One*. 2011; 6:e27840. [PubMed: 22132151]
- You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*. 2012; 22:9–20. [PubMed: 22789535]
- Youssef EM, Lotan D, Issa JP, Wakasa K, Fan YH, Mao L, Hassan K, Feng L, Lee JJ, Lippman SM, et al. Hypermethylation of the retinoic acid receptor-beta(2) gene in head and neck carcinogenesis. *Clin Cancer Res*. 2004; 10:1733–1742. [PubMed: 15014026]
- Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem*. 2011; 286:7983–7989. [PubMed: 21239496]

Figure 1. (a) mRNA levels of homeobox genes expressed differentially in OKF6-TERT1R and SCC-9 cells and (b) levels of the H3K9me3 epigenetic marks proximal to homeobox genes expressed differentially between OKF6-TERT1R and SCC-9 cells
 n=3 independent biological repeats; * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, **** indicates $p < 0.0001$; note differences in scales of y-axes in different panels (a) mRNA levels are represented as mean RPKM values (Reads per Kilobase per Million Reads) obtained from an RNAseq experiment; differences in transcript levels between the cell lines were analyzed by unpaired t test. A to D: mRNA levels of HOX genes expressed at higher levels in SCC-9 vs. OKF6-TERT1R cells. E to H: mRNA levels of homeobox genes expressed at lower levels in SCC-9 vs. OKF6-TERT1R cells. *Note: These data have been reported in an earlier publication (Marcinkiewicz and Gudas, 2013) and are shown here for*

reference. (b) Levels of H3K9me3 epigenetic marks proximal to homeobox genes expressed differentially between OKF6-TERT1R and SCC-9 cells, as assessed by ChIP; data are represented relative to chromatin sample used as input in corresponding IPs. Differences in immunoprecipitated chromatin levels between the OKF6-TERT1R (clear bars) and SCC-9 (black bars) cells were analyzed by two-way ANOVA followed by Bonferroni post-test. A to D: Levels of the H3K9me3 marks at HOX genes expressed at higher levels in SCC-9 vs. OKF6-TERT1R cells; E to H: Levels of the H3K9me3 mark at HOX genes expressed at lower levels in SCC-9 vs. OKF6-TERT1R cells; I and J: Negative control: chromatin immunoprecipitated using normal rabbit IgG; sequences amplified are indicated by the graph legend. Note differences in scales of y-axes in different graphs.

Figure 2. Levels of the H3K79me3 epigenetic mark proximal to IRX1 and IRX4 genes in OKF6-TERT1R and SCC-9 cells

n=3 independent biological repeats; differences in immunoprecipitated chromatin levels between the cell lines were analyzed by unpaired t-test; **A and B:** Levels of the H3K79me3 marks at IRX1 (A) and IRX4 (B) in OKF6-TERT1R (clear bars) and SCC-9 (black bars) cells, as assessed by ChIP; data are represented relative to chromatin sample used as input in corresponding IPs; **C:** Negative control: chromatin immunoprecipitated using normal rabbit IgG; sequences amplified are indicated by the graph legend.

Figure 3. Levels of the H3K4me3 (a) and H3K36me3 (b) epigenetic marks proximal to homeobox genes expressed differentially between OKF6-TERT1R and SCC-9 cells

n=3 independent biological repeats; * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, and **** indicates $p < 0.0001$; Note differences in scales of y-axes in different panels; (a) levels of the H3K4me3 epigenetic marks proximal to homeobox genes expressed differentially between OKF6-TERT1R (clear bars) and SCC-9 (black bars) cells, as assessed by ChIP; data are represented relative to chromatin sample used as input in corresponding IPs, differences in immunoprecipitated chromatin levels between the cell lines were analyzed two-way ANOVA followed by Bonferroni post-test: A to D: Levels of the H3K4me3 marks at HOX genes expressed at higher levels in SCC-9 vs. OKF6-TERT1R cells; E to H: Levels of the H3K4me3 at HOX genes expressed at lower levels in SCC-9 vs.

OKF6-TERT1R cells; I and J: Negative control: chromatin immunoprecipitated using normal rabbit IgG; sequences amplified are indicated by the graph legend. **(b)** levels of the H3K36me3 epigenetic marks proximal to homeobox genes expressed differentially between OKF6-TERT1R (clear bars) and SCC-9 (black bars) cells, as assessed by ChIP; data are represented relative to chromatin sample used as input in corresponding IPs, differences in immunoprecipitated chromatin levels between the cell lines were analyzed two-way ANOVA followed by Bonferroni post-test: A to D: Levels of the H3K36me3 marks at HOX genes expressed at higher levels in SCC-9 vs. OKF6-TERT1R cells; E to H: Levels of the H3K36me3 at HOX genes expressed at lower levels in SCC-9 vs. OKF6-TERT1R cells; I and J: Negative control: chromatin immunoprecipitated using normal rabbit IgG; sequences amplified are indicated by the graph legend. Note differences in scales of y-axes in different graphs.

Figure 4. DNA methylation levels along annotated gene bodies and proximal promoter regions with at least a 20% point difference in methylation levels between OKF6-TERT1R and SCC-9 cells

Methylation levels expressed as % (see: Methods section) along annotated gene bodies (top panel) or proximal promoter regions ((defined as a 2000 bp sequence immediately upstream of the first TSS; bottom panel) with at least a 20 percent point difference in methylation levels between the OKF6-TERT1R and SCC-9 cells are shown in OKF6-TERT1R (x-axis) versus SCC-9 cells (y-axis). This shows the lower methylation levels along gene bodies and gene proximal promoter regions in SCC-9 as compared to OKF6-TERT1R cells.

Figure 5. Methylation of individual CpGs within genomic regions of homeobox genes expressed differentially between OKF6-TERT1R and SCC-9 cells (a and b) and entire HOX gene clusters (c)

Methylation expressed as % of detected methylated cytosine residues is shown as black dots for OKF6-TERT1R and as red dots for SCC-9 cells. Each dot represents an individual CpG dinucleotide. Genomic regions encoding the entire transcript (including UTRs) are indicated by light grey shading of the graph background; CpG islands, as defined by the UCSC Genome Browser, are indicated by pinstripe diagonal patterns in the graph background; (a) DNA methylation of HOX genes expressed at higher levels in SCC-9 vs. OKF6-TERT1R cells; (b) DNA methylation of homeobox genes expressed at lower levels in SCC-9 vs. OKF6-TERT1R cells; (c) DNA methylation of HOX gene clusters in SCC-9 and OKF6-TERT1R cells; locations of individual HOX genes within the cluster are indicated along the x- axis.

Table 1

Descriptive summary of ChIP results shown in Figures 1, 2 and 3.

Gene name	H3K9me3		H3K79me3		H3K4me3		H3K36me3	
	Fold increase OKF6- TERT1R /SCC-9	p value	Fold increase SCC-9/ OKF6- TERT1R	p value	Fold increase SCC-9/ OKF6- TERT1R	p value	Fold increase OKF6- TERT1R /SCC-9	p value
HOXB7	14.8	< 0.0001	N/D	-	25	< 0.001	2.7	< 0.0001
HOXC10	3.8	< 0.0001	N/D	-	2.6	< 0.01	9.6	< 0.001
HOXC13	6.2	< 0.01	N/D	-	3.3	< 0.0001	0.4	< 0.05
HOXD8	7.7	< 0.0001	N/D	-	3.0	< 0.0001	17.2	< 0.0001

Gene name	Fold increase SCC-9/ OKF6- TERT1R	p value	Fold increase OKF6- TERT1R /SCC-9	p value	Fold increase OKF6- TERT1R /SCC-9	p value	Fold increase OKF6- TERT1R /SCC-9	p value
	SCC-9/ OKF6- TERT1R	p value	Fold increase OKF6- TERT1R /SCC-9	p value	Fold increase OKF6- TERT1R /SCC-9	p value	Fold increase OKF6- TERT1R /SCC-9	p value
IRX1	2.7	< 0.05	4.6	n.s.	16.5	< 0.0001	5.9	< 0.0001
IRX4	1.0	ns	0.08	n.s.	0.9	n.s.	4.5	< 0.0001
SIX2	2.1	< 0.0001	N/D	-	6.0	< 0.0001	5.9	< 0.0001
TSHZ3	1.3	ns	N/D	-	1.5	n.s.	3.1	< 0.01

N/D not detected at levels above control IgG

Table 2

Methylation levels along gene body and proximal promoter regions for genes frequently methylated in oral squamous cell carcinoma.

Gene symbol	Gene name	Gene body ERRBS				Promoter ERRBS				References
		ML (OKF)	ML (SCC)	# CpG	FDR	ML (OKF)	ML (SCC)	# CpG	FDR	
AIM2	absent in melanoma 2	40.3	7.7	20	1.2E-87	<i>no data</i>				(Jithesh et al., 2013; Lleras et al., 2013; Poage et al., 2012; Poage et al., 2011)
CDH1	cadherin 1, type 1, E-cadherin (epithelial)	25.2	27.0	158	0.002	1.0	1.2	115	0.385	(Ha and Califano, 2006; Shaw et al., 2006; Steinmann et al., 2009)
CDKN2A	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	32.8	69.6	132 (OKF) 31 (SCC)	7.9E-153	56.4	76.7	23	1.8E-25	(Ha and Califano, 2006; Jithesh et al., 2013; Maruya et al., 2004; Steinmann et al., 2009; Su et al., 2010)
DAPK1	death-associated protein kinase 1	28.2	27.1	294	0.01	4.4	10.8	168	4.4E-97	(Ha and Califano, 2006; Jithesh et al., 2013; Steinmann et al., 2009)
DBC1	deleted in bladder cancer 1	14.4	26.1	112	3.1E-84	48.6	2.3	5	1.7E-115	(Ha and Califano, 2006)
DCC	deleted in colorectal carcinoma	31.6	2.9	107	0	1.5	11.6	21	1.9E-29	(Ha and Califano, 2006)
EBF3	early B-cell factor 3	28.1	61.7	706	0	33.9	59.14	17	2.6E-29	(Bennett et al., 2008)
GSTP1	glutathione S-transferase pi 1	0.3	1.8	95	8.9E-27	0.3	1.5	116	1.4E-22	(Steinmann et al., 2009)
HOXA11	homeobox A11	1.8	10.3	116	8.1E-168	2.1	11.3	51	4.8E-74	(Jithesh et al., 2013)
HOXA9	homeobox A9	12.1	70.5	174	0	6.3	18.9	64	5.9E-102	(Jithesh et al., 2013)
IRF8	interferon regulatory factor 8	37.4	26.4	159	1.1E-75	19.7	38.2	84	2.5E-111	(Hennessey et al., 2011)
IRX1	iroquois homeobox 1	45.9	72.9	247	0	2.7	66.1	318	0	(Bennett et al., 2008)
MAGEA3	melanoma antigen family A, 3	47.3	18.5	12	1.3E-33	54.1	17.7	2	0.001	(Hennessey et al., 2011; Lleras et al., 2013)

Gene symbol	Gene name	Gene body ERRBS				Promoter ERRBS				References
		ML (OKF)	ML (SCC)	# CpG	FDR	ML (OKF)	ML (SCC)	# CpG	FDR	
MGMT	O-6-methylguanine-DNA methyltransferase	39.4	6.2	629	0	23.1	28.1	80	6.9E-08	(Steinmann et al., 2009)
MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	19.5	20.0	42	0.66	1.6	1.1	43	0.065	(Steinmann et al., 2009)
MST1	macrophage stimulating 1 (hepatocyte growth factor-like)	21.6	15.1	40	1.5E-08	87.5	64.6	9	1.1E-20	(Steinmann et al., 2009)
P73	tumor protein p73 (TP73), transcript variant 10	39.8	43.5	775	6.0E-39	13.7	14.5	313	0.021	(Ha and Califano, 2006; Maruya et al., 2004)
RARB	retinoic acid receptor, beta	46.8	23.1	106	3.5E-183	11.1	20.2	24	7.6E-11	(Ha and Califano, 2006; Maruya et al., 2004; Shaw et al., 2006; Steinmann et al., 2009; Youssef et al., 2004)
RASSF1A	Ras association (RalGDS/AF-6) domain family member 1, (RASSF1), transcript variant A	4.0	4.6	211	0.01	8.4	3.6	26	0.00001	(Ha and Califano, 2006; Maruya et al., 2004; Steinmann et al., 2009)
RASSF5	Ras association (RalGDS/AF-6) domain family member 5	19.3	23.6	202	1.9E-19	22.6	5.2	108	4.4E-294	(Steinmann et al., 2009)
SERPINA12	serpin peptidase inhibitor, clade A (alpha-1 antitrypsin), member 12	26.7	9.3	23	2.4E-35	17.7	17.3	8	0.947	(Hennessey et al., 2011)
SERPINB5	serpin peptidase inhibitor, clade B (ovalbumin), member 5	61.2	83.9	32	1.3E-75	no data				(Ha and Califano, 2006)
TM4SF19	transmembrane 4 L six family member 19 (TM4SF19), transcript variant 1	15.3	13.5	97	0.002	3.1	5.7	34	0.00001	(Hennessey et al., 2011)
TRIM58	tripartite motif-containing 58	38.6	38.9	98	0.83	38.2	33.4	89	9.6E-08	(Hennessey et al., 2011)
ZNF671	zinc finger protein 671	3.4	49.9	31	0	100.0	98.6	1	0.419	(Hennessey et al., 2011; Lleras et al., 2013)

ML (methylation level)(**OKF**)- methylation levels expressed as % in OKF6-TERTIR cells; **ML (SCC)**- methylation levels expressed as % in SCC-9 cells; **#CpG**- number of CpG dinucleotides present within the region; **FDR**- False Discovery Rate expressed as a q value in the Benjamini and Hochberg adjusted Fisher exact test. The q value equals (i/m)Q and is calculated by assigning p-values from individual Fisher tests a rank in order, from smallest (rank of i=1) to largest (rank of i=m), m is the total number of tests and Q is the chosen false discovery rate.

Table 3

Gene body and promoter methylation data for homeobox genes with transcript levels higher (top) or lower (bottom) in SCC-9 than in OKF6-TERT1R cells (RNAseq, at least 3 fold transcript level differences).

#	Gene ID	Gene symbol	RNAseq		Gene level ERRBS				Promoter level ERRBS			
			FC	FDR (FC)	ML (OKF)	ML (SCC)	# CpG	FDR	ML (OKF)	ML (SCC)	# CpG	FDR
1	ENSG00000120093	HOXB3	2519	2.46E-34	5.2	69.2	555	0	13.4	78.4	41	0
2	ENSG00000108511	HOXB6	1799	5.63E-36	3.7	30.3	170	0	1.4	3.7	81	4.16E-17
3	ENSG00000123364	HOXC13	910	1.45E-64	3.6	16.0	157	3.23E-285	5.2	11.1	45	2.39E-16
4	ENSG00000123388	HOXC11	608	1.69E-42	5.8	10.6	151	1.92E-42	51.9	1.8	5	5.4E-110
5	ENSG00000180818	HOXC10	493	1.15E-58	7.6	7.1	43	0.532872	1.6	8.7	38	7.77E-35
6	ENSG00000180806	HOXC9	473	1.16E-38	5.2	10.6	164	5.55E-69	0.3	5.7	14	2.46E-20
7	ENSG00000120087	HOXB7	373	1.23E-64	2.5	20.6	539	0	9.1	77.0	47	0
8	ENSG00000253293	HOXA10	287	6.24E-30	20.9	28.7	195	5.16E-70	10.5	13.3	98	6.09E-09
9	ENSG00000172789	HOXC5	239	1.38E-27	6.6	33.7	738	0	10.7	63.6	54	0
10	ENSG00000170166	HOXD4	229	1.32E-45	11.3	54.5	83	0	22.7	72.6	22	4.0E-226
11	ENSG00000198353	HOXC4	218	2.02E-26	9.0	46.2	389	0	9.4	50.2	9	5.44E-58
12	ENSG00000128652	HOXD3	218	2.87E-41	7.1	48.6	615	0	0.6	27.8	60	0
13	ENSG00000152192	POU4F1	212	5.68E-37	3.4	12.8	168	3.08E-119	1.9	0.3	43	7.05E-08
14	ENSG00000258518	SHOX2	166	5.81E-23	8.0	34.0	171	0	11.2	27.9	26	1.00E-57
15	ENSG00000132130	LHX1	159	1.88E-35	4.4	40.7	218	0	1.4	5.7	109	5.58E-50
16	ENSG00000175879	HOXD8	155	1.86E-52	1.5	0.1	120	1.88E-26	2.2	26.8	39	9.1E-237
17	ENSG00000170689	HOXB9	107	2.11E-27	1.7	23.8	197	0	14.8	44.9	19	2.06E-49
18	ENSG00000168779	SHOX2	100	6.47E-33	8.8	30.9	184	0	16.0	80.8	72	0
19	ENSG00000106536	POU6F2	94	3.15E-22	38.9	7.3	177	0	3.4	9.4	14	5.30E-11
20	ENSG00000128710	HOXD10	94	8.40E-34	11.8	59.9	160	0	2.1	17.2	161	0
21	ENSG00000163623	NKX6-1	94	2.99E-21	7.0	29.3	273	0	5.4	41.2	70	0
22	ENSG00000134138	MEIS2	54	9.08E-34	9.6	23.7	172	5.97E-160	0.5	0.1	52	0.032018
23	ENSG00000128713	HOXD11	38	1.54E-19	2.5	29.1	233	0	2.2	21.0	189	0
24	ENSG00000144355	DLX1	35	1.31E-24	4.9	47.7	112	0	15.4	7.4	54	2.38E-30
25	ENSG00000119042	SATB2	10	1.47E-23	5.7	22.1	607	0	14.7	80.0	7	1.0E-101

#	Gene ID	Gene symbol	RNAseq		Gene level ERRBS				Promoter level ERRBS			
			FC	FDR (FC)	ML (OKF)	ML (SCC)	# CpG	FDR	ML (OKF)	ML (SCC)	# CpG	FDR
26	ENSG00000180438	TPRXL	5	2.05E-07	66.6	70.9	33	0.003211	0.0	15.0	1	0.014093
27	ENSG00000119547	ONECUT2	4	1.40E-06	8.5	59.2	269	0	1.6	0.3	99	5.95E-15
28	ENSG000001106852	LHX6	3	3.59E-05	9.4	26.3	408	0	<i>no data</i>			
29	ENSG00000105717	PBX4	3	0.001775	54.9	68.6	123	4.64E-64	44.3	68.7	32	1.27E-64

#	Gene ID	Gene symbol	RNAseq		Gene level ERRBS				Promoter level ERRBS			
			FC	FDR (FC)	ML (OKF)	ML (SCC)	# CpG	FDR	ML (OKF)	ML (SCC)	# CpG	FDR
1	ENSG00000170549	IRX1	338	4.00E-76	45.9	72.9	247	0	3.8	63.9	211	0
2	ENSG00000105419	MEIS3	111	6.92E-75	18.8	14.6	203	1.91E-20	3.7	43.3	30	4.66E-269
3	ENSG00000170577	SIX2	105	5.70E-72	8.7	28.1	122	2.48E-223	2.0	2.7	82	0.026130
4	ENSG00000182463	TSHZ2	91	3.06E-46	40.8	10.8	224	0	0.8	3.0	33	6.78E-08
5	ENSG00000185630	PBX1	46	2.34E-30	49.7	13.1	154	0	48.3	0.6	6	7.63E-32
6	ENSG00000121297	TSHZ3	22	6.44E-50	40.2	35.3	254	1.15E-21	84.6	75.3	14	1.32E-06
7	ENSG00000113430	IRX4	20	3.67E-48	22.4	25.0	532	5.47E-19	10.2	8.6	82	0.001055
8	ENSG0000028277	POU2F2	18	2.16E-29	35.5	29.3	277	9.18E-35	0.4	1.0	40	0.014617
9	ENSG00000126778	SIX1	16	7.47E-28	11.4	23.2	277	2.6834E-319	0.4	9.3	27	2.85E-64
10	ENSG00000148516	ZEB1	9	2.49E-17	12.3	7.5	312	3.51E-52	28.5	33.8	9	0.028639
11	ENSG00000184271	POU6F1	7	1.49E-13	25.7	25.9	127	0.8161854	7.6	10.0	71	6.11E-04
12	ENSG00000143995	MEIS1	6	5.45E-20	7.5	63.4	258	0	90.2	79.3	4	0.003921
13	ENSG00000159556	ISL2	6	4.10E-10	27.3	43.7	234	1.67E-181	33.3	43.6	39	6.56E-15
14	ENSG00000171476	HOPX	5	3.90E-11	9.3	2.6	125	2.05E-99	74.3	64.3	6	0.002940
15	ENSG00000169554	ZEB2	4	1.40E-06	10.5	22.9	239	1.50E-198	<i>no data</i>			
16	ENSG00000182568	SATB1	3	1.93E-06	5.6	5.0	309	0.0113757	0.2	0.6	109	3.40E-04
17	ENSG0000016082	ISL1	3	6.91E-06	75.1	31.9	90	0	6.4	0.5	51	7.49E-47
18	ENSG00000064195	DLX3	3	0.005921	6.3	6.4	131	0.791207	0.8	0.2	70	5.14E-08
19	ENSG00000167081	PBX3	3	2.55E-10	15.1	12.5	257	.43E-14	0.6	0.3	350	1.32E-06

Unfiltered methylation data are included for all differentially expressed homeobox genes. **FC**- transcript Fold Change in SCC-9 vs. OKF6-TERT1R cells (top) or in OKF6-TERT1R vs. SCC-9 cells (bottom). **FDR (FC)** - False Discovery Rate (for RNAseq analysis) expressed as the q value in the Benjamini and Hochberg adjusted t-test. **ML (OKF)**- methylation levels expressed as % in OKF6-TERT1R cells; **ML (SCC)**- methylation levels expressed as % in SCC-9 cells; **#CpG**- number of CpG dinucleotides present within the region; **FDR**- False Discovery Rate expressed as the q value in the Benjamini and Hochberg adjusted Fisher exact test.