Identification of HOX Signatures Contributing to Oral Cancer Phenotype

Kanaka Padam

Manipal Academy of Higher Education

Richard Morgan

University of West London

Keith Hunter

University of Sheffield

Sanjiban Chakrabarty

Manipal Academy of Higher Education

Naveena Kumar

Manipal Academy of Higher Education

Raghu Radhakrishnan (✉ raghu.ar@manipal.edu)

Manipal Academy of Higher Education

Research Article

Keywords: HOX, Homeobox, phenotype, oral cancer, dysplasia, temporospatial

Posted Date: December 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1189533/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: Evolutionarily conserved homeobox-containing *HOX* genes as transcriptional regulators in the developmental specification of organisms is well known. The contribution of *HOX* genes involvement in oral cancer phenotype has yet to be fully ascertained.

Methods: GEO datasets (GSE72627, GSE30784, GSE37991) were accessed and analyzed using GEO2R. TCGA-HNSC HTSeq-counts and clinical data were retrieved from the GDC portal for oral cavity neoplasms. Differential *HOX* gene expression was profiled using the DESeq2 R package with a log2 fold change cut-off (-1 and +1) and Benjamini-Hochberg *p*-adjusted value at <0.01. Gene set over-representation analysis and semantic analysis associated with the disease ontology were performed using ClusterProfiler R package and pathway over-representation analysis was performed using IMPaLa. *HOX* protein interaction network was constructed using the Pathfind R package. *HOX* phenotype associations were performed using Mammalian Phenotype Ontology, Human Phenotype Ontology, PhenGenI associations, Jensen tissues, and OMIM entries. Drug connectivity mapping was carried out with Dr. Insight R Package.

Results:

HOXB2 and *HOXA5* genes were upregulated in oral dysplasia but silenced during tumor progression. Loss of *HOXB2* expression was consistent through potentially malignant dysplastic oral lesions (PMOL) to primary tumor formation. *HOXA10*, *HOXB7*, *HOXC6*, *HOXC10* and *HOXD10* showed consistent upregulation from premalignancy to malignancy and were notably associated with risk factors. Overrepresentation analysis suggested *HOXA10* was involved in the transcriptional misregulation leading to oral cancer phenotype. *HOX* subnetwork analysis showed crucial interactions with cell cycle regulators, growth responsive elements, and proto-oncogenes.

Conclusion: Phenotype associations specific to the oral region involving *HOX* genes provide intrinsic cues to tumor development. The 5' *HOX* genes were aberrantly deregulated which reflects their posterior prevalence during oral carcinogenesis.

1. Introduction

HOX genes are a subset of homeobox genes, which function as transcriptional regulators specifying the anteroposterior (A-P) axis of the animal body plan and developmental organization of tissues and organs (Darbellay et al. 2019). The expression of *HOX* genes is coordinated from 3' to 5', temporally corresponding to their position along the A-P axis during vertebrate development. The 3' and 5' *HOX* genes are localized to the anterior and posterior tissues/regions, respectively (Shah and Sukumar 2010). A total of 39 *HOX* genes are segregated into four clusters: *HOXA* (7p15), *HOXB* (17q21.2), *HOXC* (12q13), and *HOXD* (2q31) (Holland 2013). The clustered topology of *HOX* genes in the mammalian genome corresponds to their magnitude and order of expression in a collinear fashion, which would be either regulated or coordinated (Gaunt 2015; Lewis 1978).

Patterns of disruption of *HOX* gene expression due to temporospatial deregulation was first described by Abate-Shen (Abate-shen 2002). The differential expression in tumor tissues has shown to be associated with a perturbation of normal organogenesis and differentiation. Epigenetic interplay involved in the process of gene transcription promoting neoplastic transformation may be the underlying factor leading to altered expression of *HOX* genes in tumor tissues (Abate-Shen 2002). Typically the 5' positioned *HOX* genes show a dominant phenotype due to posterior prevalence in disease states. These observations are evident in oesophageal squamous cell carcinoma where the normal foregut shows 3' *HOX* dominated expression which is lost in tumor tissues and instead exhibits a dominant expression of 5' *HOX* genes (Takahashi et al. 2007).

Several studies have shown that deregulated *HOX* expression in oral cancer (Abate-Shen 2002; Carrera et al. 2015; Dai et al. 2019; Hassan et al. 2006; Platais et al. 2016; Shah and Sukumar 2010) is either due to loss of tissue specificity or epigenetically-mediated loss of function (Abate-Shen 2002; Shah and Sukumar 2010; Xavier et al. 2014). The regulatory role of *HOX* genes in determining tumor characteristics and factors contributing to oral cancer phenotype, in particular, is the basis of this study. In this paper, the crucial interactions of *HOX* subnetworks were computationally analyzed to understand the role of deregulated *HOX* genes in transition to the oral tumor phenotype and identify potential therapeutic targets.

2. Methodology

2.1 Data acquisition:

Publicly available gene expression datasets deposited in the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) were accessed to query curated gene expression profiles. The following search terms “(oral cancer) OR dysplasia OR leukoplakia AND/OR transcriptome AND (areca nut/betel quid) OR smoking OR alcohol OR Tobacco” were used. The GEO datasets, GSE72627 consisted of paired normal and tumor samples (n=3) derived from three patients (Zhang et al. 2016). GSE30784 included 45 normal, 17 dysplasia, and 167 oral cancer samples (C. Chen et al. 2008) and GSE37991 consisted of paired normal and oral tumor samples (n=40) associated with risk factors such as smoking tobacco, drinking alcohol and chewing areca nut (Lee et al. 2015).

Oral cancer-causing genomic alterations, cataloged in The Cancer Genome Atlas (TCGA) (Wang, Jensen, and Zenklusen 2016) oral cancer datasets consisted of 18 normal and 294 primary tumor cases. HTSeq-counts data for RNA-Seq data was accessed and downloaded from Genomic Data Commons (GDC) repository (<https://portal.gdc.cancer.gov/>). The 10th revision of the International Classification of Diseases (ICD-10) was used to define cancer of the mouth, which included malignancy of the lip (C01); the base of the tongue (C02); other/unspecified parts of the tongue (C03); gingiva (C03), the floor of the mouth (C04), the palate (C05); other/unspecified parts of the mouth (C06) and oropharynx (C11). Cancers of the parotid gland and pharynx were excluded. Patients with insufficient or missing clinicopathological data were excluded from the subsequent analysis. The workflow employed has been illustrated in Fig. 1.

2.2 Identification of Differential Expression of HOX genes:

The GSE30784 and GSE37991 oral cancer datasets were analyzed for profiling the differential *HOX* gene expression using the interactive web tool GEO2R, (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) to compare two or more groups of samples (Edgar, Domrachev, and Lash 2002). The transcriptome profiles of paired normal and moderate dysplasia samples (n=3) derived from three patients were analyzed using “ggpubr” for data visualization in R. The paired analysis specific to *HOX* genes was illustrated using “ggboxplot” and “ggpaired” functions. Genome-wide annotation for the HTseq-counts data derived from TCGA oral cancer cohort was analyzed using org.HS.eg.db, Bioconductor version: Release (3.13) and R package (v4.1) (R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>). Quality control, normalization and differential expression of the *HOX* genes in the sample cohorts were performed using DESeq2 package (Love, Huber, and Anders 2014), which uses negative binomial distribution to model RNA-seq counts data. A log₂ fold change of +1 and -1 was maintained as a threshold with 1% FDR corrected with the Benjamini-Hochberg procedure to determine the differentially expressed *HOX* genes. An adjusted *p*-value of <0.01 was considered to be statistically significant.

2.3 Oral cancer cell line analysis:

The RNA-seq datasets (DepMap, Broad (2021): DepMap 21Q2 Public) were accessed and downloaded from the Dependency Map portal (<https://depmap.org/portal/>) to analyze the expression patterns and regulation of *HOX* genes in oral cancer-derived cell lines. A total of 39 oral cancer cell lines were screened for the expression patterns of *HOX* genes. The cell lines with missing clinical characteristics were excluded from the subsequent analysis. A list of cell line-specific gene expressions of *HOX* genes in oral cancer cell lines profiled in this study along with the biological origin of cell lines is provided as Supplementary File S2.

2.4 Gene set over-representation analysis:

Over-representation analysis (ORA) was performed to identify the biological process and the semantic similarities of disease ontology (DO) associated with the significant differentially expressed *HOX* genes in oral cancer using clusterProfiler (Yu et al. 2012, 2015). The categories relevant to the study were filtered (adjusted *p*-value<0.05) and depicted as a heatmap (Fig. 5 (a-b)). Further, the pathway over-representation analysis was performed using IMPaLa (Integrated Molecular Pathway Level analysis) (Kamburov et al. 2011) to identify the list of pathways involved. A *p*-value of <0.05 was considered statistically significant (Table 2).

Table 2
HOX genes involved in the over represented pathways using integrated molecular pathway level analysis (ImPaLa) tool

Pathway	Source	Genes involved	p-value
Differentiation of white and brown adipocyte	Wikipathways	<i>HOXC8; HOXC9</i>	0.000181
btg family proteins and cell cycle regulation	BioCarta	<i>HOXB9</i>	0.00735
Transcriptional misregulation in cancer	KEGG	<i>HOXA10; HOXA11</i>	0.0102
Signaling events mediated by HDAC Class III	PID	<i>HOXA10</i>	0.0315
keratinocyte differentiation	BioCarta	<i>HOXA7</i>	0.0426

2.5 HOX subnetwork analysis:

Underlying disease states were further predicted by constructing a protein interaction network (PIN). Anticipated protein-protein interaction (PPI) information was complemented using STRING (Szklarczyk et al. 2019) and genes prioritized for functional assays using GeneMania (Warde-Farley et al. 2010). To identify the distinct active subnetwork(s) associated with the HOX, pathfindR (Ulgen, Ozisik, and Sezerman 2019) package was used. Further, the subnetwork-oriented pathway enrichment analysis was performed using KEGG to exploit the disease alterations by this subset of genes in interaction with HOX using KEGG pathways.

2.6 HOX phenotypic and disease associations:

The deregulated HOX genes were further processed downstream to predict their phenotypic associations with oral cancer development. Data and text mining was performed for integration of the query *HOX* genes to identify their association with phenotypes accessing the Mammalian Phenotype Ontology (Smith and Eppig 2009), Human Phenotype Ontology (HPO) (Köhler et al. 2021), PhenGenI associations, Jensen diseases (Pletscher-Frankild et al. 2015), Jensen tissues (Palasca et al. 2018) and OMIM (Amberger et al. 2019) using Enrichr (E. Y. Chen et al. 2013)(Kuleshov et al. 2016). It was further supplemented with harmonizome (Rouillard et al. 2016), an integrated analysis tool about genes and proteins for massive mining of the publicly available RNA-seq data, ARCHS4 (Lachmann et al. 2018) resource. The biomolecular-phenotype network of association of the *HOX* genes and phenotypes specific to the oral cavity has been illustrated using Cytoscape (Shannon et al. 2003).

2.7 Drug Connectivity Mapping:

Drug connectivity mapping was carried out using CMap (Lamb et al. 2006) drug rank matrix dataset and the query dataset matrix containing significantly differentially expressed *HOX* genes with t-test statistic score computed using DESeq2 (Love, Huber, and Anders 2014). Drug identification analysis was performed in an attempt to identify the drugs that could reverse the query disease phenotype (negative connectivity) following the perturbation using Dr. Insight R package (Chan et al. 2019). The query results

were adjusted to and sorted by a *p*-value of <0.05 as a measure of significance and illustrated using Cytoscape (Shannon et al. 2003).

3. Results

3.1 Differentially expressed HOX genes show posterior prevalence in oral cancer:

Significantly upregulated and downregulated *HOX* genes through the oral cancer progression were identified (Table 1 and Fig. 2(a-b)) with a cut-off of 1% FDR. The complete list of *HOX* genes analyzed using TCGA oral cancer datasets and GEO datasets has been provided in supplementary file S1. *HOX* gene expression from normal and oral dysplasia conditions was represented as paired box plots connecting the paired data points across the samples analyzed (Fig. 3(a-b)).

Table 1

List of differentially upregulated and downregulated *HOX* genes in a panel of normal, dysplastic and primary tumor through the disease progression of oral cancer

Dysplasia versus Normal	
Upregulated <i>HOX</i> genes	<i>HOXA2, HOXA5, HOXA10, HOXB7, HOXC6, HOXC10, HOXD10</i>
Downregulated <i>HOX</i> genes	<i>HOXB2</i>
Tumor vs normal (Habituated with risk factors)	
Upregulated <i>HOX</i> genes	<i>HOXA1, HOXA3, HOXA5, HOXA9, HOXA11, HOXA13, HOXB3, HOXB5, HOXB7, HOXC6, HOXC8, HOXC9, HOXC10, HOXC13, HOXD9, HOXD10, HOXD11, HOXD13</i>
Downregulated <i>HOX</i> genes	-
Primary tumor versus Normal	
Upregulated <i>HOX</i> genes	<i>HOXA1, HOXA6, HOXA7, HOXA10, HOXA11, HOXA13, HOXB7, HOXB9, HOXC4, HOXC6, HOXC8, HOXC9, HOXC10, HOXC11, HOXD10, HOXD11, HOXD13</i>
Downregulated <i>HOX</i> genes	<i>HOXB2, HOXB4</i>

HOXB2 and *HOXA5* expression was upregulated in PMOL but their expression was lost during primary tumor progression. *HOXA10, HOXB7, HOXC6, HOXC10*, and *HOXD10* showed consistent upregulation from premalignancy to malignancy which indicated the posterior prevalence of the *HOX* genes during cancer progression. However, *HOXB2* was downregulated in both dysplasia and primary tumor samples which was consistent with the disease progression. Among the cohort of patients who had a history of habits (areca nut chewing, smoking and alcohol consumption) upregulation of *HOXA1, HOXA11*,

HOXA13, *HOXB7*, *HOXC6*, *HOXC8*, *HOXC10*, *HOXD10*, *HOXD11* and *HOXD13* were notably increased through the progression from dysplasia to oral cancer (Fig. 3c). These findings confirmed that the differential expression of *HOX* genes was associated with the progression of oral cancer although the associated risk factors influenced the clinical outcome.

3.2 *HOX* gene expression in oral cancer cell lines:

The expression pattern of *HOX* genes screened in 39 oral cancer cell lines is provided as supplementary File S2. Further, the expression states of differentially expressed *HOX* genes in a panel of normal and oral tumor specimens were compared with the oral cancer cell lines for translating and validating the *in silico* data with the *in vitro* data. The expression patterns were diverse across each cell line screened. This was perhaps due to the cancer cell heterogeneity with respect to the primary site of the tumor and its biological behavior. However, the average expression profile of the *HOX* genes in the cell lines when pooled together were similar to those of the patients' tissue panel validated experimentally (Fig. 4).

3.3 *HOX* genes as developmental cues in oral carcinogenesis:

The differentially expressed *HOX* genes had hits against the developmentally related events such as epithelial morphogenesis, monocyte differentiation including cell fate specification and commitment (Fig. 5a). Any alteration in the functional state had pathological consequences leading to phenotypic abnormalities. *HOXA10* participated in the transcriptional misregulation in cancer (Table 2) influencing upstream and downstream target interactions as reported previously (Padam et al. 2021). *HOXB9* is involved in the regulation of the *btg* family of proteins whose deregulation has implicit biological consequences during cancer progression (Yuniati et al. 2019) and was also noted to function as an angiogenic switch along with *HOXB7* (Fig. 5b). The other posterior *HOX* genes such as *HOXD10*, *HOXD11* were noted to be associated with the mouth neoplasm based on the semantic analysis performed using DO (Fig. 5b). Altered expression of *HOXA11*, *HOXC4* and *HOXC6* was associated with smoking/tobacco habits (Fig. 5b) and their expression varied in patients with positive habit history (Fig. 3c).

3.4 *HOX* showed interactions with notable key regulators:

HOX subnetwork analysis showed crucial interactions with cell cycle regulators such as Geminin (GMNN), CDKN2A (cyclin-dependent kinase inhibitor 2A), cell division cycle (CDC) associated proteins such as CDC6, 14A, 20, 27, transcriptional regulators like TBX4 (T-box transcription factor 4), FOXO1 (Forkhead box protein O1), FOXC1 (Forkhead box C1), EGR (early growth response), RARA (retinoic acid receptor alpha), proto-oncogenes such as JUN, HRAS, BRAF, markers such as PD-L1 (programmed cell death ligand 1), SMAD family, FGF10 and other homeobox families of genes such as POU2F1, PITX2, PDX1, MEIS, and PBX (Fig. 6a). The downstream functional KEGG pathway enrichment analysis revealed their involvement in various cancer-related pathways (Fig. 6b) and the differential expression of these genes in KEGG pathways was also identified (Fig. 6c).

3.5 HOX gene deregulation contributes to the oral carcinoma phenotype:

Based on the phenotypic analysis of differentially regulated *HOX* genes, *HOXA6*, *HOXA10*, *HOXC6*, *HOXC8*, *HOXC10*, *HOXC11*, and *HOXB9* were involved in cancer development of the oral cavity during the onset of carcinogenesis. *HOXC6* was differentially upregulated at the time of the onset of dysplasia, which however had a role in the development of secondary malignant neoplasm of the lymph node. On other hand, deregulated *HOXB4* has been associated with abnormal development of the nasopharynx and oral frenula. These findings indicate that deregulation in the expression of homeobox genes contributes to the carcinoma phenotype as an effect of disruption in developmental coordination during the onset of oncogenesis. This reflects the need for normal expression of these *HOX* genes localized to their developmental sites. It could also be inferred that several deregulated *HOX* genes have different phenotypic alterations during their action (Fig. 7a). These findings indicate that the role of *HOX* genes in oral carcinogenesis is tightly modulated and transcriptionally active as previously reported (Padam et al. 2021).

3.6 Identification of HOX proteins as potential therapeutic targets:

HOXB2, *HOXA10*, *HOXC10* and *HOXD11* were noted to be the central hub *HOX* genes involved in promoting oral tumor phenotype (Fig. 7a) amenable to the antagonistic effect of the drugs (Fig. 7b). The wavy arrows indicate the drugs that upregulate the target genes whereas the inhibitory arrows are the drugs that downregulate the target genes. The queried drug-gene interactions ($p < 0.05$) could potentially be used as a part of a chemotherapy regime to reverse the phenotypic abnormalities associated with the oral tumor development driven by the deregulation of *HOX* genes. The mechanistic action and insights into the queried drugs reported are discussed further.

4. Discussion

HOX genes are well established in specifying the developmental states whose deregulation has been noted to be involved in carcinogenesis. Our analysis showed that these deregulated *HOX* genes which have been described as being involved in developmental aberrations could lead to morphogenetic changes during oral carcinogenesis (Fig. 5(a-b)). Moreover, *HOX* genes are epigenetically regulated via DNA methylation and altered histone modifications, which when dysregulated could potentially be driving the normal cell towards the neoplastic phenotype (Marcinkiewicz and Gudas 2014; Xavier et al. 2014).

The occurrence of oral squamous cell carcinoma has been characterized as a step-wise process in which the normal oral healthy mucosa, after prolonged carcinogenic influences, undergoes a series of changes to develop into the primary invasive tumor (Califano et al. 1996). Some studies have shown that *HOXA10* functions in the regulation of proliferation, migration, and invasion and has been reported as a less aggressive tumor phenotype (Carrera et al. 2015) whereas *HOXC10*, which regulates oral tumorigenesis

through Wnt-EMT signaling pathways, might play a pivotal role in metastasis of OSCC as studied in an *in vivo* xenograft model (Dai et al. 2019). *HOXD10* expression is different across the stages of tumor progression inducing invasion with its overexpression and causing reciprocal effects with a knockdown in oral cancer cell lines (Hakami et al. 2014).

In our analysis, *HOXA2*, *HOXA5*, *HOXA10*, *HOXB2*, *HOXB7*, *HOXC6*, *HOXC10* and *HOXD10* were differentially expressed in potentially malignant dysplastic oral lesions (PMOL) compared to the normal mucosa indicating their early responsive role in the malignant transformation to oral cancer. Of these, *HOXA10*, *HOXB2*, *HOXB7*, *HOXC10* and *HOXD10* were commonly dysregulated in both the PMOL and primary tumors (Fig. 2 (a-b) & 3 (a-c)) indicating the role of the posterior prevalent 5' *HOX* genes in the development of tumor phenotype.

HOX gene expression concerning oral cancer phenotype was noted to be varied among the subset of *HOX* genes screened. This could be attributable to the tumor heterogeneity and tissue specificity of *HOX* genes (Abate-Shen 2002; Shah and Sukumar 2010) causing locoregional aberrations during tumor development (Fig. 5a). Of the *HOX* genes analyzed, *HOXA6*, *HOXA10*, *HOXC6*, *HOXC10* and *HOXC11* have a common phenotypic association in relation to the oral mucosa. This confirms the reasons for morphological alterations noted in the oral cavity during tumorigenesis due to altered *HOX* gene expression (Fig. 7a).

Further, the subnetwork driven regulatory analysis showed *HOX* interactions with the cell cycle regulators (*GMNN*, *CDKN2A*, *CDC* associated proteins), other families of homeobox genes (*TBX4*, *FOXO1*, *FOXC1*, *MEIS*, *PBX*, *POU2F1*) and the proto-oncogenes (*JUN*, *HRAS*, *BRAF*) leading to the alterations of various cancer related signaling pathways (Fig. 6 (a-c)). It was reported that *POU2F1*, identified through the subnetwork analysis, regulates both *HOXD10* and *HOXD11* activity which drives the proliferative and invasive phenotype (Sharpe et al. 2014).

Cell lines are crucial for performing the *in vitro* validation of specific cancer-associated genes, which could be potential diagnostic markers and therapeutic targets. The choice of a suitable cell line with a careful interpretation of the clinical data is necessary to derive clinically significant results from the *in vitro* work. In addition to the clinical studies and descriptive studies of samples, assumptions derived about cancer can be tested in cell lines of interest, which are originally isolated from carcinoma patients and also from transgenic models via genetic manipulation. To understand the key *HOX* genes among the cluster that are deregulated from their normal regulation patterns during oral cancer disease progression, we selected 39 oral cancer cell lines based on their characteristic features (Supplementary File S2, Fig. 4) for screening *HOX* expression patterns in comparison to the patient samples.

Among the drugs screened to reverse the acquired phenotype (Fig. 7b), MG-262 (Z-Leu-Leu-Leu-B(OH)₂), a proteasome inhibitor, was reported to be involved in cell growth arrest by promoting the expression of cell cycle inhibitors (p21 and p27), and by driving cell death through the activation of mitogen-activated protein kinase phosphatase 1 (MAPK1) and c-Jun phosphorylation (Wong et al. 2008). Chlorzoxazone, a relaxant, was reported to function in the alteration of Ca⁺² signaling and cell viability by promoting Ca⁺²

independent cell death in human oral cancer cells (Lu et al. 2019). A study (Lippman et al. 1990) reported that low doses of isotretinoin (13-cis retinoic acid) resulted in the regression of potentially malignant oral lesions (PMOL). Furosemide, a modulator of cellular pumps, was potentially found to be an effective therapeutic regimen that reverses the multidrug resistance in cancer cells and could be used as an adjunct in cancer therapy (Speers et al. 2006). S-Propranolol (PRO), a non-selective beta-adrenergic receptor antagonist (beta-blocker) could be a novel adjunctive treatment for HNSCC as it has been shown to inhibit proliferation, invasion, and angiogenesis, and modulate tumor cell sensitivity (Wolter et al. 2012). Alvespimycin was reported to function as an antitumor agent which sequesters the target proteins mediated through HSP90 inhibition directing selective proteasomal degradation of BRAF, a proto-oncogene (Pacey et al. 2011). GW-8510, also belonging to the class of antineoplastics, acts as a cyclin-dependent kinase inhibitor (Hsieh et al. 2016). These repurposed drugs queried from CMap dataset could be used as adjuncts of the therapeutic regimen and/or for targeted therapy of *HOX* genes studied.

5. Conclusion

The posterior prevalent 5' *HOX* genes were expressed more commonly in oral tumor samples compared to the anterior *HOX* genes, reflecting the loss of expression of anterior *HOX* genes upon the onset of tumorigenesis and deregulation of temporospatial patterning. The development of suitable orthotopic models targeting the specific *HOX* genes identified here would help facilitate future functional studies. Several *HOX* genes, including *HOXA10*, *HOXB2*, *HOXC10*, and *HOXD11*, could serve as potential therapeutic targets, which would reverse the oral tumor phenotype.

Declarations

Funding

This work was supported by the DBT/Wellcome Trust India Alliance Fellowship, (Grant number - IA/CPHI/18/1/503927) awarded to Raghu Radhakrishnan.

Conflict of Interest: The authors state that they have no conflict of interest.

Contributions

Conceptualization, methodology, formal analysis and investigation: Kanaka Sai Ram Padam, Raghu Radhakrishnan; Writing – original draft preparation: Kanaka Sai Ram Padam; Writing – review and editing: Raghu Radhakrishnan, Sanjiban Chakrabarty, Keith Hunter, Richard Morgan, Naveena A N Kumar; Clinical interpretation: Naveena A N Kumar; Funding acquisition: Raghu Radhakrishnan; Supervision: Raghu Radhakrishnan. All authors read and approved the final manuscript.

Data Availability: All the data used in the present study are freely available for the research community to access from publicly archived datasets, analyzed and/ generated from DepMap repository (<https://depmap.org/portal/download/>), Genomic Data Commons portal (<https://portal.gdc.cancer.gov/>),

Gene Expression Omnibus Datasets (<https://www.ncbi.nlm.nih.gov/geo/>) - GSE72627, GSE30784, GSE37991).

Code Availability: All source codes used are publicly available on GitHub for the research community to access and reproduce the data using DepMap (https://github.com/broadinstitute/depmap_omics), clusterProfiler (<https://github.com/YuLab-SMU/clusterProfiler>), pathfinder (<https://github.com/egeulgen/pathfindR>), DrInsight (<https://github.com/cran/DrInsight>) and Cytoscape (<https://github.com/cytoscape/cytoscape.js>).

Ethics Approval: Data was freely available from the public domains which are properly anonymised and informed consent was obtained at the time of data collection did not require any ethical approval.

References

1. Abate-shen, Cory. 2002. "Deregulated Homeobox Gene Expression in Cancer: Cause or Consequence?" *Nat Rev Cancer* 2(October): 777–85.
2. Abate-Shen, Cory. 2002. "Deregulated Homeobox Gene Expression in Cancer: Cause or Consequence?" *Nature reviews. Cancer* 2(10): 777–85.
3. Amberger, Joanna S, Carol A Bocchini, Alan F Scott, and Ada Hamosh. 2019. "OMIM.Org: Leveraging Knowledge across Phenotype-Genes Relationships." *Nucleic acids research* 47(D1): D1038–43.
4. Califano, J et al. 1996. "Genetic Progression Model for Head and Neck Cancer: Implications for Field Cancerization." *Cancer research* 56(11): 2488–92.
5. Carrera, Manoela et al. 2015. "HOXA10 Controls Proliferation, Migration and Invasion in Oral Squamous Cell Carcinoma." *International journal of clinical and experimental pathology* 8(4): 3613–23.
6. Chan, Jinyan et al. 2019. "Breaking the Paradigm: Dr Insight Empowers Signature-Free, Enhanced Drug Repurposing." *Bioinformatics* 35(16): 2818–26.
7. Chen, Chu et al. 2008. "Gene Expression Profiling Identifies Genes Predictive of Oral Squamous Cell Carcinoma." *Cancer Epidemiol Biomarkers Prev* 17(8): 2152–62.
8. Chen, Edward Y et al. 2013. "Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool." *BMC bioinformatics* 14: 128.
9. Dai, Bo Wen et al. 2019. "HOXC10 Promotes Migration and Invasion via the WNT-EMT Signaling Pathway in Oral Squamous Cell Carcinoma." *Journal of Cancer* 10(19): 4540–51.
10. Darbellay, Fabrice et al. 2019. "The Constrained Architecture of Mammalian Hox Gene Clusters." *Proceedings of the National Academy of Sciences* 116(27): 13424 LP – 13433.
11. Edgar, Ron, Michael Domrachev, and Alex E Lash. 2002. "Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository." *Nucleic acids research* 30(1): 207–10.
12. Gaunt, Stephen J. 2015. "The Significance of Hox Gene Collinearity." *The International journal of developmental biology* 59(4–6): 159–70.

13. Hakami, F et al. 2014. "The Roles of HOXD10 in the Development and Progression of Head and Neck Squamous Cell Carcinoma (HNSCC)." *British Journal of Cancer* 111(4): 807–16.
14. Hassan, Nur Mohammad Monsur et al. 2006. "Aberrant Expression of HOX Genes in Oral Dysplasia and Squamous Cell Carcinoma Tissues." *Oncology research* 16(5): 217–24.
15. Holland, Peter W H. 2013. "Evolution of Homeobox Genes." *WIREs Developmental Biology* 2(1): 31–45.
16. Hsieh, Y-Y, C-J Chou, H-L Lo, and P-M Yang. 2016. "Repositioning of a Cyclin-Dependent Kinase Inhibitor GW8510 as a Ribonucleotide Reductase M2 Inhibitor to Treat Human Colorectal Cancer." *Cell death discovery* 2: 16027.
17. Kamburov, Atanas et al. 2011. "Integrated Pathway-Level Analysis of Transcriptomics and Metabolomics Data with IMPaLA." *Bioinformatics* 27(20): 2917–18.
18. Köhler, Sebastian et al. 2021. "The Human Phenotype Ontology in 2021." *Nucleic acids research* 49(D1): D1207–17.
19. Kuleshov, Maxim V et al. 2016. "Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update." *Nucleic acids research* 44(W1): W90-7.
20. Lachmann, Alexander et al. 2018. "Massive Mining of Publicly Available RNA-Seq Data from Human and Mouse." *Nature Communications* 9(1): 1–10.
21. Lamb, Justin et al. 2006. "The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease." *Science (New York, N.Y.)* 313(5795): 1929–35.
22. Lee, Chia-Huei et al. 2015. "IL-1 β Promotes Malignant Transformation and Tumor Aggressiveness in Oral Cancer." *Journal of cellular physiology* 230(4): 875–84.
23. Lewis, E B. 1978. "A Gene Complex Controlling Segmentation in Drosophila." *Nature* 276(5688): 565–70.
24. Lippman, S M et al. 1990. "Low Dose 13-Cisretinoic Acid (13cRA) Maintains Remission in Oral Premalignancy: More Effective than b-Carotene in Randomized Trial." In *Proc Am Soc Clin Oncol*, , 59.
25. Love, Michael I, Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15(12): 550.
26. Lu, Ti et al. 2019. "Action of Chlorzoxazone on Ca(2)(+)Movement and Viability in Human Oral Cancer Cells." *The Chinese journal of physiology* 62(3): 123–30.
27. Marcinkiewicz, Katarzyna M, and Lorraine J Gudas. 2014. "Altered Histone Mark Deposition and DNA Methylation at Homeobox Genes in Human Oral Squamous Cell Carcinoma." *Journal of cellular physiology* 229(10): 1405–16.
28. Pacey, Simon et al. 2011. "A Phase I Study of the Heat Shock Protein 90 Inhibitor Alvospimycin (17-DMAG) given Intravenously to Patients with Advanced Solid Tumors." *Clinical cancer research: an official journal of the American Association for Cancer Research* 17(6): 1561–70.
29. Padam, Kanaka Sai Ram et al. 2021. "In Silico Analysis of HOX-Associated Transcription Factors as Potential Regulators of Oral Cancer." *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*

132(1): 72–79.

30. Palasca, Oana et al. 2018. "TISSUES 2.0: An Integrative Web Resource on Mammalian Tissue Expression." *Database* 2018.
31. Platais, Christopher et al. 2016. "The Role of HOX Genes in Head and Neck Squamous Cell Carcinoma." *Journal of Oral Pathology and Medicine* 45(4): 239–47.
32. Pletscher-Frankild, Sune et al. 2015. "DISEASES: Text Mining and Data Integration of Disease–Gene Associations." *Methods* 74: 83–89.
33. Rouillard, Andrew D. et al. 2016. "The Harmonizome: A Collection of Processed Datasets Gathered to Serve and Mine Knowledge about Genes and Proteins." *Database: the journal of biological databases and curation* 2016.
34. Shah, Nilay, and Saraswati Sukumar. 2010. "The Hox Genes and Their Roles in Oncogenesis." *Nature Publishing Group* 10(MAY): 885–90.
35. Shannon, Paul et al. 2003. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome research* 13(11): 2498–2504.
36. Sharpe, Daniel J et al. 2014. "POU2F1 Activity Regulates HOXD10 and HOXD11 Promoting a Proliferative and Invasive Phenotype in Head and Neck Cancer." *Oncotarget* 5(18): 8803–15.
37. Smith, Cynthia L, and Janan T Eppig. 2009. "The Mammalian Phenotype Ontology: Enabling Robust Annotation and Comparative Analysis." *Wiley interdisciplinary reviews. Systems biology and medicine* 1(3): 390–99.
38. Speers, A G et al. 2006. "Furosemide Reverses Multidrug Resistance Status in Bladder Cancer Cells in Vitro." *Journal of clinical pathology* 59(9): 912–15.
39. Szklarczyk, Damian et al. 2019. "STRING V11: Protein–Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets." *Nucleic Acids Research* 47(D1): D607–13.
40. Takahashi, Osamu et al. 2007. "Dysregulated Expression of HOX and ParaHOX Genes in Human Esophageal Squamous Cell Carcinoma." *Oncology reports* 17(4): 753–60.
41. Ulgen, Ege, Ozan Ozisik, and Osman Ugur Sezer. 2019. "PathfindR: An R Package for Comprehensive Identification of Enriched Pathways in Omics Data Through Active Subnetworks ." *Frontiers in Genetics* 10: 858.
42. Wang, Zhining, Mark A Jensen, and Jean Claude Zenklusen. 2016. "A Practical Guide to The Cancer Genome Atlas (TCGA)." *Methods in molecular biology (Clifton, N.J.)* 1418: 111–41.
43. Warde-Farley, David et al. 2010. "The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function." *Nucleic Acids Research* 38(suppl_2): W214–20.
44. Wolter, Nikolaus E, Jennifer K Wolter, Danny J Enepekides, and Meredith S Irwin. 2012. "Propranolol as a Novel Adjunctive Treatment for Head and Neck Squamous Cell Carcinoma." *Journal of*

otolaryngology - head & neck surgery = Le Journal d'oto-rhino-laryngologie et de chirurgie cervico-faciale 41(5): 334–44.

45. Wong, Thian Sze et al. 2008. "Mature MiR-184 as Potential Oncogenic MicroRNA of Squamous Cell Carcinoma of Tongue." *Clinical Cancer Research* 14(9): 2588–92.
46. Xavier, Flavia Calo Aquino, Maria Fernanda de Souza Setubal Destro, Carina Magalhaes Esteves Duarte, and Fabio Daumas Nunes. 2014. "Epigenetic Repression of HOXB Cluster in Oral Cancer Cell Lines." *Archives of oral biology* 59(8): 783–89.
47. Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. "ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters." *Omics: a journal of integrative biology* 16(5): 284–87.
48. Yu, Guangchuang, Li-Gen Wang, Guang-Rong Yan, and Qing-Yu He. 2015. "DOSE: An R/Bioconductor Package for Disease Ontology Semantic and Enrichment Analysis." *Bioinformatics* 31(4): 608–9.
49. Yuniati, Laurensia, Blanca Scheijen, Laurens T van der Meer, and Frank N van Leeuwen. 2019. "Tumor Suppressors BTG1 and BTG2: Beyond Growth Control." *Journal of cellular physiology* 234(5): 5379–89.
50. Zhang, Jiali, Si Chen, Lan Yin, and Xinming Chen. 2016. "Transcriptome Profiles of Moderate Dysplasia in Oral Mucosa Associated with Malignant Conversion." *International Journal of Clinical and Experimental Pathology* 9(6): 6107–16.

Figures

Figure 1

Schematic workflow employed to identify the HOX involved with oral tumor phenotype

Figure 2

(a-b): a) Illustrative heatmap of *HOX* cluster gene expressions in a panel of normal (n=18) and tumor (n=294) specimens analyzed from the TCGA-HNSC oral cancer datasets. **b)** Differentially expressed *HOX* genes in primary tumors compared to the normal samples were identified using DESeq2 with a log2FC cut-off (-1 and +1) factoring 1% FDR to minimize false positives. Results are depicted in the form of a volcano plot.

Figure 3

(a-c): *HOX* genes showed deregulated expression pattern in oral dysplasia compared to the normal. **a)** *HOXA2*, *HOXA10* and *HOXC10* expression showed an increasing trend whereas *HOXB2* showed a decreased trend of expression during the early onset of dysplasia (n=3 paired samples). The connected lines represent the paired data points analyzed for each sample. **b)** *HOXB7*, *HOXC6* and *HOXD10* were shown to be significantly differentially expressed ($p<0.01$) in the GSE30784 dataset in a panel of 17 dysplasia and 45 oral mucosa tissues. **c)** *HOXA1*, *HOXA11*, *HOXA13*, *HOXB7*, *HOXC6*, *HOXC8*, *HOXC10*, *HOXD10*, *HOXD11* and *HOXD13* were identified to be significantly upregulated in the cohort of patients who are habituated with the risk factors, betel quid chewing, and smoking or alcohol consumption daily. These results implicate the role of external hazards in contributing to oral carcinogenesis.

Figure 4

Scattered dot plots illustrating the gene expression level of *HOX* expressed as a $\log_2(\text{TPM}+1)$ in a panel of 39 oral cancer cell lines screened. The expression states of *HOX* genes have been noted to be modestly varied across the cell lines which are in concordance with the clinical data alluding to their biological origin and genomic variation. The pooled mean average of the *HOX* genes in the cell lines is more or less alike with the differentially expressed *HOX* genes analyzed from the panel of patient tissues.

Figure 5

(a-b): Overrepresentation analysis of differentially expressed *HOX* genes using **a)** GO biological processes and **b)** disease ontology (DO). *HOXA11*, *HOXC4* and *HOXC6* have been noted to be associated with smoking. *HOXB7* and *HOXB9* may regulate the angiogenesis whereas *HOXD10*, which regulates the cell fate commitment and specification and *HOXD11*, are noted to be associated with the mouth neoplasm. *HOXA7* and *HOXB2* were involved in the epithelial morphogenesis whose expression was deregulated in the oral carcinoma which implicates the altered expression states of these *HOX* genes contribution to the downstream effects.

Figure 6

(a-c): **a)** *HOX* subnetwork analysis depicting the protein interactions of *HOX* downstream to the key cell cycle regulators (*GMNN*, *CDKN2A*, *CDC* group of proteins), proto-oncogenes such as *HRAS*, *BRAF*, *JUN* and *PD-L1*. **b)** Bubble chart depicting the fold enrichment of *HOX* associated subnetwork query using the KEGG pathways and **c)** the heatmap illustration of potential alterations noted in the KEGG pathways. Top 30 hits were displayed. These results implicate the downstream effects of *HOX* leading to the various alterations in the pathways studied.

Figure 7

(a-b): a) Disease-phenotype association of *HOX* genes pertaining to oral cancer. *HOXA2*, *HOXA6*, *HOXB7*, *HOXA10*, *HOXC10*, *HOXC11*, *HOXD11* and *HOXD13* have been shown to have common phenotypic associations pertaining to the oral mucosa. *HOXC6* was shown to be involved in regulating the metastatic phenotype whose expression was observed to be upregulated in oral cancer. Upregulated *HOX* genes in oral cancer are represented in red color whereas downregulated *HOX* genes are represented in green color. **b)** Drug gene interaction network computed using CMap drug rank matrix to identify the potential therapeutic targets. The interaction of drugs that represses the regulation of *HOXA10*, *HOXC10* and *HOXD11* was represented with an inhibition line type arrow whereas the drug that promotes the activity of *HOXB2* was represented with a curved arrow.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFileS1.docx](#)
- [SupplementaryFileS2.xlsx](#)