

Aus der Klinik für Hals-Nasen-Ohren-Heilkunde der Universität Heidelberg

(Geschäftsführender Direktor: Prof. Dr. med. Dr. h.c. Peter K. Plinkert)

Sektion für Experimentelle und Translationale Kopf-Hals Onkologie

(Sektionsleiter: Prof. Dr. rer. nat. Jochen Hess)

## **The Role of SOX2 and SOX9 as Regulatory Genes of the Head and Neck Squamous Cell Carcinoma Microenvironment**

Inauguraldissertation

zur Erlangung des Doctor scientiarum humanarum

an der

Medizinischen Fakultät Heidelberg

der

Ruprecht-Karls-Universität

vorgelegt von

Silvia Barbosa

aus

Porto Alegre, Brasilien

2022

Dekan: Herr Prof. Dr. med. Dr. h.c. Hans-Georg Kräusslich

Doktorvater: Herr Prof. Dr. rer. nat. Jochen Hess

*To my family with all my love.*

## **Table of Contents**

<b>Abbreviation List.....</b>	<b>9</b>
<b>1. Introduction.....</b>	<b>11</b>
1.1. Head and Neck Squamous Cell Carcinoma (HNSCC).....	11
1.2. SOX Family of Transcription Factors.....	14
1.3. Tumor Microenvironment in HNSCC.....	19
1.4. Treatment and Challenges of Resistance.....	20
<b>2. Aims of the Study.....</b>	<b>24</b>
<b>3. Material &amp; Methods.....</b>	<b>25</b>
3.1. Computational Data Resources.....	25
3.1.1. Validation Analyses.....	25
3.1.2. Expression and Clinical Data sets.....	25
3.1.3. Co-expression Analysis.....	25
3.1.4. Differential Gene Expression Analysis.....	25
3.1.5. Survival Analysis.....	26
3.1.6. Integrative Analysis.....	26
3.1.7. Clustering Analysis.....	26
3.1.8. QuPath Analysis.....	26
3.2. Experimental Resources.....	26
3.2.1. Human Head and Neck Squamous Cell Carcinoma (HNSCC) Cell Lines.....	26
3.2.2. Cellular Culture.....	27
3.2.3. Colony Formation Assay.....	27
3.2.4. Real-time Quantitative Polymerase Chain Reaction.....	28
3.2.4.1. RNA extraction.....	28
3.2.4.2. cDNA synthesis.....	28
3.2.4.3. SYBR™ green PCR.....	28
3.2.5. Western Blotting.....	29
3.2.5.1. Protein Extraction.....	29
3.2.5.2. SDS-PAGE (Sodiumdodecylsulfate polyacrylamide gel electrophoresis).....	29
3.2.5.3. Polyvinylidene Difluoride Membrane (PVDF).....	30
3.2.6. Immunofluorescence (IF).....	30
3.2.7. Immunohistochemistry (IHC).....	31
3.2.8. Dissociation of Tumor Samples.....	31
3.2.9. Flow Cytometry Analysis (FACS).....	32
3.2.10. Multiplex Cytokine Assay.....	32
3.2.11. 51Chromium Release Assay.....	32
3.2.12. Statistical Analysis.....	33
3.2.13. Study Approval and Patient Samples.....	33

<b>4. Results</b> .....	<b>34</b>
4.1. Survival and SOX2/SOX9 Gene Expression in HNSCC Patients .....	34
4.1.1. Inverse Association of SOX2 and SOX9 in HNSCC .....	34
4.1.2. Gene Network Based on the Inverse SOX2 and SOX9 Gene Expression .....	43
4.1.3. Differential Gene Expression Analysis .....	44
4.1.4. Integrative Gene Set Analysis and Final Signature Gene Set Based on Inverse SOX2 and SOX9 Expression.....	44
4.1.5. TGF- $\beta$ Signaling as one of The Top Canonical Pathways Regulated by Inverse SOX2/SOX9 Expression .....	46
4.2. Tumor Samples of The Heidelberg Center for Personalized Oncology Head and Neck Cancer (HIPO-HNC) Cohort.....	48
4.2.1. SOX2 and SOX9 Protein Expression in Tumor Samples.....	48
4.2.2. SOX9 Expression in the Tumor Microenvironment.....	51
4.2.3. SOX2 and SOX9 Gene and Protein Expression are Associated in Patient Samples.....	52
4.3. Impact of SOX2 Low Expression in Cellular Models .....	53
4.3.1. Characterization of SOX2 and SOX9 Expression Pattern in HNSCC Cell Lines .....	53
4.3.2. SOX2 Expression and Response to Irradiation .....	55
4.3.3. Impact of Irradiation on SOX2 and SOX9 Expression .....	61
4.4. Impact of Altered SOX9 Expression in Cellular Models .....	63
4.5. Downregulation of SOX2 and Immune Response .....	67
4.6. Single Cell Profiling with Tumor Samples.....	71
<b>5. Discussion</b> .....	<b>75</b>
5.1. Computational Analysis of SOX2 and SOX9 as Potential Biomarkers in HNSCC.....	75
5.2. The Role of SOX2 and SOX9 Expression in Tumor Tissue and in vitro .....	77
5.3. Radiosensitivity in HNSCC .....	80
5.4. The Role of SOX9 in Cancer .....	82
5.5. SOX2 and SOX9 Expression and Immune Modulatory Relation .....	83
<b>6. Conclusion and Perspectives</b> .....	<b>85</b>
<b>7. Summary</b> .....	<b>87</b>
<b>8. Reference</b> .....	<b>91</b>
<b>Supplements</b> .....	<b>101</b>
<b>Curriculum Vitae</b> .....	<b>123</b>
<b>Acknowledgement</b> .....	<b>124</b>
<b>Eidesstattliche Versicherung</b> .....	<b>125</b>

## List of Figures

<b>1. Introduction.....</b>	<b>9</b>
Figure 1. Head and Neck Squamous Cell Carcinoma Anatomical sites.....	9
Figure 2. Illustrative Representation of the Transcription Complex Driven by Transcriptional Factor Activity.....	12
Figure 3. Schematic Illustration of the SOX Family including groups, regions and domains.....	13
Figure 4. Role of SOX Genes in Cellular Lineage Development.....	14
Figure 5. SOX2 Main Role in Embryonic Development, Pluripotency and Homeostasis in Healthy Tissues.....	15
Figure 6. Cell Cycle Phases and Radiosensitive Properties.....	19
Figure 7. Radiation Treatment in HNSCC Based on Stage.....	20
<b>2. Aims of the Study.....</b>	<b>22</b>
Figure 8. Schematic Representation of the aims of this study.....	22
<b>4. Results.....</b>	<b>32</b>
Figure 9. Characteristics of SOX2 and SOX9 Expression in HNSCC Patient Cohorts.....	33
Figure 10. Inverse SOX2 and SOX9 Expression Pattern.....	34
Figure 11. Cut-off Point Strategy and Survival Analysis.....	36
Figure 12. Inverse SOX2 and SOX9 Group Definition Based on the Gene Expression.....	38
Figure 13. Survival Analysis of Patients in the TCGA-HNSCC Cohort with HPV Negative Status and Radiotherapy.....	40
Figure 14. Co-expression Analysis of Genes Related with SOX2 and SOX9 in the TCGA-HNSCC Cohort.....	41
Figure 15. Differential Gene Expression Analysis for the TCGA-HNSCC Cohort.....	42
Figure 16. Signature Gene Set in Clinical and Biological Relevance.....	43
Figure 17. Signature Gene Set and SOX Group.....	44
Figure 18. Signature Gene Set Biological Functions.....	45
Figure 19. Characterization of SOX2 and SOX9 Protein Expression in Tumor Sections of the HIPO-HNC Cohort.....	47
Figure 20. SOX2 and SOX9 Expression in Tumor Samples.....	48
Figure 21. SOX9 Expression in The Tumor Microenvironment of samples from HIPO-HNC.....	50
Figure 22. SOX2 and SOX9 Expression Pattern in HNSCC Cell Lines.....	52
Figure 23. SOX2 and SOX9 Expression Pattern in OSCC Cell Lines.....	53
Figure 24. SOX2 Silencing in a HNSCC Cell Line.....	55
Figure 25. SOX2 Silencing in a HNSCC Cell Lines and Response to irradiation.....	56
Figure 26. Irradiation Effect on HNO223 Cell Lines.....	57
Figure 27. Tumor Initiating Capacity of HNO223 Cell Lines.....	58
Figure 28. SOX2 and SOX9 Expression after fractionated irradiation.....	60
Figure 29: INHBA and SERPINE1 Expression after Fractionated Irradiation.....	60
Figure 30. SOX9 Silencing in a HNSCC Cell Line.....	62

Figure 31. SOX2 Expression in SOX9 Silenced HNSCC Cell Lines. ....	63
Figure 32. Slug Expression in SOX9 Silenced HNSCC Cell Lines.....	64
Figure 33. Antigen Presentation Machinery. ....	66
Figure 34: HLA-E Expression in HNO223 Cell Lines.....	67
Figure 35. NK-related Cytotoxicity and HLA-E Expression in HNSCC Cell Lines under Irradiation. 68	
Figure 36: Tumor Dissociation Single Cell Analysis. ....	69
Figure 37: Cytokine Profile in Tumor Samples. ....	71
<b>5. Discussion</b> .....	<b>73</b>
Figure 38. Schematic Diagram of the Role of Inverse SOX2 and SOX9 Expression Predicted by in Silico Analysis in HNSCC.....	83
<b>Supplements</b> .....	<b>100</b>
Supplementary Figure 1. Morphological characteristic of HNSCC cells.....	121

## ***List of Tables***

Table 1: Histopathological and Clinical Data of the TCGA-HNSCC Cohort.....	41
Table 2. Molecular and Cellular Functions Affected by the Signature Gene Set.....	47
Table 3. Association between Gene and Protein Expression in the HIPO-HNC Cohort Samples. ....	53
Table 4. List of Genes related to the Antigen Presentation Pathway after SOX2 Silencing as Assessed by Ingenuity Pathway Analysis.. .....	67
Table 5: Tumor Samples Collection. ....	72

## ***List of Supplementary Tables***

Supplementary Table 1. Gene set Signature Based in Inverse SOX2/SOX9 Expression.....	102
Supplementary Table 2. Differential Expressed Genes by Inverse SOX2/SOX9 Expression among G1 versus G2 of the TCGA-HNSCC Cohort .....	104
Supplementary Table 3. Co-expressed Genes by Inverse SOX2/SOX9 Expression in TCGA-HNSCC Cohort.....	116
Supplementary Table 4. IPA Gene Set Signature – Top 50 Canonical Pathways .....	119
Supplementary Table 5. Summary of Clinical and Histopathological Features of the HIPO Cohort ..	121



## Abbreviation List

AJCC	American Joint Commission on Cancer
ANOVA	Analysis of Variance
APC	Antigen-Presenting Cells
APC(label)	Adenomatous Polyposis Coli
ATCC	American Type Culture Collection
BCA	Bicinchoninic Acid Assay
CAFs	Cancer-associated Fibroblast
CD45	Leucocyte Common Antigen
CD56	Neural Cell Adhesion Molecule (N-CAM)
CREs	<i>cis</i> -regulatory elements
CRT	Chemoradiation Therapy
CSC	Cancer Stem Cells
CT	Chemotherapy
DEG	Differential Gene Expression Analysis
DNA	Deoxyribonucleic Acid
DSS	Disease Specific Survival
EBV	Epstein-Barr Virus
ECM	Extracellular Matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-Mesenchymal Transition
FACS	Flow Cytometry Analysis
FFPE	Formalin-Fixed Paraffin-Embedded
GEO	Gene Expression Omnibus
GO	Gene Ontology
GSVA	Gene Set Variation Analysis
Gy	Gray
HIPO	Heidelberg Center for Personalized Oncology
HLA	Human Leukocyte Antigen
HLA-E	HLA class I, alpha chain E
HNSCC	Head and Neck Squamous Cell Carcinoma
HPV	Human Papillomavirus
IF	Immunofluorescence

IHC	Immunohistochemistry
IL	Interleukin
IPA	Ingenuity Pathway Analysis
IRS	Immune-reactivity Score
KM	Kaplan Meier Analysis
min	Minutes
NK	Natural Killer cells
°C	Degrees Celsius
OPSCC	Oropharyngeal Squamous Cell Carcinoma
OS	Overall Survival
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PFS	Progression-Free Survival
RT	Radiotherapy or Radiation Therapy
s	Second
SOX2	Sex determining region Y-box 2
SOX9	Sex determining region Y-box 9
TCGA	The Cancer Genome Atlas
TF	Transcription Factor
TFs	Transcription Factors
TGF- $\beta$ 1	Transforming Growth Factor $\beta$ 1
TME	Tumor Microenvironment
TNM	Tumor-Node-Metastasis
UICC	Union for International Cancer Control
VST	Variance Stabilizing Transformation
WHO	World Health Organization

## 1. Introduction

### 1.1. Head and Neck Squamous Cell Carcinoma (HNSCC)

Head and neck squamous cell carcinoma is the 6<sup>th</sup> most common solid malignant tumor worldwide with 5-years survival rate of 66% in the last decades<sup>1</sup>. It is estimated that ~562.000 people were diagnosed with HNSCC in 2020 and ~277.000 people died from the disease worldwide<sup>2</sup>. HNSCC is one of the most common malignancies that arise from the epithelium of the oral cavity, pharynx, larynx and comprehends a heterogeneous group of tumors with site-specific molecular, cellular and clinical behaviors<sup>3,4</sup>.

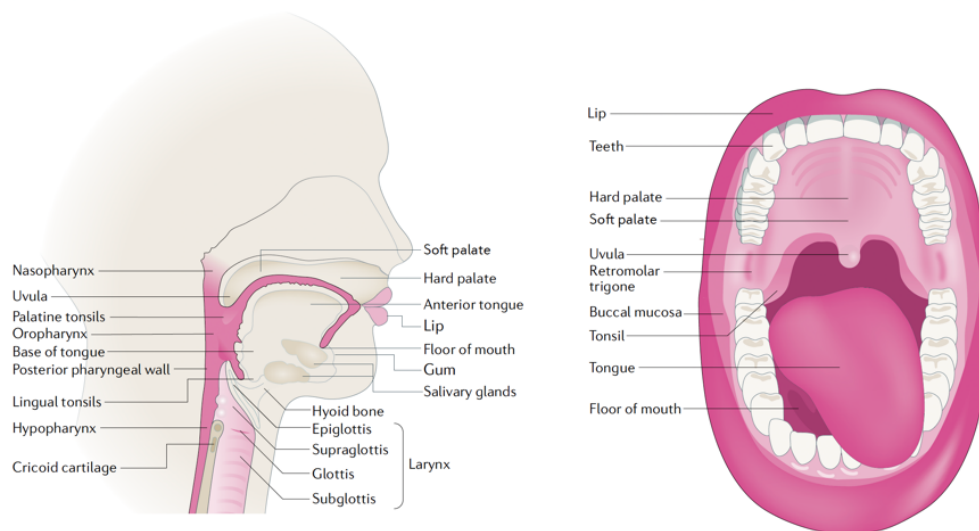


Figure 1. Head and Neck Squamous Cell Carcinoma Anatomical Sites. Extracted from Johnson *et al.*, 2020<sup>3</sup>.

The World Health Organization (WHO) classified a several of risk factors for HNSCC, such as tobacco, alcohol consumption, environmental exposure to carcinogens and viral infections (HPV and EBV)<sup>5</sup>. Oropharynx is the anatomical site with highest infection by Human Papilloma Virus (HPV), while nasopharynx is the most common anatomical site affected by Epstein-Barr Virus (EBV) infection<sup>6,7</sup>. The infection by HPV represents a distinct group of HNSCC that differ in molecular signature, disease progression and clinical response. Large number of studies reported the HPV status works as an independent risk factor and elucidated the differences between HPV-positive and HPV-negative tumors<sup>8-10</sup>. The most frequent type of HPV viral infection in HNSCC is the HPV16/18 with almost 90% of the infections of oropharynx. Changes in lifestyle are causes for the increased number of HPV-positive tumors<sup>7,10</sup>.

A large body of molecular studies with tumor samples and cell lines from HNSCC demonstrated differences in molecular signatures among HPV negative and positive tumors<sup>7,9</sup>. The difference in gene expression profile involves genes related to apoptosis, survival, cell cycle, DNA replication and repair, immune response, signal transductions, transcription factors and epigenetic mechanisms which are often present in HPV-positive oropharyngeal SCC (OPSCC). Mutations in tumor suppressor genes and oncogenes are often related to HPV-negative tumors<sup>9,11-13</sup>.

Histologically, the epithelium of the mucosal linings, upper aerodigestive tract is composed by stratified squamous epithelium with thickness respecting anatomical site and keratinization. During the development of HNSCC the tumor may progress from a non-malignant lesion (Leukoplakia) to a premalignant lesion (dysplasia). Dysplasia, presents neoplastic alteration on the surface epithelium, such as an increase in mitotic activity, nuclear enlargement and loss of cellular organization, but did not progress to an invasion lesion. Once the cells break through the lamina propria and infiltrate into the stroma this process is designated carcinoma *in situ* which normally progresses to more advanced stages of tumors<sup>7,14</sup>.

Hanahan and Weinberg described that cancer usually arises through an accumulation of genomic and epigenetic alterations in genes, causing disturbances in signaling pathways<sup>15,16</sup>. Carcinogenesis starts with an accumulation of mutations in adult stem cells producing clonal expansion which culminates in a tumor mass<sup>14</sup>. Enormous efforts have been applied to identify candidate genes which serve as biomarkers to stratify HNSCC patients. For instance, the earliest mutations in HNSCC are the loss of chromosome 3p, 9p and TP53 genes<sup>7,17</sup>. In addition, tumors present many mutations and chromosomal alterations providing a big number of candidate driver genes in HNSCC which requires a deep and careful biological interpretation<sup>7</sup>.

When a tumor cell promotes invasion and migration to a secondary site, metastasis is characterized. The tumor microenvironment (TME) plays a role in promoting the remodeling of the extracellular matrix (ECM) through productions of metalloproteinases (MMPs) from the stromal and tumor cells facilitating the cellular invasion<sup>3,18,19</sup>. The epithelial-mesenchymal transition (EMT) has a key role in the metastatic process where the cells go through a remodeling process which leads to detachment of epithelial cells from each other and invade the connecting tissues

promoting activation of several transcriptional programs to promote a mesenchymal phenotype<sup>20</sup>. The transition from an epithelial cell to a mesenchymal phenotype occurs with repression of E-cadherin and activation of markers associated with mesenchymal states, such as vimentin, N-cadherin and fibronectin leading to a reduction in cellular adhesion and enhancing a migratory behavior<sup>21,22</sup>. Also, this process confers an increase of tumor-initiating capacity, high metastatic potential and a higher potential of resistance to killing by several therapeutic approaches<sup>20</sup>. The EMT is a plastic process being able to occur at different steps of malignancy and cell populations. Puram *et al.*, demonstrated that tumor cells localized in the edges of the tumor exhibit partial EMT, where cells partially change the classical markers or properties<sup>23</sup>. These cells presented both epithelial and mesenchymal states with a mixture of corresponding marker expressions<sup>20</sup>. The EMT process also is related to the acquisition of stem cell-like properties. These stem-like cells in carcinoma are designated cancer stem cells (CSC) with the property to self-renew and give rise to well-differentiated cell types that form the tumor<sup>20,24,25</sup>. CSCs present resistance to treatment which they may persist and after the treatment they cause locoregional recurrence or metastasis leading to tumor relapse<sup>20,24,25</sup>.

The diagnostic of HNSCC is usually performed by histopathological analysis of hematoxylin-eosin staining. Immunohistochemistry analysis is applied to perform a differential diagnostic. Cases originated from oropharynx undergo to immunohistochemistry for HPV testing followed by detection of viral DNA<sup>3,26</sup>. The stage of HNSCC is performed following the UICC/AJCC Cancer Staging Manual, which classify the tumor according the anatomical subsite, tumor-node-metastasis (TNM) system and HPV status<sup>27</sup>. Furthermore, all patients have complete head and neck examination with clinical inspection and imaging to evaluate locoregional disease. The PET-CT is used to evaluate distant metastasis<sup>3</sup>.

## 1.2. SOX Family of Transcription Factors

Gene expression is a complex and well-regulated process by which information encoded in deoxyribonucleic acid (DNA) is translated to a functional product<sup>28</sup>. Transcription factors (TFs) are key proteins that regulate gene expression by binding a specific DNA sequence<sup>29-32</sup>. The regulation of TFs activity is tissue/cell specific and demands a correct balance between repression and activation in order to promote particular gene expression<sup>28</sup>. The regulation of TFs is influenced by several mechanisms, such as epigenetic regulation, molecular co-factors and gene regulatory elements. The cis-regulatory elements (CREs) is the main mechanism controlling the activation and expression of TFs<sup>33</sup>. By definition, the TFs contains: One DNA-Binding Domain (DBD) which bind specific sequences of DNA (promoters and enhancers) close to the regulated gene, one Activation Domain (AD) which contains binding sites for other proteins (co-factors) and one Signal-Sensing Domain (SSD) which recognize external signals to promote gene up- or down-regulation<sup>28</sup>. Promoters are cis-regulatory elements located upstream of the initiation transcription start site and is responsible for recruitment of the RNA polymerase II. In order to initiate the process of transcription the TFs must bind the promoter regions<sup>34,35</sup>. Enhancers are cis-acting elements that bind TFs to increase gene expression. They can be located upstream, downstream or far from the initiation transcription start site<sup>36,37</sup>.

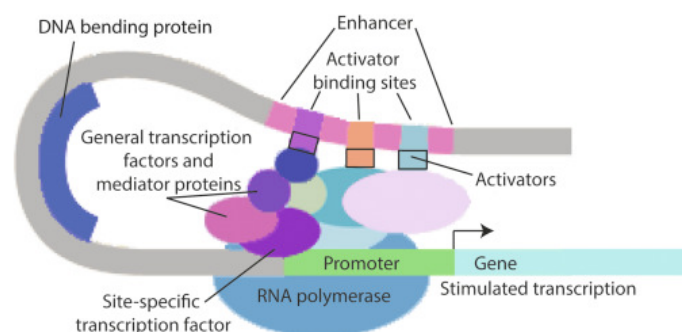


Figure 2. Illustrative Representation of the Transcription Complex Driven by Transcriptional Factor Activity. Extracted from Arie S. Mobley<sup>38</sup>.

TFs have two main features: 1 - The ability to bind short specific DNA sequences; 2 - Recruit proteins involved in the transcriptional regulation process<sup>28,30,32</sup>. The spatiotemporal regulation of TFs is linked to the cellular or tissue specificity, promoting their functions<sup>39</sup>. Dysregulation of TFs contribute to pathogenesis, such as inflammatory disorders, cardiovascular disease and a variety of malignant tumors<sup>31</sup>.

The SOX family is divided into eight groups from A to H, presenting 20 members in total<sup>40,41</sup> (Figure 3). The classification of SOX members is based on comparisons of their homology within the DNA-binding element high-mobility group (HMG) domain and other conserved regions<sup>40</sup>. The HMG group domain present two autonomous nuclear localization signals (NLSs) and one leucine-rich nuclear export signal (NES). These structures facilitate the flux of SOX proteins through nucleus and diverse subcellular localizations<sup>42-44</sup>. Members of the same SOX family also present other characteristic structural organization domains for trans-activation, trans-repression and dimerization<sup>45</sup>. The action of SOX TFs is determined by DNA-binding properties and interaction with other SOX members and/or cofactors. Their function must be adaptable during the developmental process and cellular context, making these proteins multifaced regulators of gene expression<sup>40</sup>.

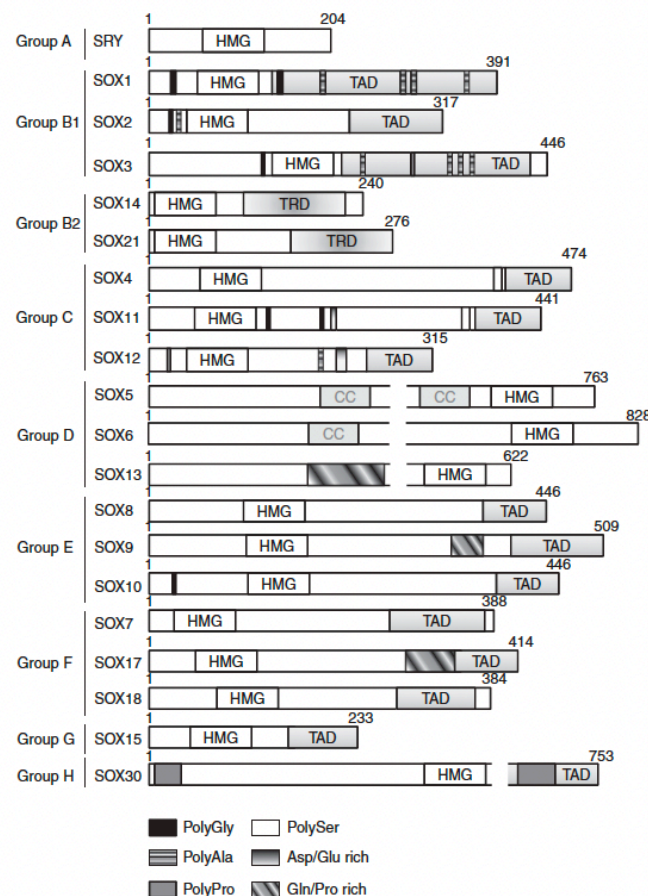


Figure 3. Schematic Illustration of the SOX Family Including Groups, Regions and Domains. Specific regions are indicated: CC: Coiled coil domain; HMG: High-mobility-group box DNA-binding domain; TAD: Trans-activation domain; TRD: Trans-repression domain. Amino acid positions at the beginning and end of each protein are shown. Extracted from Castillo *et al.*, 2012<sup>45</sup>.

The SOX genes main function by activating or repressing gene transcription during the development process, control of homeostasis in adult tissues, and cellular differentiation<sup>40,42,45</sup>. Extensive research has elucidated the involvement of SOX genes in cell-lineage fate and differentiation during the embryonic development<sup>42</sup>. Diverse SOX genes have overlapping functions with other SOX members, specially inside of the same subfamily. These close functions have argued that cellular mechanisms are associated in the same molecular network and that the regulation of SOX proteins might be depending on spatiotemporal and tissue specific traits<sup>42,45</sup>. SOX protein binding to a regulatory region of DNA is insufficient to activate transcription. A cooperation with a specific cofactor is necessary to create a functionally active complex that will be responsible to promote transcription. Thus, the combinations of SOX and co-factors provide a code for cell specification<sup>46</sup>.

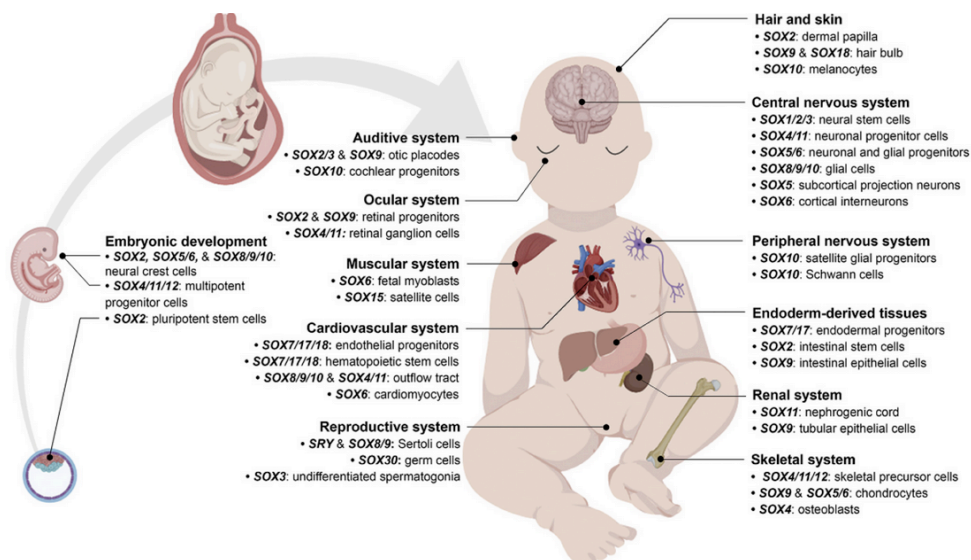


Figure 4. Role of SOX Genes in Cellular Lineage Development. Illustration with key examples of action mediated by SOX genes during the developmental process. Extracted from Dehshahri et al., 2021<sup>47</sup> and originally represented by Angelozzi et al., 2019<sup>48</sup>.

Sex determining region Y-box 2 (SOX2) belongs to the SOX protein family and is encoded the chromosome 3q26.3. Structurally it is composed of three main domains: high mobility group (HMG), dimerization (DIM) and transactivation (TAD) domains. SOX2 binds at the promoter region via its TAD to promote or repress expression of target genes<sup>49,50</sup>. It is a master regulator of embryonic stem cells being the first lineage of cell differentiation during embryogenesis and playing an important role for the development of the three germ layers: endoderm, ectoderm and mesoderm<sup>51</sup>. SOX2 also regulates the pluripotency and self-renewal of adult stem cells<sup>52</sup>. As TFs, SOX2



expression needs to be regulated more tightly at the transcriptional and post-transcriptional levels<sup>53</sup>.

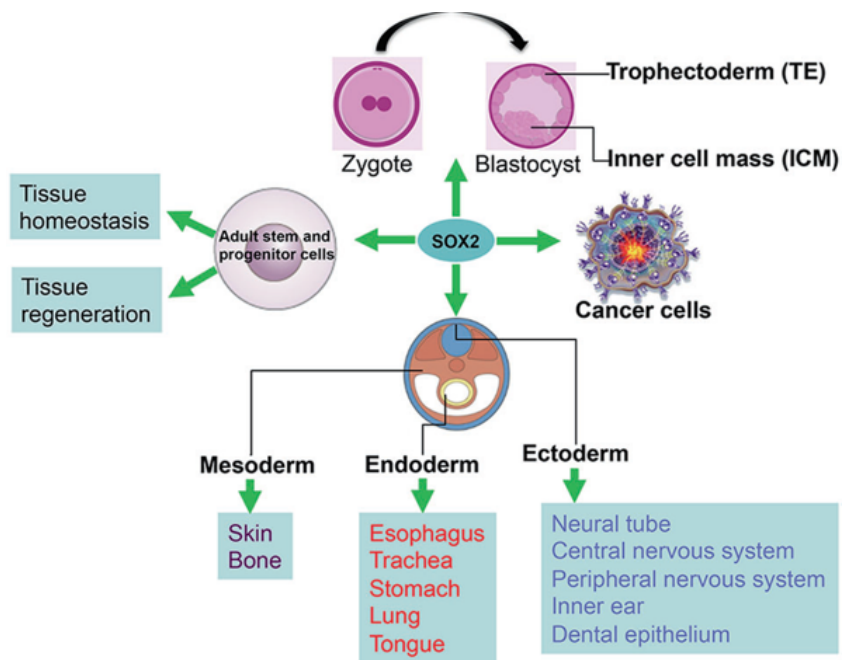


Figure 5. SOX2 Main Role in Embryonic Development, Pluripotency and Homeostasis in Healthy Tissues. Extracted from Zhang *et al*<sup>51</sup>.

Dysregulations of SOX genes in cancer have been implicated in tumorigenesis, changes in the tumor microenvironment, metastasis, and treatment resistance. Aberrant expression of SOX2 has been reported in many tumors. As an example, in gastric cancer SOX2 expression was correlated with tumor suppressor activity by modulating the WNT/ $\beta$ -catenin signaling in a mice model<sup>54</sup>. In breast cancer, SOX2 expression was linked with stemness properties which lead to hormone therapy resistance<sup>55</sup>. In lung, the regulation of stemness by SOX2 has been linked to the release of cytokines in the TME which promotes tumor cells self-renewal and/or differentiation. This plasticity of SOX2 was related to the acquisition of EMT characteristics and tolerance to therapy<sup>56</sup>. In HNSCC, low SOX2 expression was correlated with poor clinical prognosis and with a migratory behavior of the tumor cells<sup>57</sup>. Moreover, high SOX2 expression in HNSCC served as a prognostic marker together with EpCAM and vimentin in patients treated with radio-chemo-therapy<sup>58</sup>.

Sex determining region Y-box 9 (SOX9) is another gene, member of the SOXE superfamily. Its gene is located at the chromosome 17q24.3. It is structurally composed of three main domains: high mobility group (HMG), self-dimerization (DIM) and transactivation (TAD) domains<sup>59</sup>. SOX9 expression is required for the

development of the three germ layers during embryogenesis and it is precisely regulated in a spatiotemporal manner at a cellular and tissue specific context<sup>40,60</sup>. SOX9 is involved in many functions, such as: lineage restriction and temporal differentiation, maintenance of progenitor and adult stem cells, postnatal injury repair of endodermic and ectodermic organs<sup>59,60</sup>. It is subject to post-translational modification resulting in repression of expression<sup>59</sup>. SOX9 initially was studied during the embryonic development as part of chondrogenesis and male gonadal differentiation. Recently, many reports have been elucidating the involvement of SOX9 as either an oncogene or a tumor suppressor<sup>61</sup>. In breast cancer, SOX9 expression has been linked to regulation of cancer stem cell properties, EMT, metastasis and poor clinical prognostic<sup>62</sup>. In non-small cell lung carcinoma (NSCLC), SOX9 contributes to tumor development and growth<sup>63</sup>. In gastric, the SOX9 expression was associated with collagen type X alpha 1 (COL10A1) to promote migration and invasion of tumor cells<sup>64</sup>. It has been described that SOX9 present one activation and repression mode which depends of the DNA target site, the cofactor and cellular context to transcriptionally activate expression<sup>46</sup>. Moreover, the role and the molecular mechanism of SOX9 in tumor is not well established and need further research to elucidate their role in tumor development and treatment resistance.

In 2016, Lin *et al.* described that an epigenetic switch between SOX2 and SOX9 modulates the plasticity of lung cancer cells. Tumor cells with reduced SOX2 levels start to increase the expression of mesenchymal markers, which is accompanied by the loss of morphological characteristics of epithelial cells and an increase in SOX9 expression<sup>65</sup>. Another study demonstrated, how SOX2 and SOX9 TFs are regulated during therapy-induced cellular plasticity and metastasis, providing new insight into underlying principles of chemo-resistance in oral OSCC<sup>66</sup>. This study has suggested the action of stressing factors in promoting the switch in expression of SOX2 and SOX9 TFs promoting EMT and drug resistance<sup>66</sup>.

Nonetheless, Malladi *et al.*, demonstrated the imbalance in SOX2-SOX9 expression as an essential factor for stem cell identity, pluripotency, immune surveillance and metastasis<sup>67</sup>. They have shown that the existence of latency competent cancer cells in lung and breast tumors can be one mechanism suppressing outgrowth, long-term survival and maintenance of stemness property. The difference in expression of SOX2/9 and the silencing of WNT signaling enables tumor cells to

enter a quiescent state and to evade the immune clearance by NK cells, finally resulting into propagation into metastasis-establishing cells<sup>67</sup>. More recently, Laughney *et al.* showed that the selective pressure of the immune surveillance can modulate the expression of SOX2 and SOX9 to enable tumor cell adaptation and immune escape. In a mouse model, they identified a quiescent state in which tumor cells remain undetectable for a period of time. These data suggest a higher SOX2 expression of during regeneration of the tissue and during earlier stages of tumor formation. On the other hand, high expression of SOX9 is present in proliferative, regenerative stages and inhibits the killing by NK cells<sup>68</sup>.

### **1.3. Tumor Microenvironment in HNSCC**

Recently, with the advance of genomic and cellular studies, it has become evident that the nature of different signals affects cell behavior. The mixture of signals originating from the tumor microenvironment alter the epithelial cells and induce expression of genes and pathways in different combinations<sup>18,69</sup>. The TME has been shown to promote aggressive behavior of the tumor cells and therapy resistance<sup>70</sup>. The secretion of soluble factors, growth factors, nutrients, cytokines and immunomodulatory signals by tumor cells or stromal cells promote clonal selection of cells and acquisition of many hallmarks in cancer<sup>69</sup>.

A variety of pathways are activated and/or repressed by signals from the TME, such as: cell cycle, differentiation, EMT, apoptosis, stemness, immune surveillance, metastatic niches and therapeutic response<sup>69</sup>. Stromal cells can provide a support to tumor cells to proliferate, invade and resist to treatment. Cancer-associated fibroblasts (CAFs) are the most abundant cellular type found in the TME. CAFs are responsible to produce metalloproteinases (MMP), promoting the remodeling of the EMC and facilitating tumor cell migration<sup>71</sup>. Tumor associated macrophages (TAMs) play a role in controlling EMT, cellular invasion, migration and immunosuppression<sup>72</sup>. The TME in HNSCC presents overexpression of cytokines, NK cells, TAMs and T cells, even so HNSCCs are often classified as a cold tumor with an immunosuppressive phenotype<sup>7,73</sup>.

The secretion of TGF- $\beta$  by regulatory T cells (Tregs), CAFs, TAMs, and myeloid derived suppressor cells (MSDCs) in the TME of HNSCC promotes angiogenesis, mediates cell-cell communication, and regulates the immune response. In HNSCC, the

mechanisms used by tumor cells to prevent immune recognition are down-regulation of major histocompatibility complex (MHC) and inactivation of antigen presentation machinery (APM)<sup>74</sup>. The high expression of TGF- $\beta$  is known to be associated with decrease of NK cells activity<sup>8,75,76</sup>. Also, CAFs produce high levels of TGF- $\beta$ , IL-10 and IL-6 which are linked to recruitment of immune cells by promoting an inflammatory environment through secretion of chemokines<sup>77</sup>.

Cancer stem cells (CSCs) are self-renewing cells that compose one small part of the TME<sup>78</sup>. These cells play a key role in maintaining tumor heterogeneity and tumor-initiating potential, which might contribute to metastasis<sup>79</sup>. The plasticity of these cells has been correlated with a quiescent status, survival, recurrence and resistance to therapy<sup>66,68</sup>. The hypoxic TME serves as niche for CSCs<sup>80,81</sup> and it has been shown that particularly hypoxia plays a crucial role in impeding treatment response of HNSC<sup>82,83</sup>. Furthermore, the involvement of the microbiota in cancer has been proposed to present a significance on tumor growth and development, probably promoting changes in the microenvironment and immune surveillance<sup>84,85</sup>.

#### **1.4. Treatment and Challenges of Resistance**

Radiotherapy (RT) is widely recommended in oncological therapy. Approximately 50% of all cancer patients will receive one kind of radiation during the course of the treatment. The ionizing radiation has the property to deposit energy in the cells, causing breaks to the chemical bonds, forming reactive free radicals, thereby damaging biological molecules, such as DNA<sup>86,87</sup>. The most commonly applied form of radiation in clinical settings are photons. Different protocols and doses of radiation are required depending on the treated tumor entity. The main goal is to deliver the maximum dose of radiation to the tumor cells and preserve the surrounding normal tissues<sup>88,89</sup>.

The radiation acts mainly during the cell cycle. Cells during the replication process present high metabolism and increased levels of oxygen, fomenting the number of free radicals, which produce damages in the DNA. During G0 and S phase the cells present the most radioresistant properties. Benefits of providing intervals between RT is to promote cell repair in the normal tissues and allow the cells to move from a resistance phase to a sensitive phase<sup>86,87,90</sup>.

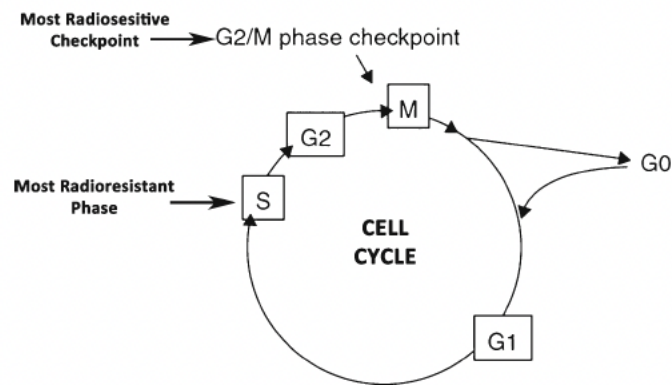


Figure 6. Cell Cycle Phases and Radiosensitive Properties. Extracted from Koushik *et al.*, 2013<sup>86</sup>.

Radiation works in many ways to promote cell death. The major effect is to induce irreversible DNA breaks. Cells with DNA damage have compromised metabolism, pausing proliferation/replication and inducing cell death<sup>89</sup>. A variety of pathways have been discussed to be related with radiation effects on cellular death. Apoptosis and cell cycle arrest have been associated with P53 pathways<sup>91,92</sup>, while senescence has been associated with the Myc pathway<sup>93</sup>.

Recent advancements in RT incorporate new technologies by using more powerful computers, better images modalities, and delivery systems with millimeters of accuracy. Radiation has great potential to promote cure, but also cause severe comorbidities limiting the quality of life for the patients. To improve the clinical outcome, radiation has been combined with other treatment modalities, focusing on quality of life for cancer patients and the therapeutic ratio<sup>86,87,94</sup>.

For HNSCC, the state-of-the-art in treatment for early stages consists of a single treatment modality that usually comprises primary surgery or definitive RT. Advanced locoregional stages make use of combined strategies of surgery followed by RT or chemo-radiotherapy (CRT)<sup>95</sup>. The platinum-based chemotherapy (CT) is the standard association with radiation<sup>95,96</sup>. Recently, immunotherapy emerged as a new line of treatment for HNSCC<sup>97</sup>. However, the tumor progression, metastasis and limited response to therapy are clearly affected by the heterogeneity of HNSCC. Many new approaches have been largely studied and applied in clinical trials to try to overcome cellular strategies to resist to treatment.

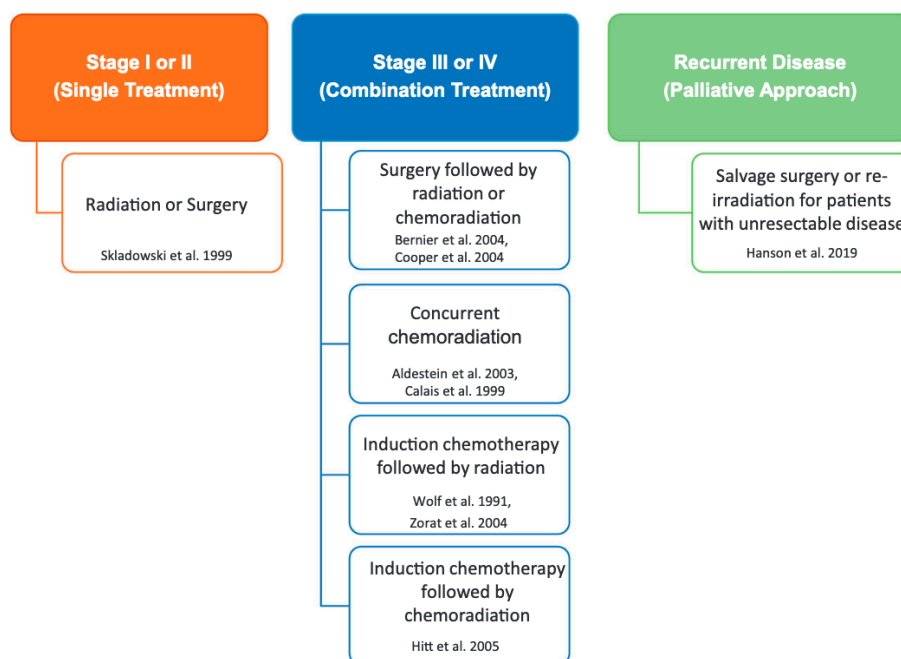


Figure 7. Radiation Treatment in HNSCC Based on Stage. Extracted from Hutchinson *et al.*, 2020<sup>90</sup>.

For the RT, the usual therapeutic dose ranges from 70 Gy based in a protocol of delivery fractionated radiation of 2 Gy once per day for several weeks. The protocol of radiation treatment depends on tumor characteristics (Figure 7). Limitations to treatment response are the overexpression of epidermal growth factor receptor (EGFR) and P53 mutation<sup>90,98</sup>. Cetuximab, a monoclonal EGFR IgG antibody, has been widely applied in combination with radiation therapy for HNSCC patients<sup>97</sup> however with limited success. Although RT eradicates the majority of tumor cells, a small amount of tumor cells are able to resist the treatment and promote a loco-regional relapse or distant metastasis<sup>99</sup>. These cells represent CSC characteristics, including resistance to conventional therapy, which can be intrinsic or acquired<sup>100</sup>. One property of the CSCs is entering in a quiescent state for an indeterminate period protecting the

cells from the radiation effects. Another property is the activation of essential pathways, which leads to repopulation of tumor cells after radiation treatment<sup>100</sup>. TGF- $\beta$  is produced by radioresistant cells and plays a role in clonal expansion by cell proliferation<sup>101</sup>. The PI3K/AKT/mTOR pathway modulates radioresistance by stimulating downstream effectors to promote antioxidative processes and a quiescent status of the CSCs<sup>102-104</sup>.

A growing body of evidence demonstrates that the association between RT and the immune system might produce synergistic effects in treatment response<sup>105</sup>. The lack of presentation, production and deregulation of immune checkpoints are known and are one obstacle to promote immune response in cold tumors. The combination of both therapies presented promising results to overcome the antitumor effects<sup>69</sup>. Dual combination with Nivolumab and Ipilimumab, an IgG antagonist to CTLA-4 is currently being investigated (NCT02551159, NCT02369874, and NCT02319044). The Food and Drug Administration (FDA) approved pembrolizumab and Nivolumab, both are anti PD-1 antibodies for patients with recurrent or metastatic HNSCC<sup>69,106,107</sup>. The combination of therapies has demonstrated improvements in patient response. However, it is necessary to understand the underlying cellular and molecular complexity of HNSCCs, the involvement of the TME, the therapeutic resistance of the tumor cells, the immunosuppressive environment and how the therapies could contribute to better response.

The role of SOX2 and SOX9 in HNSCC tumor evolution of HNSCCs and their relationship with immune cells in the microenvironment are not yet clearly understood. The modulation between the different morphological states and differential gene expression causing radio/chemo-resistance can be a possible mechanism that enables HNSCC cells to proliferate, promote metastases and escape from the immune system. Understanding, how both transcription factors, SOX2 and SOX9, and related genes can interfere with immune escape mechanisms might help us to predict which patient will respond to treatment with success and which will relapse.

## 2. Aims of the Study

Main objective of this work was to investigate the inverse expression of SOX2 and SOX9 as key nodes of a gene regulatory network in HNSCC and its potential use as biomarkers. Moreover, molecular mechanisms influencing the treatment response of HNSCC should be explored using public data repositories in order to identify a signature gene set and key pathways acting in regulation of the tumor growth, escape to immune surveillance and resistance to radiation treatment. Hence, the main aims of this study were:

1. Topological characteristics of the inverse regulation between SOX2 and SOX9 in HNSCC cohorts unrevealing an ideal classifier for patient stratification;
2. Establishment of a signature gene set based in the inverse SOX2 and SOX9 expression in association with clinical parameters;
3. Association of the signature gene set with gene-regulatory networks and biological pathways;
4. Explore the relation of inverse SOX2 and SOX9 expression in patient material and *in vitro* assays;
5. Investigate the relation between signature gene set, SOX2/9 expression and response to radiation therapy;
6. Explore the composition of tumor and immune cells in the TME and corresponding signaling in patient material.

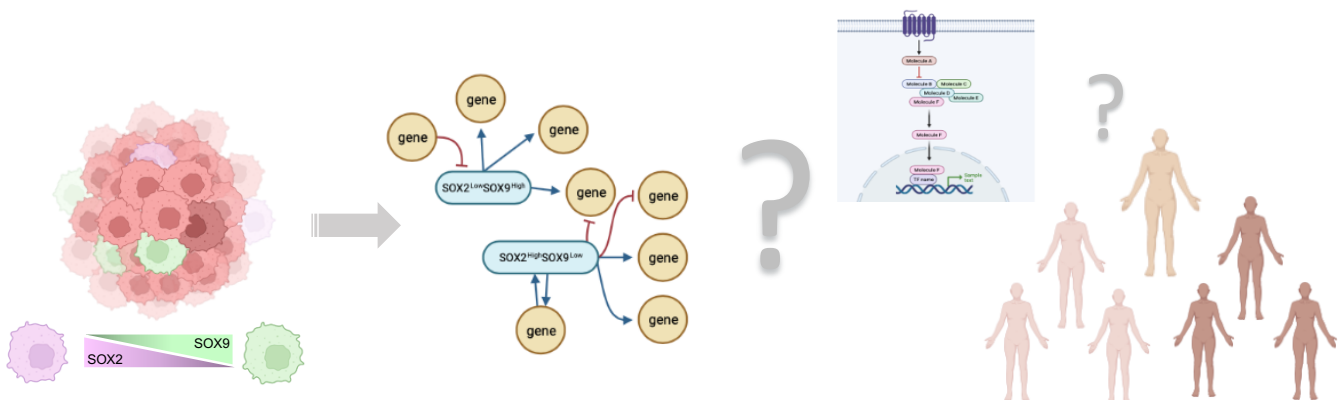


Figure 8. Schematic Representation of the aims of this study. Illustration was draw using the Biorender webtool (<https://app.biorender.com>).



### **3. Material & Methods**

#### **3.1. Computational Data Resources**

##### **3.1.1. Validation Analyses**

Since database platform, tool parameters and versions change without notice, the validation analyses were performed between September and November 2021 in order to confirm the data sources. As a proof-of-concept, all analyses were resubmitted to the pipelines and the results data were confirmed and updated to date. The analyses were performed by Amisha Dhawan (internship student) under supervision of Dr. Wahyu W. Hadiwikarta (Bioinformatician at Kurth's Lab/ Radio oncology Department/ DKFZ) and myself.

##### **3.1.2. Expression and Clinical Data sets**

The RNA-seq count data of The Cancer Genome Atlas (TCGA) were downloaded from a public repository, GDC portal (<https://portal.gdc.cancer.gov>) in January 2019<sup>108</sup>. The clinical as well as pathological data were downloaded from the cBioPortal (<http://www.cbioportal.org>) upon selecting Head and Neck Study cohorts based on the TCGA PanCancer data set (TCGA-HNSCC, n=530) in February 2019. The Heidelberg Center for Personalized Oncology (HIPO) HNC data set was used as an independent validation cohort (GSE117973)<sup>109</sup>.

##### **3.1.3. Co-expression Analysis**

Co-expressed candidate genes for either SOX2 or SOX9 were extracted according to the Spearman correlation coefficients for the TCGA-HNSCC cohort and were downloaded from cBioPortal (<https://portal.gdc.cancer.gov>)<sup>110,111</sup>. The files were loaded in R software and were analyzed using packages survminer, survival and ggplot2.

##### **3.1.4. Differential Gene Expression Analysis**

The analysis of differentially expressed genes (DEGs) was performed on TCGA-HNSCC using R studio version 3.6.1 (2019-07-05) with the DESeq2 package from Bioconductor<sup>112</sup>. For the comparison in the DESeq2 method, the patient data was

classified based on the rank of the inverse expression of SOX2 and SOX9 genes after normalization.

### **3.1.5. Survival Analysis**

Overall Survival (OS), Disease Specific Survival (DSS) or Progression-Free Interval (PFI) were calculated using the Kaplan-Meier method. Log-rank test was used to evaluate the differences between groups. P-value was considered statistically significant when less than 0.05. Survival analysis was performed in R software, using packages survminer, survival and ggplot2.

### **3.1.6. Integrative Analysis**

The integration of the gene sets generated by the previous analyses was intersected by Venn diagram using the webtool Draw Venn Diagram available at <http://bioinformatics.psb.ugent.be/webtools/Venn/>.

### **3.1.7. Clustering Analysis**

The expression values of each gene were normalized according to the variance stabilizing transformation (VST) and plotted using clustering calculation according to Euclidean distance and Manhattan (complete) linkage. Heatmaps with hierarchical trees were performed by a webtool ClustVis<sup>113</sup>.

### **3.1.8. QuPath Analysis**

Image acquisition were made using Axio slide scanner (Zeiss) where the whole slides were scanned at 20x magnification. The images were imported into QuPath software and pipeline of analyzes according software documentation<sup>114</sup>.

## **3.2. Experimental Resources**

### **3.2.1. Human Head and Neck Squamous Cell Carcinoma (HNSCC) Cell Lines**

HNSCC cell lines Cal27, FaDu, SCC4, SCC25, were purchased from ATCC (<https://www.atcc.org>). HNO223 was purchased from CSL (Cell Line Service, GmbH, Germany). HNO97, HNO199, HNO210 and HNO222 were accessed from Herold-Mendes Lab at the Heidelberg University Hospital. HNO223-Luci and HNO223-

shSOX2 cell lines were described previously<sup>57</sup> and were kindly provided by Dr. Adriana Jou. HNO223-shSOX9, non-target and positive control cells were generated at the DKFZ Proteomics Core Facility using Dharmacon™ Inducible shRNA according to the manufacturer's instruction.

### **3.2.2. Cellular Culture**

The HNSCC cell lines were cultivated with Dulbecco's Modified Eagle's Medium (DMEM; Sigma, Germany) supplemented with 10% fetal bovine serum (Invitrogen, Germany), penicillin-streptomycin 50µg/ml (Invitrogen, Germany) and 2M glutamine (Invitrogen, Germany). The cells were kept in humidified and sterile conditions with 5% CO<sub>2</sub> at 37°C. When the cells reached 80-90% confluency, they were detached by adding Trypsin-EDTA solution (Sigma, Germany) for 5 minutes at 37°C. Trypsin reaction was neutralized by adding a 2x complete medium. The viability of cells was checked with trypan blue.

The HNO223-Luci and HNO223-shSOX2 cell lines were kept under selection adding Zeocin (60 µg/ml) in DMEM complete medium (Invitrogen, Germany). The HNO223-shSOX9 were treated with Doxycycline (1µg/ml) in DMEM complete medium (Invitrogen, Germany).

### **3.2.3. Colony Formation Assay**

For the two-dimensional (2D) colony formation assay, cells were seeded in a 6 well plate with cell concentration ranging from  $5 \times 10^2$  to  $1.5 \times 10^3$  cells/well depending on the cell line. Protocol of a single dose of irradiation was applied after one overnight cultivation upon seeding. The plates were submitted to a single dose of 2Gy, 4Gy, 6Gy or 8Gy using MultiRad Faxitron (200kV and 17.8mA). The control plates were omitted from irradiation (0Gy). After irradiation all cells were cultured for 10 days to recover. For the multiple doses protocol, after one overnight cultivation upon seeding, the plates were submitted to a daily dose of 2Gy during 5 consecutive days resulting in a total dose of 10Gy of radiation. The cells were cultured for 5 days to recover. At the end, the colonies were fixed, stained with crystal violet and the total number of colonies was quantified as described<sup>115</sup>.

For the three-dimensional (3D) colony formation assay, 150µl of a media-Matrigel mixture (1:1) containing  $1.5 \times 10^3$  cells was added to each well of the 96-well

plates (ultra-low attachment, Corning). The plates were immediately irradiated with either single dose or multiple doses according to irradiation protocols. After one overnight cultivation, 50µl of media was carefully added into each well. The plates were kept at 37°C and 5% CO<sub>2</sub> in a humidified incubator for 10 days. Plates were scanned in a Nikon Eclipse Ti with 2x objective and colonies with a diameter higher than 50µm were counted using ImageJ software.

### **3.2.4. Real-time Quantitative Polymerase Chain Reaction**

#### **3.2.4.1. RNA extraction**

RNA extraction was performed by RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. Briefly, the cell pellet was lysed with RLT lysis buffer and Ethanol 70%. The lysate was transferred to a spin column and centrifuged for 15s at 10.000 rpm. RW1 buffer was added and centrifuged again for 15s at 10.000 rpm. To avoid DNA contamination 10µl DNase I (10U/µl) was applied onto the column for 15 min at room temperature, followed by serial washing steps. Finally, 30µl of RNase-free water was applied onto the column to elute the RNA. Quality control and quantification of purified RNA was done in the Nanodrop ND-100 spectrophotometer (Thermo Scientific).

#### **3.2.4.2. cDNA synthesis**

Reverse transcription of mRNA into cDNA was performed using the SuperScript<sup>®</sup> III Reverse Transcriptase Kit (Thermo Fischer). Initially, 1µg RNA, 1µl dNTP mix (10mM) and 1µl Oligo (dT) primers were added up to 10µl RNase free water. The mixture was incubated in a thermo-cycler(Thermo Scientific) for 10min at 65°C and 10min at 25°C. Then a master-mix with 2µl of DTT (0.1M), 2µl of 5x First Strand Buffer and 1µl SuperScript III reverse transcriptase (200U/µl) were added and the samples were incubated for 50min at 55°C. Samples were left for 10min at 85°C to stop the reaction. The synthesized cDNA (~50ng/ml) was diluted up to 100µl total volume and stored at -20°C.

#### **3.2.4.3. SYBR<sup>™</sup> green PCR**

To determine relative amounts of specific transcripts a quantitative real-time PCR was conducted on StepOnePlus (Applied Biosystems). 10µl of reaction mix (7.5µl

of SYBR green, 20nM forward and 20nM reverse primer) and 5µl of cDNA (diluted 1:5) were added to a 96-well plate. Samples were initially denatured at 95°C for 15 sec followed by 40 cycles of: denaturation (15 sec at 95°C), primer annealing (10 sec at 60°C) and elongation (60 sec at 72°C). The melting curve was recorded after complete amplification by following phases of heating to 95°C and posterior cooling to 60°C. The relative gene expression of the target genes (SOX2, SOX9, INHBA and SERPINE1) was calculated by normalization to a housekeeping gene (GAPDH, ACTB and TBP) using  $2^{-\Delta CT}$  or  $2^{-\Delta\Delta CT}$  method. Primers were designed manually and ordered from Sigma Aldrich. The primer efficiency was determined previously using a standard curve with serial dilution for cDNA.

### **3.2.5. Western Blotting**

#### **3.2.5.1. Protein Extraction**

Cells were detached by adding Trypsin-EDTA Solution (Sigma, Germany) for 5 minutes at 37°C, trypsin reaction was neutralized by adding 2x complete medium and centrifuged 15 min at 13.000 rpm at 4°C. The pellets were stored at -80°C.

Collected pellets were lysed in ice-cold Radio-Immunoprecipitation Assay (RIPA) buffer (Sigma Aldrich) plus protease and phosphatase inhibitor cocktail (Sigma Aldrich). The mixture was incubated on ice for 15 min and centrifuged for 10 min at 13.000 rpm at 4°C. Finally, the collected supernatant was transferred to a new 1.5 ml tube and the protein concentration was determined by bicinchoninic acid (BCA) assay (Pierce) according to manufacturer's instructions. The colorimetric values of each sample were measured at 562 nm with a microplate reader (SpectraMax ID3, Molecular Devices).

#### **3.2.5.2. SDS-PAGE (Sodiumdodecylsulfate polyacrylamide gel electrophoresis)**

The SDS-PAGE was performed to separate the proteins according to different sizes. Initially, a separation gel was prepared (12% Acrylamide/Bis, 1.5M Tris-HCL pH8.8, 10% SDS, 10% APS and 0.1% TEMED). After the solidification, a stacking gel was added on top (4% Acrylamide/Bis, 0.5M Tris-HCL pH6.8, 10% SDS, 10% APS and 0.1% TEMED). Equal amounts of protein samples (20 µg) and Laemmli buffer were denatured for 5 min at 95°C. Finally, the samples were loaded onto the gel and

subjected to electrophoresis. A pre-stained protein ladder (Thermo Fischer) was used as control to estimate the size of the proteins.

### **3.2.5.3. Polyvinylidene Difluoride Membrane (PVDF)**

After electrophoresis the separated proteins were transferred onto a PVDF membrane (Millipore). First, the membrane was activated in Methanol for 30s, washed in water and then incubated in a transfer buffer. The blotting sandwich consisted of one sponge, one filter paper, PVDF-membrane, gel, another filter paper, another sponge and was closed in a plastic support. The transfer was performed for 1 hour at 100V with 1x transfer buffer on ice. After the transfer, the membranes were blocked with 5% milk in PBS/0.5%Tween and then incubated with the primary antibodies against SOX2 ( $\alpha$ -rabbit; 1:1000; Cell Signaling) or SOX9 ( $\alpha$ -rabbit; 1:1000; Cell Signaling) for one over-night at 4°C on a shaker. The next day the membranes were washed with PBS/0.5%Tween for 15 min on a shaker and the secondary antibody-horseradish radish peroxidase HRP (Cell Signaling) was incubated for 1 hour at room temperature. Subsequently, the membrane was incubated for 1min with the enhanced chemiluminescence (ECL) solution (Pierce™, Thermo Fischer). The signal was measured by the ImageQuant LAS500 system (Thermo Fischer) with the appropriate time. Detection of alpha-tubulin ( $\alpha$ -mouse; 1:1000; Cell Signaling) and GAPDH ( $\alpha$ -rabbit; 1:1000; Cell Signaling), were used as loading control for quantity and quality of protein lysates.

### **3.2.6. Immunofluorescence (IF)**

Cells were seeded on coverslips in 12-well plates (Greiner Bio-One) and kept under normal growth conditions for one week. HNSCC cells were fixed with 4% paraformaldehyde (PFA) for 15 minutes and washed with PBS. Permeabilization of cells was done in X-buffer (0.5% Triton X-100 in PBS) for 30min at RT. After washing three times with PBS cells were blocked with the T-buffer (1% BSA / 0.2% Tween20 in PBS) for 30min at room temperature. First antibody was diluted in T-buffer and incubated for 1h. Antibodies for SOX2 (anti-mouse; 1:100; Cell Signaling) or SOX9 (anti-rabbit; 1:100; Cell Signaling), were used. Cells were washed three times in PBS. Second antibody was diluted in the T-buffer and incubated for 1h in the dark. Cy3-(goat anti-rabbit; 1:100; Invitrogen) and Alexa 488-conjugated antibodies (goat anti-mouse; 1:100; Invitrogen) were used. Cells were washed three times in PBS and

incubated for 5min in DAPI (100ng/ml, Carl Roth GmbH). After washing three times in PBS cells were embedded with Mowiol (Carl Roth GmbH) on glass slides. Images were captured with a Nikon Eclipse Ti microscope at 20x magnification.

### **3.2.7. Immunohistochemistry (IHC)**

IHC staining was performed as described previously<sup>57</sup> on selected FFPE tumor sections of the HIPO-HNC cohort (GSE117973) with rabbit monoclonal anti-SOX2 (Cell Signaling, D6D9) or anti-SOX9 (Cell Signaling, D8G8H or ABCAM, 76997) antibodies. The whole stained slices were scanned using a Slide Scanner Axio (Zeiss, Germany) with a 20x magnification objective. The immune-reactivity score (IRS) was computed as a product of the staining intensity (1: mild, 2: moderate and 3: strong) and the percentage of positively stained tumor cells ranging from 0-100%. The final immune-reactivity score (IRS) was calculated as described<sup>57</sup>.

### **3.2.8. Dissociation of Tumor Samples**

The non-malignant adjacent and tumor tissues were obtained from patients with SCC undergoing resection surgery at the Department of Oral and Maxillofacial Surgery of the Heidelberg University Hospital after obtaining informed consent. Samples were collected from four HNSCC patients with stage IV disease. Patient resection site, smoking and alcoholic history, primary lesion size, disease stage, histopathological diagnostic, treatment history was annotated in supplementary material. Tissue samples were collected from the tumor core and non-malignant adjacent tissue distant from the primary tumor. Tissue samples were immediately placed in Tissue storage solution (Miltenyi Biotec) on ice and dissociated using enzymatic digestion (Human Tumor Dissociation Kit, Miltenyi Biotec). Tissues were minced with a razor blade in the Miltenyi enzyme mix according to the manufacturer's instructions and transferred to a gentleMACS Octo Dissociator with heaters (37°C) for further mechanical dissociation. Upon completion, cell suspensions were passed through a 70-µm filter and washed twice with RPMI medium. Red blood cells were lysed in red blood cell lysis solution (Hybri-Max, Sigma) once or twice depending on red blood cell content, and final single-cell suspensions were prepared in RPMI. Final cell concentrations were determined with a hemocytometer.

### **3.2.9. Flow Cytometry Analysis (FACS)**

For analysis of cell surface protein expression on adherent cells, cells were detached using Accutase (GIBCO) to generate single-cell suspensions and were resuspended in FACS buffer (1x PBS, 0.25 mM EDTA and 2% FBS - Sigma). Cells were incubated with APC-conjugated antibody for HLA-E (Biolegend), Alexa 700-conjugated antibody for CD56 (NK cells) and/or APC/Cy-conjugated antibody for CD45 (Pan-immune cells) (Biolegend) for 30 min, and washed twice in FACS buffer. 7AAD (Biolegend) was used to exclude dead cells and immunoglobulin G control (Biolegend) staining was used as a negative control to set up the gate for analysis. To the single cell dissociation suspension, Human TruStain FcX (Biolegend) was added to the single cell suspension as a control to block nonspecific binding without interfering with the antibodies staining. Cell surface expression markers were analyzed by flow cytometry on a BD Canto II (BD Biosciences) using Diva software.

### **3.2.10. Multiplex Cytokine Assay**

Tissue samples of non-malignant and tumor areas from frozen material were prepared for a whole protein extraction using the Bio-Plex cell lysis kit (Bio-Rad) and protein amount was determined by Pierce BCA protein assay kit (Thermo Scientific) according to manufacturer's instructions. Cytokine levels were determined in duplicates using the Bio-Plex Pro Human Cytokine Screening 48-plex panel, the Bio-Plex Pro Human Cytokine ICAM-1 Set and Bio-Plex Pro Human Cytokine VCAM-1 Set (all Bio-Rad) according to manufacturer's instructions. Plates were read on a Luminex® 200 system (Thermo Fischer). The protocol was performed by Julia Pollmann (Dr. Rölle lab, NCT & Navitect Bio, Heidelberg) according to the laboratory workflow.

### **3.2.11. <sup>51</sup>Chromium Release Assay**

Cytotoxicity of NK cells against HNSCC target cells was measured in a 4 hours <sup>51</sup>Cr release assay. NK cells were purified from PBMC, extracted from the buffy coat, by negative selection (Human NK cell isolation kit, Miltenyi Biotec) according to the manufacturer's protocol. After the isolation, the NK cells were resuspended in RPMI medium and were activated in presence of Interleukin-2 (500U/ml) for 24 hours. The target cells were detached and resuspended in RPMI in a concentration of 1x10<sup>6</sup> cells/mL and labeled for 1 h with 100 µCi <sup>51</sup>Cr. Then, the target cells were washed



three times with medium and adjusted to  $1 \times 10^5$  target cells. Two effector / target (E:T) ratios were prepared in triplicate on a 96-well plate. As control, target cells were cultured alone (spontaneous release) or with 10 % Triton X-100 (maximum release). After 4 hours of incubation at 37°C, the plate was centrifuged at  $300 \times g$  for 5 min, 100µL of supernatant was collected from each well, and was counted in a gamma counter. Specific lysis was calculated as follows:

$$\% \text{ specific cytolysis} = [(cpm \text{ of EXP} - cpm \text{ of SR}) / (cpm \text{ of MR} - cpm \text{ of SR})] \times 100\%$$

### **3.2.12. Statistical Analysis**

Statistical analyses were performed using GraphPad Prism software (version 9.3.1) and R software (version 4.1.1). P-values were represented as: ns, not significant; \*p < 0.05; \*\*p < 0.01 and \*\*\*p < 0.001.

### **3.2.13. Study Approval and Patient Samples**

Buffly coats, collected according to the principles of the Declaration of Helsinki, were provided by Deutsches Rotes Kreuz (DRK) - Blutspendedienst Baden-Württemberg-Hessen GmbH (Mannheim, Germany).

Human adjacent non-malignant and human tumor tissue from human were acquired from patients undergoing surgical resection for squamous cell carcinoma (SCC) at the Department of Oral and Maxillofacial Surgery of the Heidelberg University Hospital. The patient samples were provided by the Tissue Bank of the National Center for Tumor Diseases (NCT) Heidelberg, Germany in accordance with the regulations of the tissue bank and the approval of the ethics committee of Heidelberg University (Tissue Bank Ethics Vote S-207/2005 under project number 3295). The patient samples and corresponding clinical data were provided in a pseudonymized form according to the standard procedures of the tissue bank.

HIPO-HNC cohort (GSE117973) patients were treated between 2012 and 2016 at the University Hospital Heidelberg. Sample collection and analysis of experimental and clinical data was approved by the Ethics Committee of Heidelberg University Hospital (under protocols S-206/2011, S-220/2016 and S-786/2021). The study was conducted in accordance with the Declaration of Helsinki.

## **4. Results**

### **4.1. Survival and SOX2/SOX9 Gene Expression in HNSCC Patients**

#### **4.1.1. Inverse Association of SOX2 and SOX9 in HNSCC**

To explore the hypothesis that the stem cell transcription factors SOX2 and SOX9 influence the prognosis of HNSCC patients, the association between both transcriptional factors (TFs) on mRNA level were explored by computational data analysis of the TCGA-HNSCC cohort. Spearman's correlation revealed a modest inverse association between both TFs on mRNA expression level (Figure 9A). To further explore their clinical relevance Kaplan Meier survival analysis was performed for each individual transcription factor and demonstrated an unfavorable disease specific survival (DSS) for HNSCC patients presenting low SOX2 (Figure 9B) and high SOX9 (Figure 9C) expression. To validate these observations transcriptome data of an independent HNSCC cohort (HIPO-HNC, GSE117973) were used. Spearman's correlation demonstrated a trend towards an inverse correlation between SOX2 and SOX9 transcript levels, which did not reach statistical significance (Figure 9D). However, Kaplan Meier survival analysis confirmed an unfavorable DSS for patients with low SOX2 (Figure 9E) and high SOX9 (Figure 9F) gene expression.

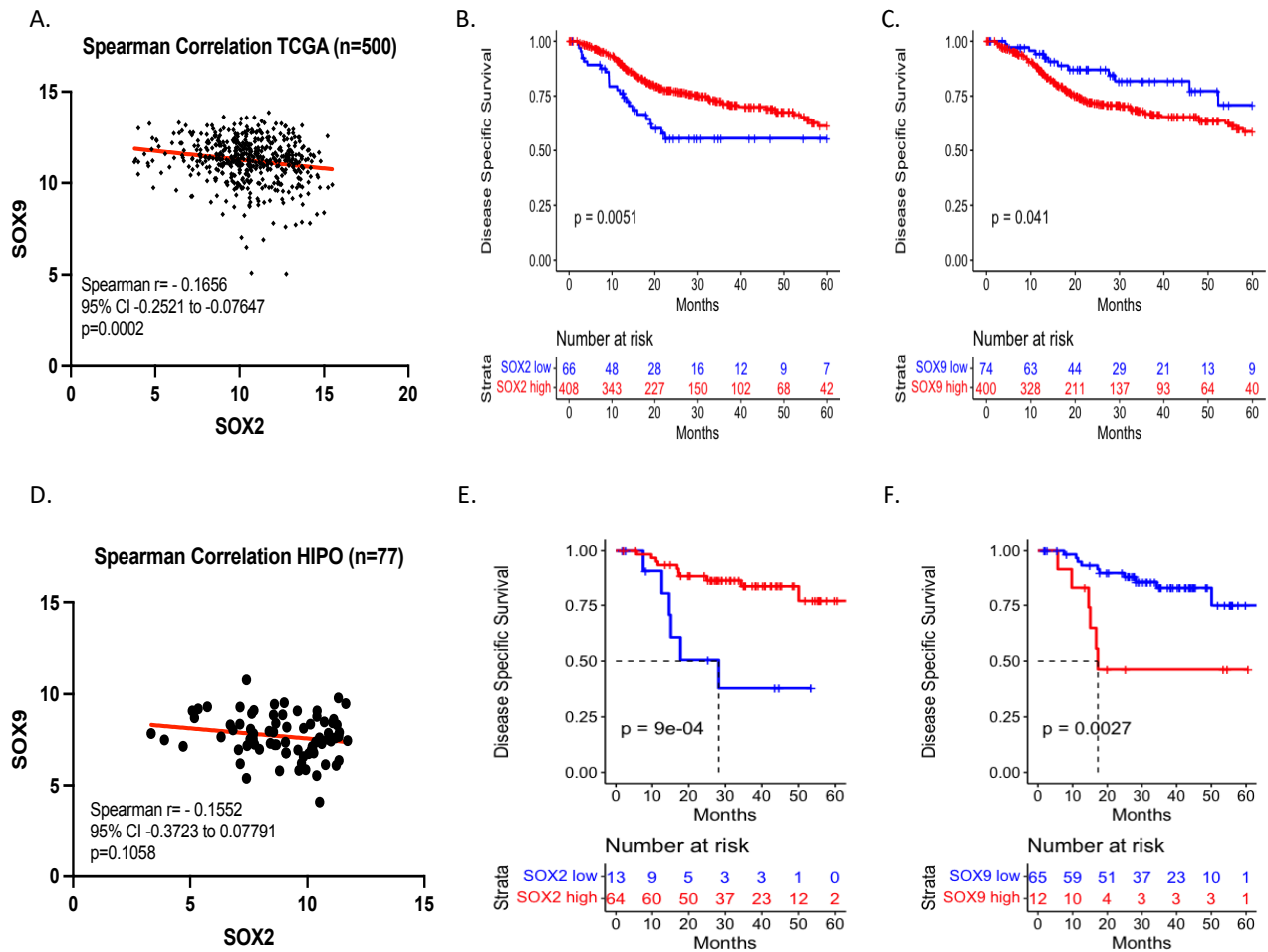


Figure 9. Characteristics of SOX2 and SOX9 Expression in HNSCC Patient Cohorts. A. Spearman's correlation coefficient was calculated between SOX2 and SOX9 gene expression values (mRNA level, RSEM). Dot plot represents patients in TCGA-HNSCC with a significantly inverse expression of SOX2 and SOX9 at mRNA level. B. Disease specific survival of TCGA-HNSCC patients (n=474) grouped according to SOX2 expression. C. Disease specific survival for TCGA-HNSCC patients grouped according to SOX9 expression (n=474). D. Spearman's correlation coefficient was calculated between SOX2 and SOX9 gene expression values of the HIPO-HNC cohort. Dot plot represents patients in HIPO-HNC with inverse expression of SOX2 and SOX9 at mRNA level. E. Disease specific survival for patients grouped according to their SOX2 expression in the HIPO cohort (n=77). F. Disease specific survival for patients grouped according to their SOX9 expression in the HIPO cohort (n=77). Missing values were removed from the analysis. Number at risk represents the number of subjects that had not yet experienced the event of interest before the time point.

Next the question was addressed, whether the inverse expression of both TFs has a prognostic value. The initial step was clustering of the gene expression values for SOX2 and SOX9 in TCGA-HNSCC and HIPO-HNC cohorts. Two main clusters of patients were observed according to the inverse gene expression of both TFs (Figure 10A and C). Second step was evaluating the variance in distribution between the gene expression values of SOX2 and SOX9. The expression data sets were loaded in R and the density plots were generated for each individual cohort (Figure 10B and D). This data analysis revealed that the TCGA-HNSCC cohort presented a higher SOX9 median transcriptional level, while the HIPO-HNC presented lower SOX9 median transcriptional levels. Thus, the results demonstrated that SOX9 expression presented a more homogeneous distribution and SOX2 was more variable.

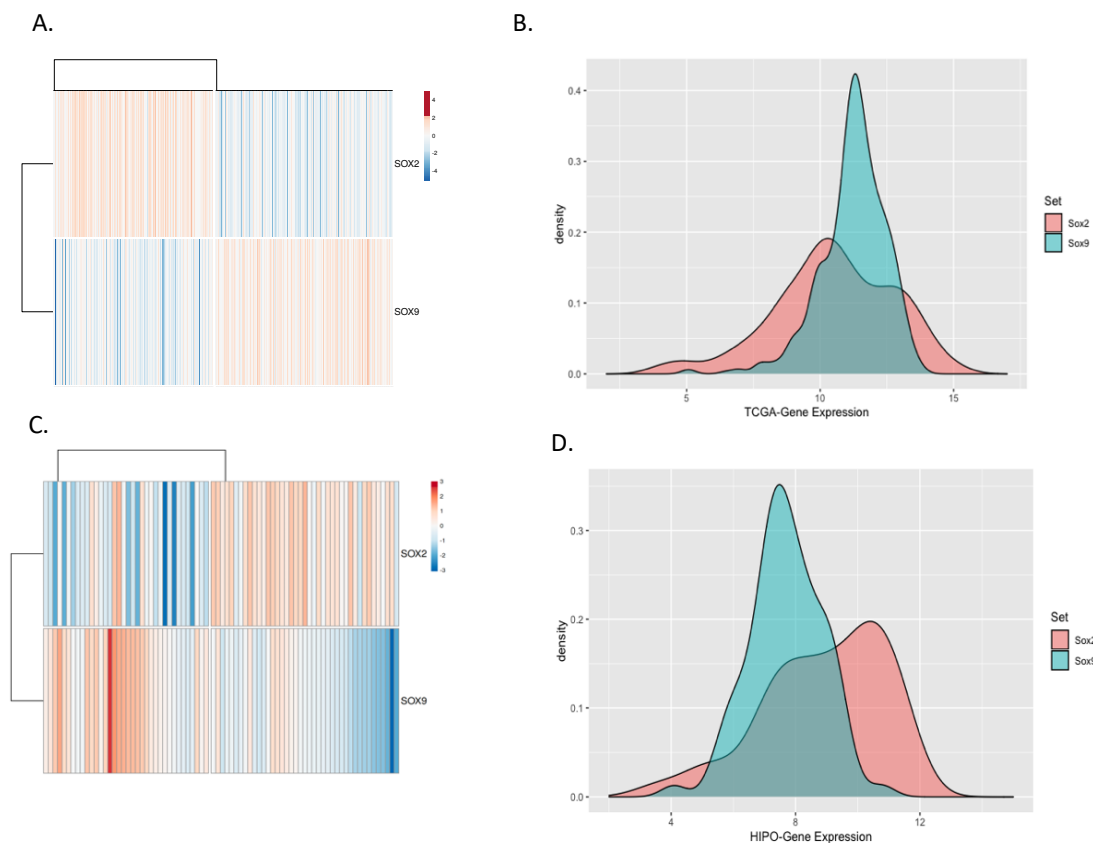


Figure 10. Inverse SOX2 and SOX9 Expression Pattern. A. Cluster analysis of SOX2 and SOX9 gene expression in TCGA-HNSCC (n=500). The values of each TF were normalized according to the variance stabilizing transformation (VST), rows were calculated using Euclidian distance and the columns using correlation distance. B. Density plot shows the distribution of SOX2 and SOX9 gene expression in the TCGA-HNSCC cohort (n=500). C. Cluster analysis of SOX2 and SOX9 gene expression in the HIPO-HNC (n=77). Rows were calculated using Euclidian distance and the columns using correlation distance. D. Density plot shows the distribution of SOX2 and SOX9 gene expression in HIPO-HNC cohort (n=77).

The final step was to find a strategy for the definition of a cut-off point to stratify patients into groups with biological significance. However, at present there are no standard methods available to define such a precise cut-off point, and this lack is even more pronounced, when two or more associated genes are involved. The most frequently used method for setting a cut-off point uses the Maxstat algorithm. With this strategy, it was possible to define individual cut-off points for SOX2 and SOX9 using the DSS as outcome parameter. As a result, the two cohorts were stratified into four patient groups according to inverse SOX2 and SOX9 expression (Figure 11A and C). Kaplan Meier analysis was applied to assess the survival probability between these groups. For TCGA-HNSCC, the group with high SOX2 and low SOX9 expression (SOX2<sup>High</sup>SOX9<sup>Low</sup>, n=68) had the most favorable prognosis. In contrast, the group with low SOX2 and high SOX9 expression presented the worst prognosis (SOX2<sup>Low</sup>SOX9<sup>High</sup>, n=117). The group with low expression values for both TFs was the smallest group (SOX2<sup>Low</sup>SOX9<sup>Low</sup>, n=6), while the largest group had high expression values for both TFs (SOX2<sup>High</sup>SOX9<sup>High</sup>, n=283). The statistical significance was only reached between the groups with inverse SOX2 and SOX9 expression (SOX2<sup>High</sup>SOX9<sup>Low</sup> vs SOX2<sup>Low</sup>SOX9<sup>High</sup>, *p* value=0.008) (Figure 11B). In HIPO-HNC, a higher number of patients had high SOX2 expression (n=55) and a small group of patients had high SOX9 expression (n=4) (Figure 11C). Again, patients with high SOX2 and low SOX9 expression showed most favorable clinical outcome. The statistical significance was reached between SOX2<sup>High</sup>SOX9<sup>Low</sup> vs SOX2<sup>Low</sup>SOX9<sup>High</sup> (*p* value=0.02), SOX2<sup>High</sup>SOX9<sup>Low</sup> vs SOX2<sup>High</sup>SOX9<sup>High</sup> (*p* value=0.001), and SOX2<sup>High</sup>SOX9<sup>Low</sup> vs SOX2<sup>Low</sup>SOX9<sup>Low</sup> (*p* value=0.0002).

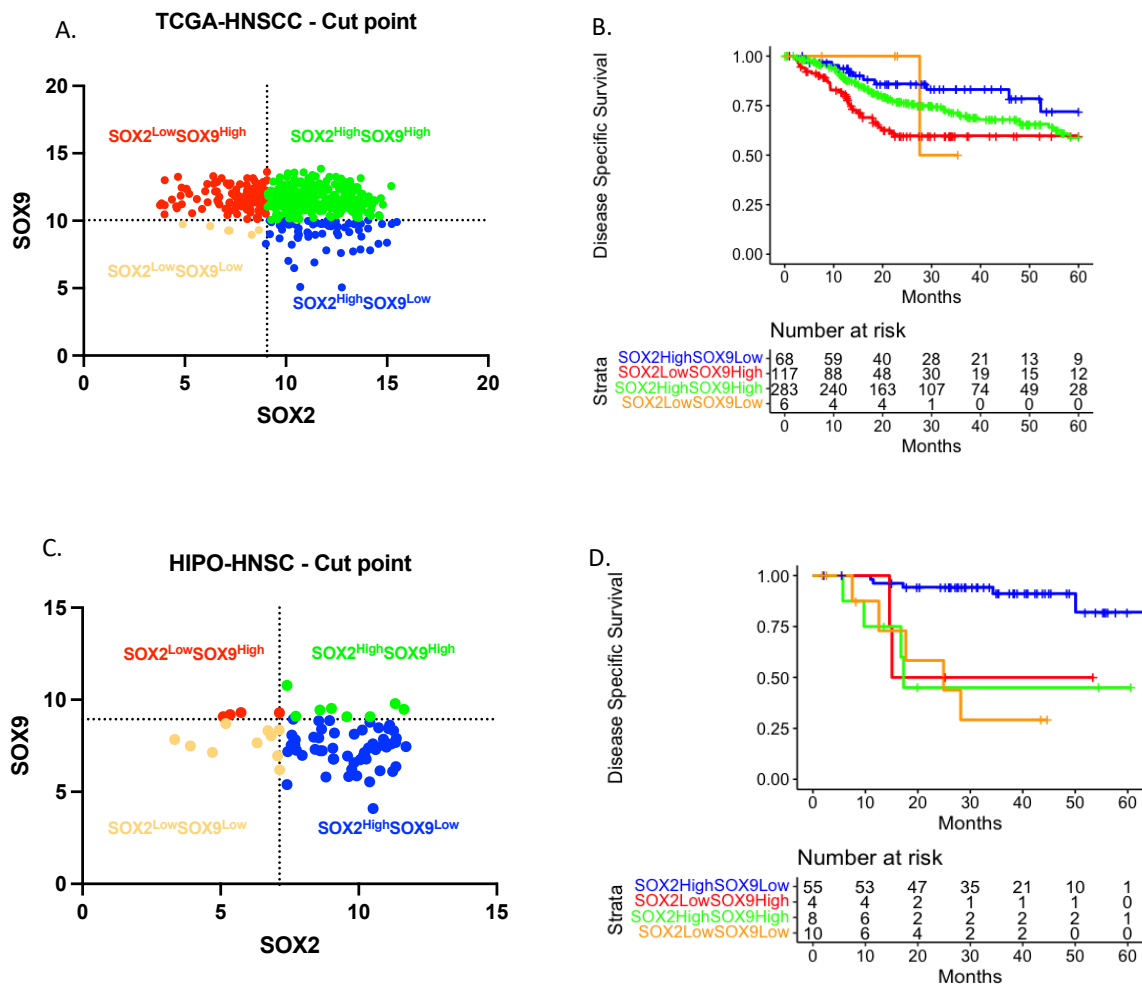


Figure 11. Cut-off Point Strategy and Survival Analysis. A. Dot plot represents the expression of SOX2 and SOX9 in the TCGA-HNSCC cohort (n=500). The values of each TF were normalized according to the variance stabilizing transformation (VST). The horizontal line represents the cut point determined by Maxstat algorithm for SOX9 (10.07) and the vertical line represents the cut point for SOX2 (9.04). The colored dots represent distinct groups according to the inverse SOX2 and SOX9 gene expression. B. Disease specific survival for TCGA-HNSCC (n=474) considering the four groups. C. Dot plot represents the expression of SOX2 and SOX9 in the HIPO-HNC cohort (n=77). The horizontal line represents the cut-off point determined by Maxstat algorithm for SOX9 (8.94) and the vertical line represented the cut point for SOX2 (7.13). The colored dots represent distinct groups according to the inverse SOX2 and SOX9 gene expression. D. Disease specific survival for the HIPO-HNC (n=77) considering the four groups. Missing values were removed from the analysis. Number at risk represents the number of subjects that had not yet experienced the event of interest before the time point. Dot plots were performed in Prism 9.0. Kaplan Meier analysis were performed in R software using the package survival and ggplot2.

In summary, the Maxstat algorithm enabled the stratification of HNSCC patients in four main groups, which differed in clinical outcome. However, this stratification into patient groups with differing clinical outcome related to the inverse expression of the TFs did not reach statistical power due to the different variance of gene expression

within the groups. Thus, with the presently available data a simple mathematical approach does not yield statistical power which would potentially allow an explanation for the relation between the inverse TF expression and the development and progression of HNSCC.

According to these limitations a second strategy was pursued. The ratio between SOX2 and SOX9 gene expression after normalization was computed for each patient. The Maxstat algorithm was applied and a new cut-off point was defined to stratify the HNSCC patients, based on the established score into two groups: G1 =  $SOX2 > SOX9$  and G2 =  $SOX2 < SOX9$  (Figure 12A and C). To evaluate the survival probability of these groups, Kaplan Meier analysis was performed and demonstrated that the inverse gene expression was not able to present a clinically relevant prognostic value in the TCGA-HNSCC cohort. In contrast, the HIPO-HNC cohort demonstrated an unfavorable clinical outcome for the patients with a low SOX2 and high SOX9 expression pattern (Figure 12B and D). Hence, this strategy presented higher stability for patient stratification, including those patients where the gene expression presented a smaller degree of variance. As a tendency, the patient group  $SOX2^{Low}SOX9^{High}$  when compared with the other groups of both investigated cohorts presented a worse survival probability. Thus, this strategy was assumed to be the best option to stratified the patients and was implemented to further analysis.

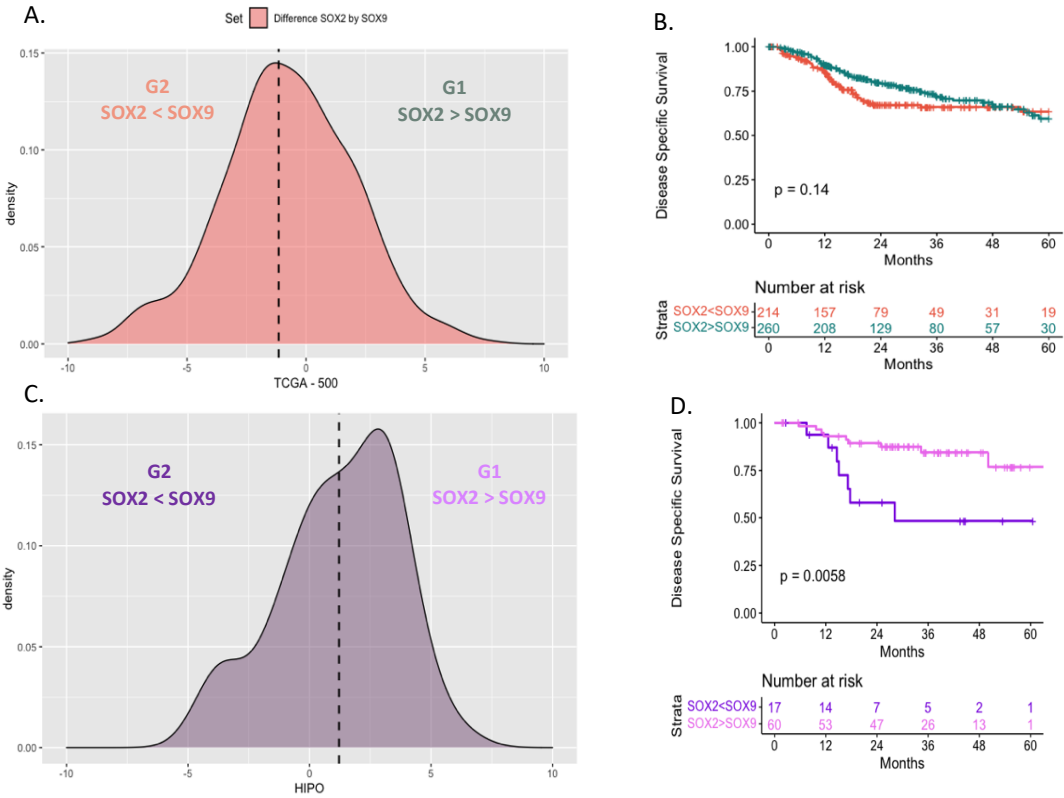


Figure 12. Inverse SOX2 and SOX9 Group Definition Based on the Gene Expression. A. Density plot of the score produced by the ratio of SOX2 and SOX9 transcript levels. Dashed line represents the cut-off point of -1.15 defined by the Maxstat algorithm based on the score. B. Disease specific survival for the patient groups of TCGA-HNSCC according to their inverse SOX2 and SOX9 expression (n=474). C. Density plot of the score produced by the ratio of SOX2 and SOX9 transcript levels. Dashed line represents the cut point of 0.94 defined by the Maxstat algorithm based on the score. D. Disease specific survival for patient groups of HIPO-HNC according to their inverse SOX2 and SOX9 expression pattern (n=77). Missing values were removed from the analysis. Number at risk represents the number of subjects that had not yet experienced the event of interest before the time point. *P value* was considered statistically significant, when lower than 0.05. Kaplan Meier analysis and density plots were performed in R software using the survival package and ggplot2.

In fact, the inverse association between both TFs was related to patient outcome and a similar pattern was observed in both cohorts investigated. However, the simple association concerning inverse gene expression did not reach a statistically significant prognostic value for the TCGA-HNSCC cohort (Figure 12B), which questions the applicability of both TFs as a biomarker for clinical outcome. Since the HNSCC is a multifactorial disease and because the inverse association between SOX2 and SOX9 might predict outcome in a context dependent manner, clinical parameters were investigated to add accuracy to survival (Table 1). Interestingly, this additional analysis revealed a significant difference in patients according to their HPV negative status and treatment by radiotherapy, respectively. Moreover, the patients with low SOX2 and high SOX9 expression (G2) presented an unfavorable DSS and were more prevalent for HNSCC originating at the oral cavity.



Table 1: Histopathological and Clinical Data of the TCGA-HNSCC Cohort.

Feature		G1 n (%)	G2 n (%)	p
<b>Patient (n)</b>		<b>260</b>	<b>214</b>	
<b>Vital Status</b>	Alive	157 (60.4)	124 (57.9)	0.63
	Dead	103 (39.6)	90 (42.1)	
<b>Overall Survival 5-year</b>	Yes	167 (64.2)	132 (61.7)	0.56
	No	93 (35.7)	82 (38.3)	
<b>Disease Specific Survival 5-years</b>	Yes	195 (75)	155 (72.4)	0.14
	No	65 (25)	59 (27.6)	
<b>HPV status<sup>1</sup></b>	Negative	194 (77)	196 (98)	<b>0.0001</b>
	Positive	58 (23)	4 (2)	
<b>Alcohol</b>	No	77 (30.3)	68 (32.4)	0.68
	Yes	177 (69.7)	142 (67.6)	
<b>Smoking</b>	No	52 (20.4)	55 (26)	0.15
	Yes	203 (79.6)	156 (74)	
<b>Gender</b>	Female	55 (21.2)	67 (31.3)	<b>0.01</b>
	Male	205 (78.8)	147 (68.7)	
<b>Age - median(range)</b>		60 (19, 84)	61 (24, 90)	> 0.99
<b>Subsite</b>	Hypopharynx	4 (1.5)	6 (2.8)	<b>0.0001</b>
	Larynx	76 (29.2)	29 (13.5)	
	Oral Cavity	122 (49.9)	169 (78.9)	
	Oropharynx	58 (19.4)	10 (4.8)	
<b>Pathological Grading</b>	G1	27 (10.4)	31 (14.5)	0.31
	G2	154 (59.3)	132 (61.7)	
	G3	67 (25.7)	47 (21.9)	
	G4	2 (0.7)	0	
	GX	19 (7.3)	14 (6.5)	
<b>Tumor size</b>	cT1	76 (29.2)	56 (26.2)	0.74
	cT2	67 (32.6)	55 (25.7)	
	cT3	91 (35)	81 (37.8)	
	cT4	6 (2.3)	5 (2.3)	
	cTX	19 (7.3)	14 (6.5)	
<b>Lymph Nodes metastasis</b>	cN0	130 (50)	93 (43.4)	0.40
	cN1-3	119 (45.7)	110 (51.4)	
	cNX	10 (3.9)	8 (3.8)	
<b>Distant metastasis</b>	M0	247 (95)	198 (92.5)	0.47
	M1	1 (0.3)	3 (1.4)	
	MX	11 (4.2)	9 (4.2)	
<b>Tumor size</b>	pT0	1 (0.4)	0	0.76
	pT1	23 (8.8)	20 (9.3)	
	pT2	73 (28)	53 (24.8)	
	pT3	49 (18.9)	40 (18.7)	
	pT4	85 (32.7)	78 (36.4)	
	pTX	20 (7.7)	12 (5.6)	
<b>Lymph Nodes metastasis</b>	pN0	90 (34.6)	71 (33.2)	0.52
	pN1	121 (46.6)	107 (50)	
	pNX	39 (15)	25 (11.7)	
<b>Distant metastasis</b>	pM0	95 (36.5)	84 (39.2)	0.42
	pM1	1 (0.4)	0	
	pMX	36 (13.8)	24 (11.2)	
<b>Pathological Staging</b>	Stage I	14 (5.3)	11 (5.1)	0.67
	Stage II	38 (14.7)	28 (13.1)	
	Stage III	43 (16.5)	30 (14.1)	
	Stage IVA	121 (46.5)	116 (54.2)	
	Stage IVB	7 (2.7)	4 (1.8)	
<b>Neoadjuvant Treatment</b>	NO	257 (98.8)	212 (99)	>0.99
	YES	3 (1.2)	2 (1)	
<b>Radiation</b>	NO	72 (32.2)	71 (36.8)	<b>0.01</b>
	YES	152 (67.8)	122 (63.2)	
<b>Additional Radiation Therapy</b>	NO	54 (58.7)	56 (70)	0.15
	YES	38 (41.3)	24 (30)	

Chi-squared test ( $\chi^2$ ) test, p values < 0.05 are indicated in bold

<sup>1</sup>HPV status – positive or negative for viral DNA and transcripts for all oncogenic variants <sup>116</sup>

Missing data were excluded for the analysis. c: clinical evaluation, p: pathological evaluation

To further explore the relation between clinical parameters and the respective SOX2/SOX9 groups, patients of the TCGA-HNSCC cohort were selected according to their HPV negative status to all oncogenic variants and radiotherapy. Survival analysis was performed considering Overall Survival (OS), Disease Specific Survival (DSS) and Progression Free Interval (PFI) (Figure 13). To investigate at what time-point during the clinical-course of the disease the inverse expression between both TFs would be clinically relevant three time points were evaluated. A significant clinical value for OS was observed only at the 18 months' time point (Figure 13A). DSS presented statistically significant difference at 18 months and 3-year follow-up (Figure 13B). PFI did not reach statistical significance at any time point evaluated (Figure 13C). In summary, the HNSCC patients group with HPV negative status, with radiotherapy and SOX2<sup>Low</sup>SOX9<sup>High</sup> expression presented a reduced survival probability when compared to the group with SOX2<sup>High</sup>SOX9<sup>Low</sup> expression, even though these difference were not statistically significant for the five years follow-up. Thus, the differences according to inverse SOX2/SOX9 were most prominent for DSS.

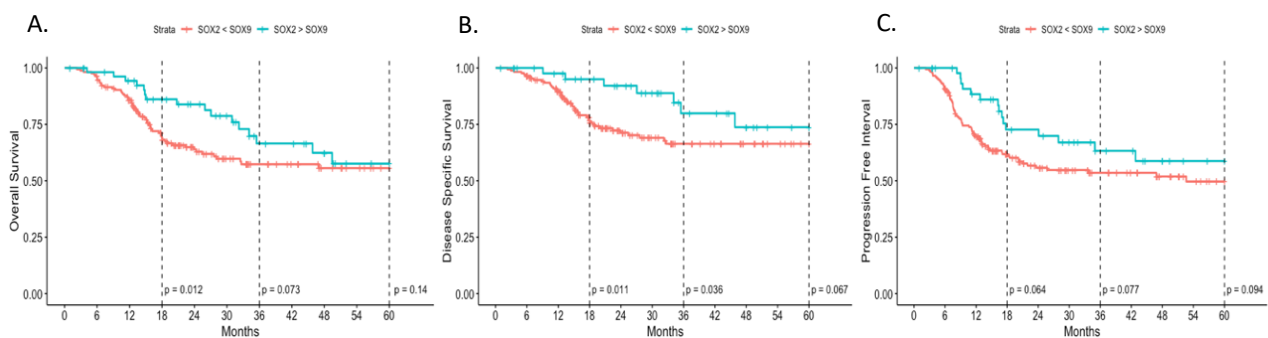


Figure 13. Survival Analysis of Patients in the TCGA-HNSCC Cohort with HPV Negative Status and Radiotherapy. A. Overall survival probability at three time points. B. Disease specific survival at three time points. C. Progression free interval at three time points. The dashed lines represent the time points selected with respective p value in the analysis. Missing values were removed from the analysis.

In summary, the definition of HNSCC patient groups based on the inverse gene expression profiling of SOX2 and SOX9 did not predict their outcome independent of clinical variables.

#### 4.1.2. Gene Network Based on the Inverse SOX2 and SOX9 Gene Expression

In order to delineate the molecular signature that could be associated with inverse expression of SOX2 and SOX9 a co-expression analysis was used to predict potential candidate genes in HPV negative status HNSCC of the TCGA cohort, which were treated with radiotherapy. For Spearman correlation analysis transcript data were downloaded from cBioPortal using the TCGA-HNSCC PanCancer data set. Files were loaded in R and density plots were generated to evaluate the correlation range for both TFs, respectively. It is worth noting that genes which correlated with SOX9 presented a smaller range of Spearman's correlations than those genes which correlated with SOX2 (Figure 14A). Based on these data a moderated correlation cut-off point of 0.2 was used to define two groups: Co1 (genes with correlations for SOX2 > 0.2 and for SOX9 < -0.2) and Co2 (genes with correlations for SOX9 > 0.2 and for SOX2 < -0.2) (Figure 14B). The resulting number of genes in Co1 was amounted to 72 genes and 129 genes were identified for Co2. The intersection of both groups revealed a final set of 201 genes which were correlated with the inverse SOX2 and SOX9 expression pattern for the TCGA-HNSCC cohort (Supplementary table 3).

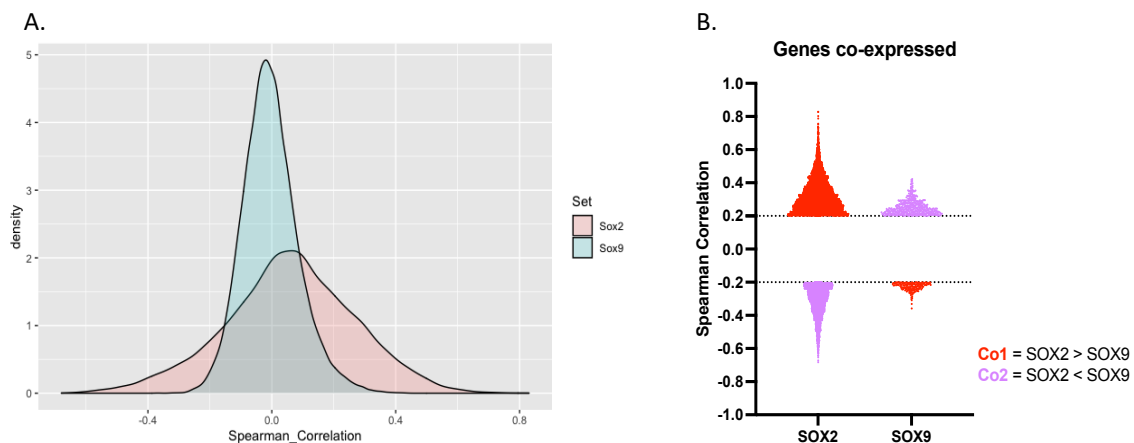


Figure 14. Co-expression Analysis of Genes Related with SOX2 and SOX9 in the TCGA-HNSCC Cohort. A. Density plot for the Spearman correlation range for SOX2 and SOX9 correlated genes in the TCGA-HNSCC cohort (n=500). B. Dot plot represents candidate genes which are correlated with SOX2 or SOX9 in patients with HPV negative status and radiation treatment in the TCGA-HNSCC (n=222). The colors indicate the group according to inverse SOX2 and SOX9 gene correlation.

### 4.1.3. Differential Gene Expression Analysis

A differential gene expression analysis was performed using the previous stratification of G1 vs G2 for the TCGA-HNSC cohort (HPV negative status HNSCC and treated with radiotherapy) to complement the co-expression analysis. The total number of differentially expressed genes (DEGs) in samples with inverse expression of SOX2 and SOX9 was 1,975 genes after applying a filter of  $-1 > \log_2FC < 1$  and adjusted p value  $< 0.01$  (Figure 15 and Supplementary table 2).

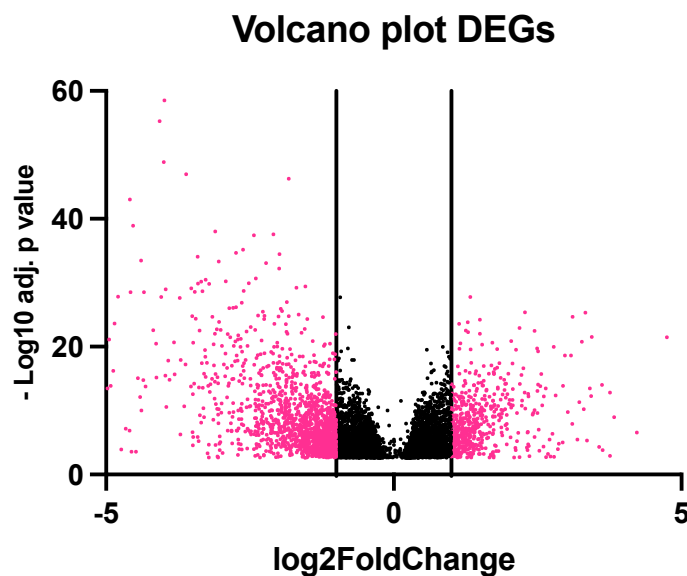


Figure 15. Differential Gene Expression Analysis for the TCGA-HNSCC Cohort. Volcano plot of the differential expressed genes in HPV negative status patients treated with radiotherapy in TCGA-HNSCC (n=222). The vertical lines indicate the filter applied to the Fold change ( $-1 > \log_2FC < 1$ ) and pink dots represents the genes in the final list of DEGs (n=1,975 genes). Analysis was performed in R software using the Desq2 package. Plot was prepared in Prims 9.0.

### 4.1.4. Integrative Gene Set Analysis and Final Signature Gene Set Based on Inverse SOX2 and SOX9 Expression

To identify a unique signature based on inverse SOX2 and SOX9 expression and respective regulation, candidate genes (n=69, Supplementary Table 1) of the intersection between co-expressed gene analysis and DEGs were identified by a Venn diagram analysis (Figure 16A). Gene Set Variation Analysis (GSVA) was used to estimate the variation of the predicted signature gene set over previously defined groups in the TCGA-HNSCC cohort. The results demonstrated that patients in the group G2 (SOX2<sup>Low</sup>SOX9<sup>High</sup>) presented a higher score for the predicted signature gene set. Survival analysis was performed to gain relevance of the signature gene set. A second score was calculated based in the average of gene expression for each

patient and Maxstat algorithm was applied to define groups with lower and higher scores of the signature gene set. The results demonstrated an unfavorable clinical outcome for patients which present a higher score of the signature gene set (Figure 16B).

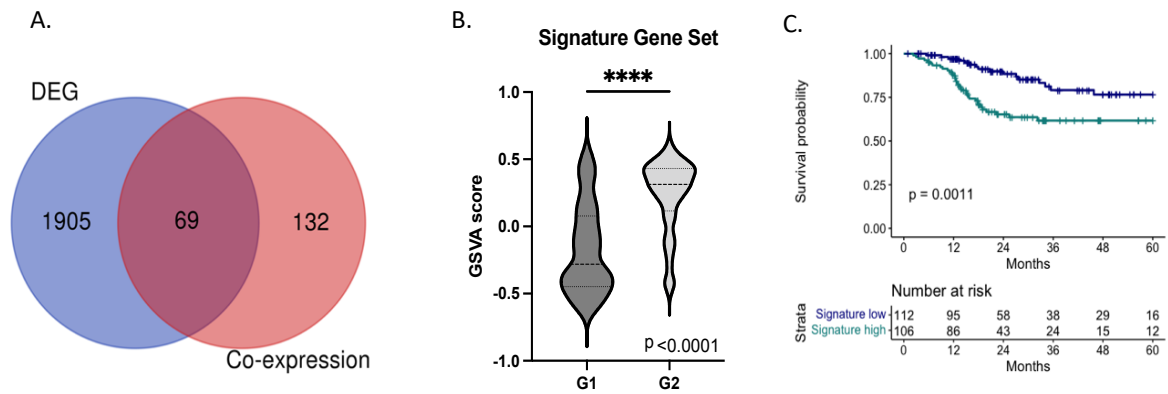


Figure 16. Signature Gene Set in Clinical and Biological Relevance. A. Venn diagram intercepting the final lists of co-expression analysis and differential gene expression analysis. The total set of common genes consisted of 69 genes. B. Violin plot for GSVA score in G1 (SOX2 > SOX9) and G2 (SOX2 < SOX9) for HPV negative status patients treated with radiotherapy in TCGA-HNSCC (n=222). Statistical analysis was performed in Prism 9.0 using t test. *p* value was considered significantly lower than 0,05. C. Kaplan Meier analysis in HPV negative status patients treated with radiotherapy in TCGA-HNSCC (n=222). Statistical analysis was performed in R software. Missing data was removed from the analysis.

In order to evaluate the relation between the previously established SOX2/SOX9 groups (G1 and G2) and the signature gene set a Principal Component Analysis (PCA) and a heatmap were generated. As a result, the PCA analysis demonstrated a separation of the TCGA-HNSCC cohort according to a high vs. low signature score. Where HNSCC with lower signature score were overlapping with the G1 group (SOX2<sup>High</sup>SOX9<sup>Low</sup>), while HNSCC with higher signature score were overlapping with G2 (SOX2<sup>Low</sup>SOX9<sup>High</sup>) (Figure 17A). A variance of 36.5% was observed for PC1 of patients which share the same group and score. The heatmap based on unsupervised hierarchical clustering revealed two main cluster of patients, and a larger number of genes with high expression were related to G2 (SOX2<sup>Low</sup>SOX9<sup>High</sup>) (Figure 17B). Thus, these analyses demonstrated that the higher expression of the signature gene set is co-related with patients who present SOX2<sup>Low</sup>SOX9<sup>High</sup> pattern, resulting in a poor clinical outcome.

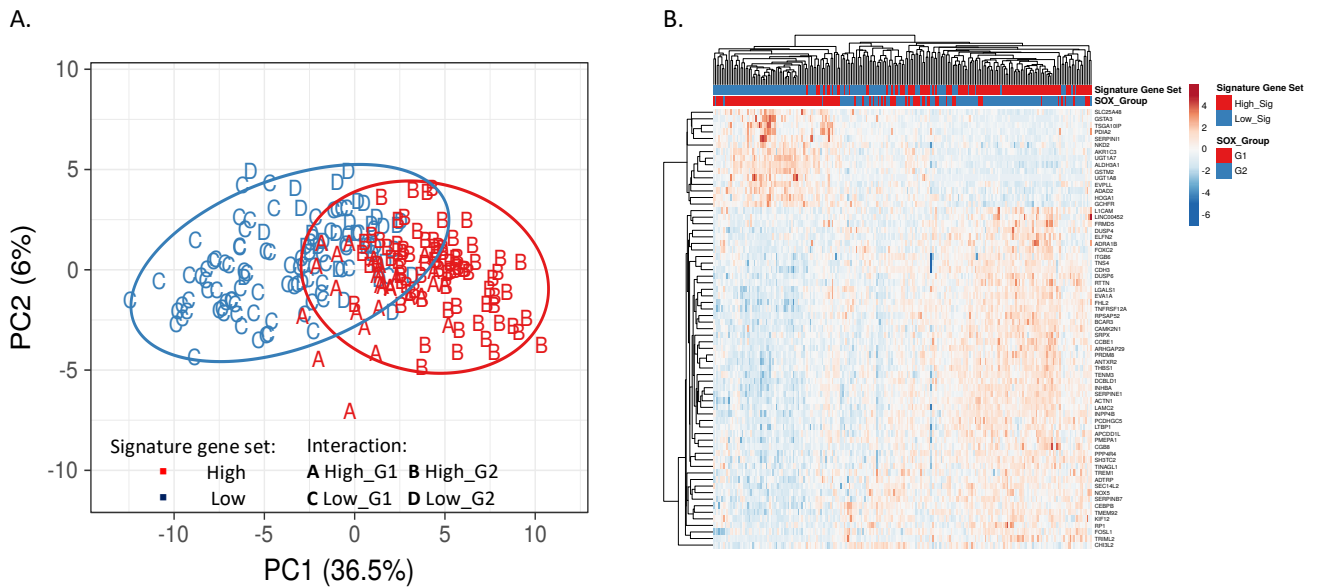


Figure 17. Signature Gene Set and SOX Group. A. The PCA-plot shows clustering of HPV negative status patients of the TCGA-HNSCC (n=222) treated with radiotherapy according to the signature gene score and the SOX2/SOX9 groups. Red color represents the high signature score and the blue color the low signature score. The letters represent the interactions between the score and the groups. B. Heatmap of the signature gene set in HPV negative status patients of the TCGA-HNSCC(n=222) treated with radiotherapy. Rows were calculated using Euclidian distance and the columns using correlation distance. Analysis was performed using the webtool ClustVis<sup>117</sup>.

#### 4.1.5. TGF- $\beta$ Signaling as one of The Top Canonical Pathways Regulated by Inverse SOX2/SOX9 Expression

To gain biological insight into the functional role of the predicted signature gene set, an Ingenuity Pathway Analysis (IPA) and STRING enrichment analysis were performed. The IPA analysis revealed a significant enrichment of differentially expressed genes involved in cellular development, cellular growth and proliferation, cellular movement, cellular function and cell-to-cell signaling (Table 2 and Supplementary Table 4).

Function Name	p-value	Number of Molecules
Cellular Development	4.70E-02 - 2.02E-06	15
Cellular Growth and Proliferation	4.70E-02 - 2.02E-06	15
Cellular Movement	3.86E-02 - 5.65E-04	12
Cellular Function and Maintenance	3.16E-02 - 9.01E-04	4
Cell-To-Cell Signaling and Interaction	4.30E-02 - 1.19E-03	9

Table 2. Molecular and Cellular Functions Affected by the Signature Gene Set.

TGF- $\beta$  signaling was the top ranked pathway in the analysis and five main genes (INHBA, SERPINE1, THBS1, ITGB6 and LTBP1) were highlighted as potentially targets for further investigation (Figure 18). Additionally, these genes were highly expressed in the group G2 (SOX2<sup>Low</sup>SOX9<sup>High</sup>) of TCGA-HNSCC.

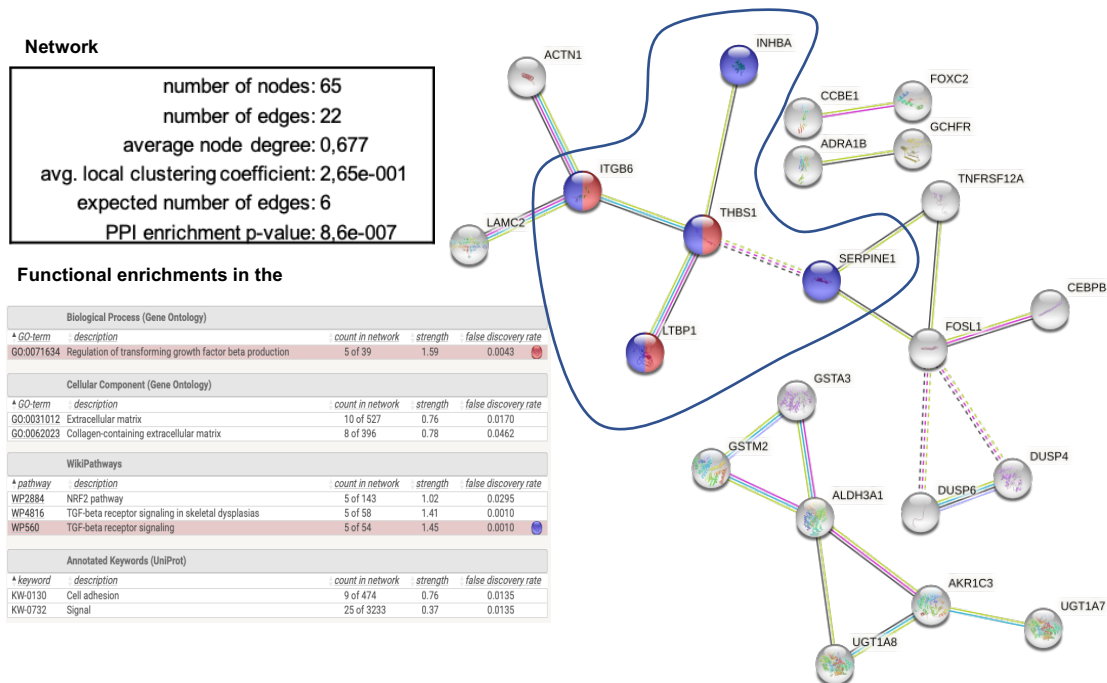


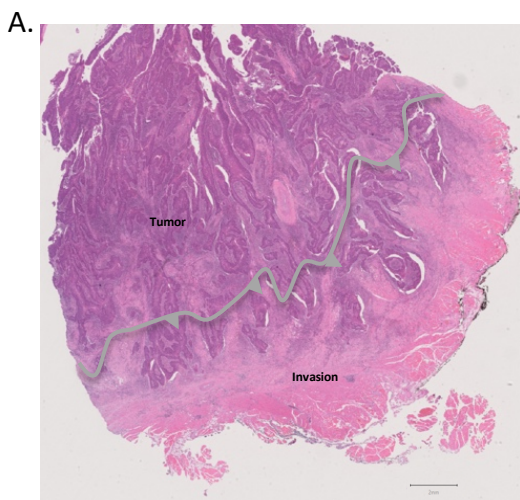
Figure 18. Signature Gene Set Biological Functions. Functional analysis of the signature gene set for Gene Ontology (GO) and biological pathways. Protein-protein interactions mapped by STRING. The blue nodes are genes connected to TGF- $\beta$  signaling and the red nodes represent genes involved in the regulation of the TGF- $\beta$  production. The analysis was performed with the STRING webtool.

Altogether, applying two different strategies to identify genes related with the inverse SOX2/SOX9 expression revealed a signature gene set of which most genes were highly expressed in the group G2 with a SOX2<sup>Low</sup>SOX9<sup>High</sup> pattern. Moreover, HPV negative status of HNSCC patients that received radiotherapy and had a higher score for this signature presented an unfavorable clinical outcome. GO analysis elucidated five candidate genes related to TGF- $\beta$  signaling, which might explain unfavorable prognosis of HNSCC with low SOX2 and high SOX9 expression pattern due to their involvement in irradiation resistance, tumor cell communication and migration, escape to immune recognition and impact on the tumor microenvironment.

## 4.2. Tumor Samples of The Heidelberg Center for Personalized Oncology Head and Neck Cancer (HIPO-HNC) Cohort

### 4.2.1. SOX2 and SOX9 Protein Expression in Tumor Samples

To further explore the inverse expression of SOX2 and SOX9 in cancer cells on protein level, immunohistochemical (IHC) staining was performed on FFPE tumor sections from the HIPO-HNC cohort<sup>118</sup>. To investigate differences in the spatial expression pattern between SOX2 and SOX9 their protein levels were assessed in two distinct areas of the tumor tissue: the invasion front and the tumor core (Figure 19A). IHC stained whole tissue sections (n=73) were scanned with the Zeiss Axion Scanner. For the semi-quantitative analysis of protein expression two parameters were defined: first, (I) five degrees of percentage for positive nuclei (1: negative staining, 2: 0-25%, 3: 26-50%, 4: 51-75% and 5: 76-100%), and second (II) three degrees of staining intensity (1: mild, 2: moderate and 3: strong) for each area. The final immune-reactivity score (IRS) was calculated as described in *Bayo and Jou et.al.*<sup>57</sup>. Furthermore, the IRS was divided in 3 levels: Low (IRS ≤ 5), Moderate (5<IRS<16) and High (IRS ≥16). A positive association was found for the IRS of both TFs at the invasion front (Spearman  $r=0.4250$ ,  $p=0.002$ ) (Figure 19B and C), while a modest inverse association was evident at the tumor core (Spearman  $r= -0.2575$ ,  $p=0.0279$ ) (Figure 19B and D). At the invasion front, a strong co-expression of both markers was observed, which is in line with the IRS score (Figure 19E). Compared to SOX9, the SOX2 protein expression was higher in the tumor core, whereas SOX9 presented higher expression at the invasive front (Figure 19F).



B.

	SOX2 n (%)	SOX9 n (%)	P value
<b>Invasion Front</b>			
Low	21 (28.76)	28 (38.35)	$p= 0.002$
Moderate	26 (35.61)	37 (50.68)	
Strong	26 (35.61)	8 (10.95)	
<b>Tumor Core</b>			
Low	12 (16.43)	21 (28.76)	$p=0.05$
Moderate	45 (61.64)	45 (61.64)	
Strong	16 (21.91)	7 (9.58)	



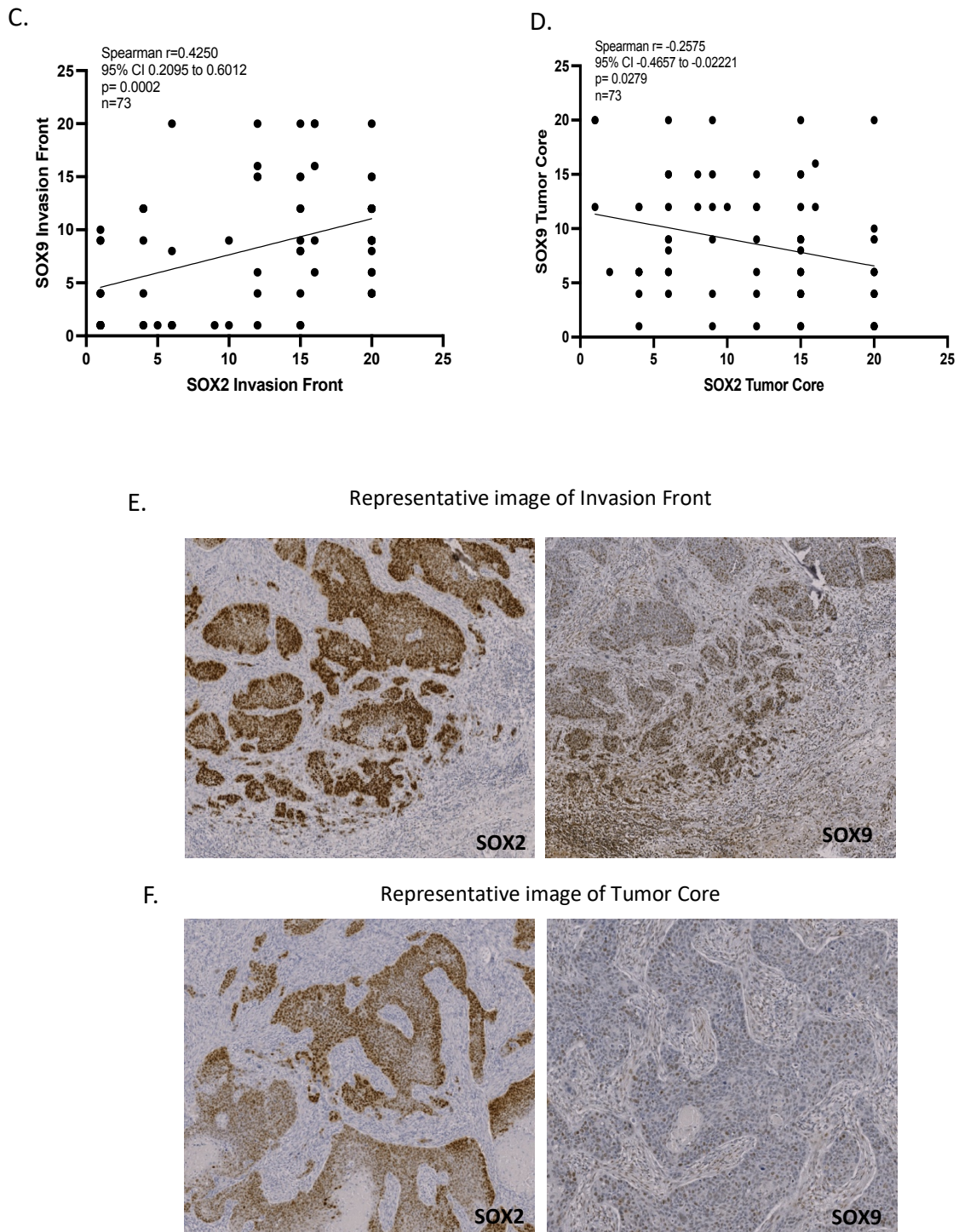


Figure 19. Characterization of SOX2 and SOX9 Protein Expression in Tumor Sections of the HIPO-HNC Cohort. A. Representative picture of a tumor sample in low magnification representing the selected areas for the analysis. B. Summary of the SOX2 and SOX9 protein expression patterns in two different areas of the tumor samples. Sample number (total  $n = 73$ ) and percentage of the total (%). Fischer's exact test. C. Spearman correlation test according to the IRS for SOX2 and SOX9 at the invasion front. D. Spearman correlation for SOX2 and SOX9 in the tumor core of the samples. Analysis was performed in GraphPad Prism 9.0,  $p$  value was considered statistical significantly lower than 0.05. E and F. Representative pictures of SOX2 and SOX9 expression at the invasion front and the tumor core.

The spatial SOX2 expression was detected homogeneously and exclusively in the nucleus of tumor cells. In contrast, SOX9 expression was heterogeneous with a positive staining in the nucleus of both tumor cells (Figure 20A) and stromal cells of the tumor microenvironment (Figure 20B). It is worth noting that cancer cells with an inverse staining ratio between SOX2 and SOX9 were located in different compartments within the tumor structure. These results demonstrated a different spatial expression profile for both TFs in a substantial amount of HNSCC samples.

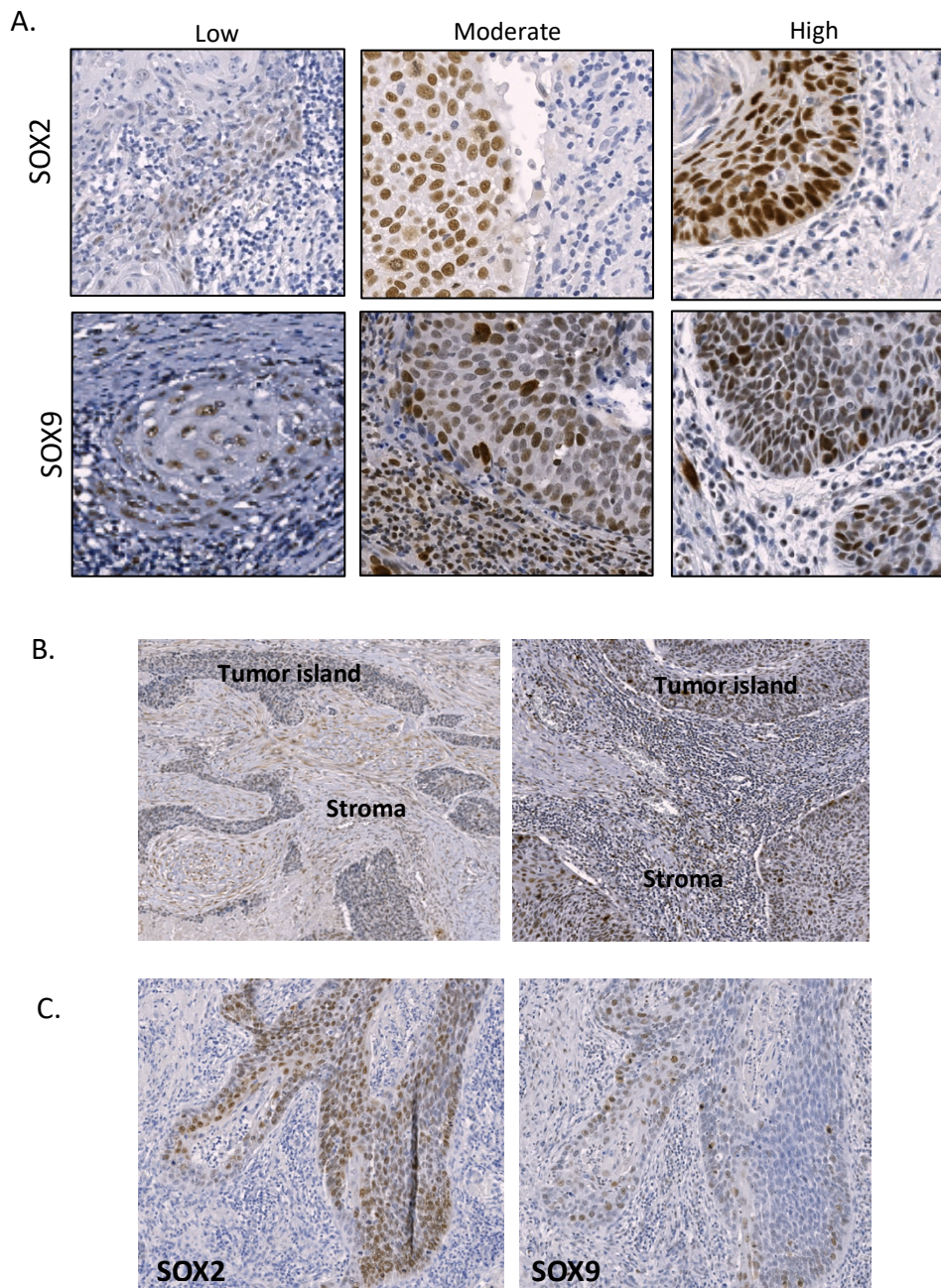


Figure 20. SOX2 and SOX9 Expression in Tumor Samples. A. IHC staining for SOX2 (upper panel) and SOX9 (lower panel) in representative cases with low, moderate and high expression. B. Selected images representing positive SOX9 expression in the stromal cells. C. Selected image representing an inverse

SOX2 and SOX9 expression pattern for tumor cells at distinct locations. Images were acquired with the Axion slide scanner 20x objective and image zoom details made in Zen Blue software.

In summary, presented data confirmed a heterogeneous expression profile for SOX2 and SOX9 proteins with prominent differences in the spatial distribution in two areas of most tumor samples. At the invasion front a positive correlation between both TFs was observed, while at the tumor core an inverse protein expression was evident.

#### **4.2.2. SOX9 Expression in the Tumor Microenvironment**

To investigate the SOX9 expression in stromal cells of the tumor microenvironment, SOX2<sup>High</sup>SOX9<sup>Low</sup> or vice versa, the percentage of SOX9 positive cells and staining intensity were quantified by the QuPath image software. Due to the differences in size of tumor samples, four areas of equal size per tumor were selected and were evaluated for SOX9 protein staining. Annotations for stromal and tumor areas, respectively, were made individually for selected areas of each sample. The software calculated the total percentage of positive cells and the intensity of staining in each cell, which finally revealed one H-score per sample based on a median of 25.000 cells per sample (Figure 21A).

Comparing SOX9 expression between stromal cells and tumor cells revealed a similar H-score for both compartments in tumor samples (Figure 21B). Moreover, this analysis demonstrated a positive correlation for the H-score between stromal cells tumor cells (Figure 21C). In summary, a significant number of stromal cells in the tumor microenvironment were SOX9 positive and the intensity and percentage were similar to tumor cells.

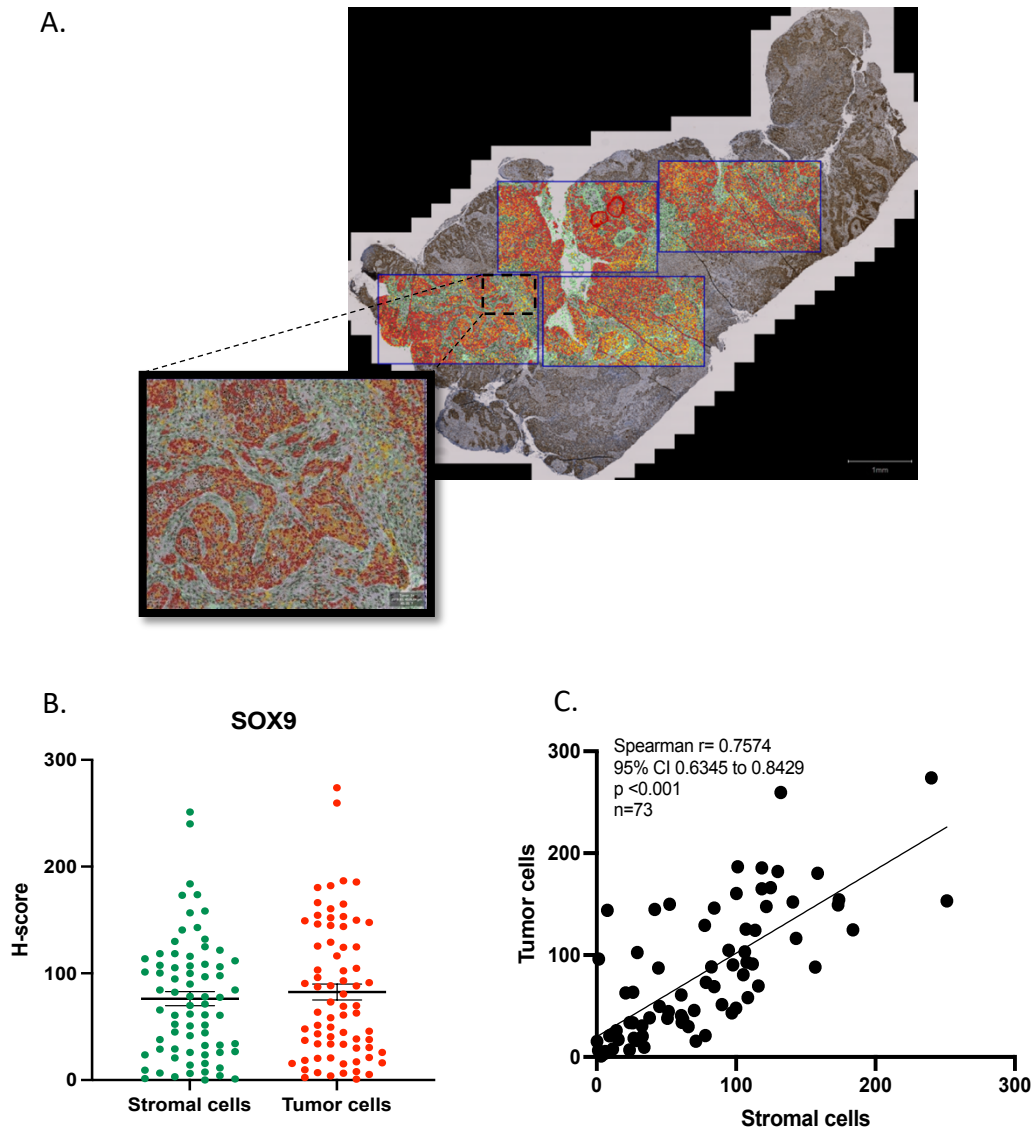


Figure 21. SOX9 Expression in The Tumor Microenvironment of samples from HIPO-HNC. A. Illustration of the representative areas of analysis (black rectangles) and annotations for stroma (green areas) and/or tumor cell (red areas) using QuPath image software. B. Dot plot shows H-scores for stromal and tumor cells in HNSCC samples. C. Spearman correlation between H-scores for SOX9 in tumor cells and stromal cells. Plots and statistical analysis were prepared in Prism 9.0. *P* value was considered statistical lower than 0.05.

#### 4.2.3. SOX2 and SOX9 Gene and Protein Expression are Associated in Patient Samples

To investigate, if gene transcript and protein expression levels were associated, transcriptome data were downloaded for the HIPO-HNC cohort (GSE117973). The Maxstat algorithm was applied to define clinically relevant cut-off points for SOX2 and SOX9 transcript levels, respectively, which were compared to IRS values at the

invasive front and tumor core. The results demonstrated that the investigated cases presented similar numbers between high/low gene and protein expression for both TFs (Table 3). In addition, a Fischer's exact test was applied to test whether the protein expression was dependent of the gene expression. These data confirmed a positive association between transcript levels and protein expression for both TFs.

FEATURE	GENE EXPRESSION	PROTEIN EXPRESSION			
		Invasion Front		Tumor Core	
	Total n=77	Total n=73	p	Total n=73	p
<b>SOX2</b>					
High	64	47	<b>0.0098</b>	51	<b>0.0363</b>
Low	13	26		22	
<b>SOX9</b>					
High	12	27	<b>0.0031</b>	26	<b>0.0081</b>
Low	65	46		47	

Table 3. Association between Gene and Protein Expression in the HIPO-HNC Cohort Samples. Fischer's exact test was performed to check a dependency between gene and protein expression in both areas analyzed in the tumor samples. Analyses were performed in R software and Prims 9.0 and p value was considered statistic lower than 0.05.

### 4.3. Impact of SOX2 Low Expression in Cellular Models

#### 4.3.1. Characterization of SOX2 and SOX9 Expression Pattern in HNSCC Cell Lines

Six well-established HNSCC cell lines (Cal27, FaDu, SCC4, SCC25, SAS and UT-SCC-5) were selected to explore SOX2 and SOX9 expression patterns *in vitro*. To assess the correlation between SOX2 and SOX9 transcription in these cell lines, gene transcript values were accessed by qRT-PCR after normalization with housekeeping genes and plotted for the six cell lines. Interesting to note that the FaDu cell line was the only one presenting a SOX2 amplification status (CCLE, 2019)<sup>119</sup>. The analysis revealed a tendency towards a negative correlation between both TFs (Spearman's  $r = -0.657$  and  $p=0.1$ ) (Figure 22A). Moreover, differences in transcript levels as determined by qRT-PCR (Figure 22B) were largely confirmed by a heterogeneous protein expression pattern for both TFs as determined by western blot analysis (Figure 22C). SCC4 and SAS cell lines presented highest SOX9 transcript and protein levels, while SOX2 was hardly detectable. Cal27 and FaDu cell lines shared a similar expression pattern of both TFs with moderate expression values. In contrast, SCC25 and UT-SCC5 cell lines presented highest SOX2 expression, but only minor SOX9 expression.

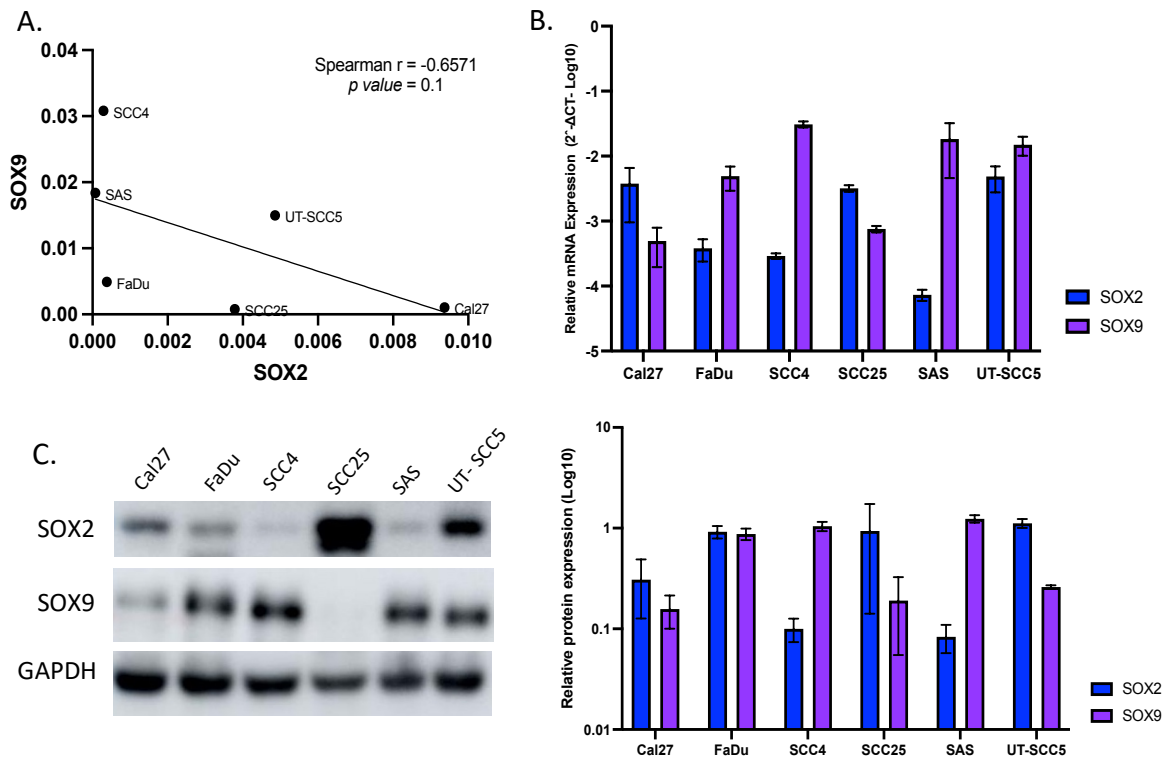


Figure 22. SOX2 and SOX9 Expression Pattern in HNSCC Cell Lines. A. Correlation of SOX2 and SOX9 mRNA transcript levels for the selected cell lines. Spearman's correlation coefficient was calculated for gene expression values. B. Relative mRNA levels determined by qRT-PCR in HNSCC cell lines. The relative gene expression of the genes of interest was calculated after normalization for the expression of a housekeeping gene (GAPDH) using  $2^{-\Delta CT}$  method. C. Protein expression of SOX2 and SOX9 in the six HNSCC cell lines. The relative protein expression for the genes of interest was calculated after normalization for the expression of GAPDH. Statistical analysis was performed in GraphPad Prism 9.0. Three biological replicates were analyzed.

The SOX9 expression was further analyzed in a set of five Oral Squamous Cell Carcinoma (OSCC) cell lines with known SOX2 amplification status (Nink *et al.* 2003). Of note, one out of five cell lines did not present SOX2 amplification status (HNO199). Western blot analysis (Figure 23A and B) demonstrated higher SOX9 protein levels for HNO97 and HNO222, while HNO199, HNO210 and HNO223 demonstrated similar levels of both TFs in protein expression. qRT-PCR (Figure 23C) analysis demonstrated higher SOX9 and lower SOX2 expression levels for HNO97, HNO199, HNO210 and HNO222 cells, while HNO223 cell line showed similar levels for both TFs. Moreover, cell lines with low SOX2 and higher SOX9 expression patterns presented a mesenchymal phenotype (Figure 23D).

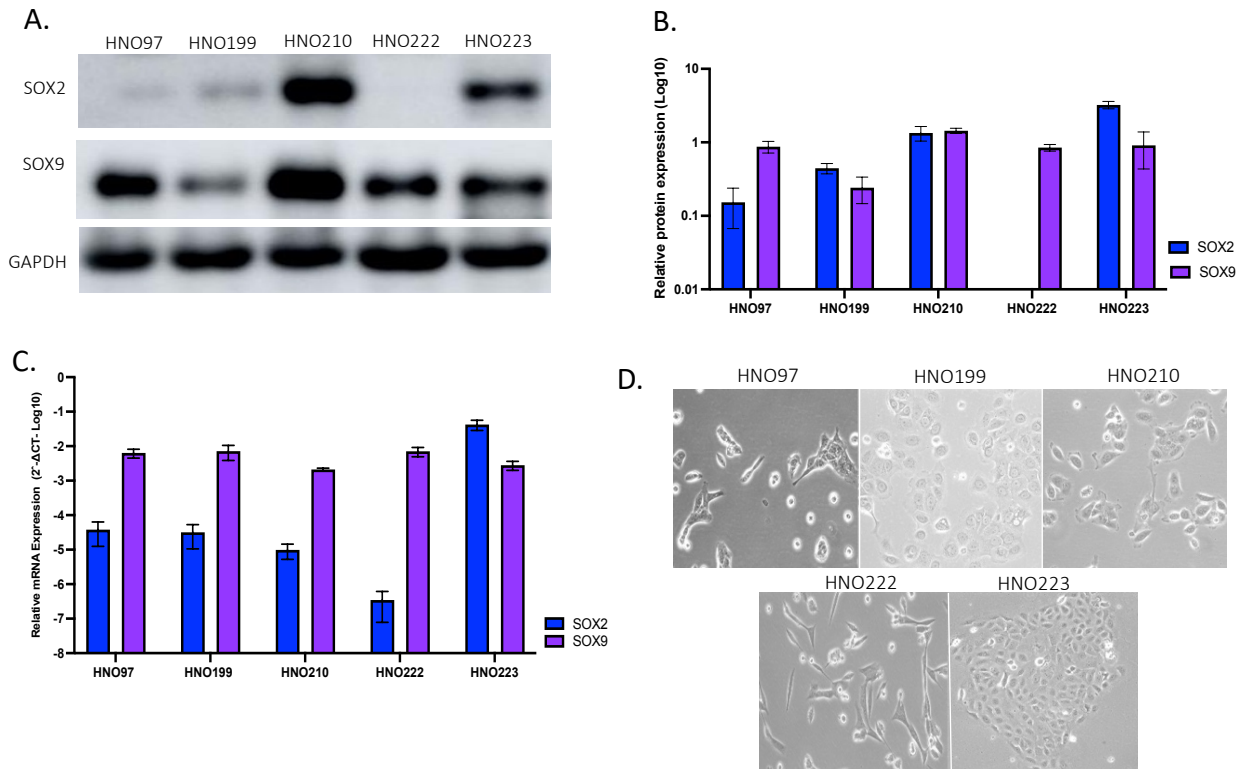


Figure 23. SOX2 and SOX9 Expression Pattern in OSCC Cell Lines. A and B. Western Blotting for SOX2 and SOX9 protein expression in HNO cell lines. The relative protein expression for the genes of interest was calculated after normalization for the expression of GAPDH. C. Relative mRNA levels determined by qRT-PCR in OSCC cell lines. The relative gene expression of the genes of interest was calculated after normalization for the expression of a housekeeping gene (GAPDH) using  $2^{-\Delta CT}$  method. D. Phase-contrast images of HNO cell lines. Statistical analysis was performed in GraphPad Prism 9.0. Three biological replicates were analyzed.

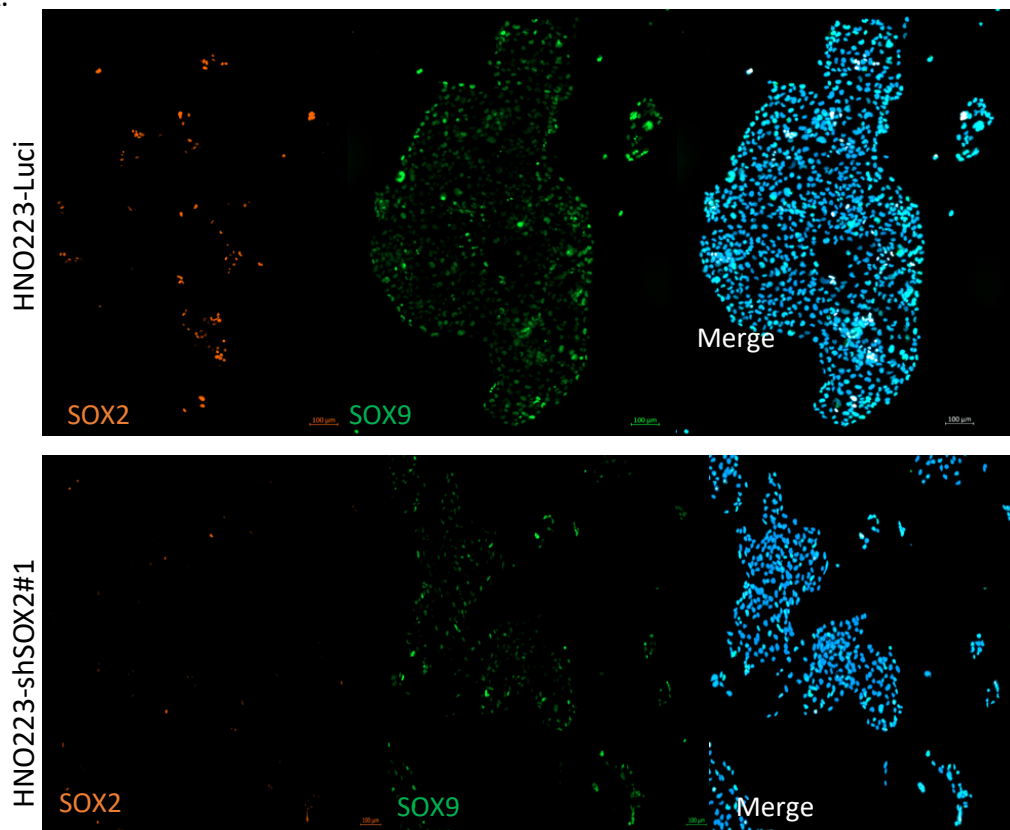
Altogether, SOX2 and SOX9 expression presented a heterogeneous expression pattern in most HNSCC cell lines investigated. A prominent and inverse expression pattern was observed in four well-established HNSCC cell lines (SCC4, SCC25, SAS and UT-SCC5), and in two OSCC cell lines (HNO97 and HNO222).

#### 4.3.2. SOX2 Expression and Response to Irradiation

HNSCC patients of the TCGA-HNSCC cohort with a  $SOX2^{Low}SOX9^{High}$  expression pattern had a reduced survival probability (Figure 13), indicating a possibly higher sensitivity of SOX2 expressing tumor cells against ionizing irradiation. To address this assumption, a previously established and published HNO223 cell line (HNO223-shSOX2) was used, in which SOX2 expression was silenced by a short hairpin RNA (shRNA) sequence and was compared to control cells (HNO223-Luci) expressing a shRNA sequence for Luciferase (non-target gene)<sup>57</sup>.

First, the impact of SOX2 silencing on SOX9 expression was determined by immunofluorescence (IF) staining, western blot (WB) and qRT-PCR analysis with HNO223-Luci and HNO223-shSOX2 cells. IF staining of HNO223-shSOX2 cells showed a reduction in SOX2-positive nuclei (orange signal) as compared to HNO223-Luci controls, while the amounts of SOX9-positive nuclei (green signal) was not affected (Figure 24A). This finding was confirmed by qRT-PCR and WB analyzes (Figure 24B). In line with a previous study, HNO223-Luci cells presented an epithelial morphology, while SOX2 silencing in HNO223-shSOX2#1 induced an mesenchymal-like morphology (Figure 24C) <sup>57</sup>. In summary, silencing of SOX2 expression did not affect SOX9 transcript or protein levels in HNO223 cells, which were selected as a model to further *in vitro* assays.

A.





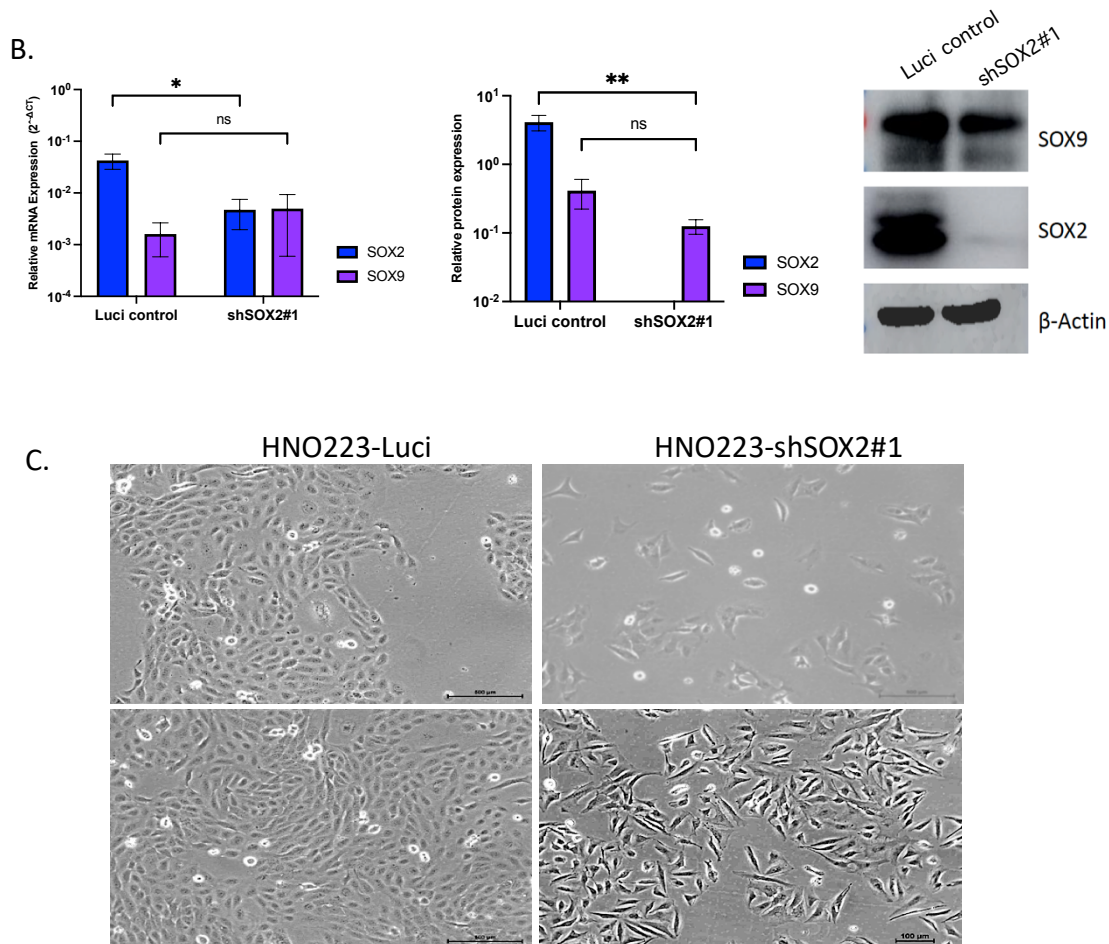


Figure 24. SOX2 Silencing in a HNSCC Cell Line. A. Representative images of an immunofluorescence staining with HNO223-Luci cells (upper panel) and HNO223-shSOX2 cells (lower panel). Images were taken with the Axion slide scanner 20x objective and image zoom details made in Zen Blue software. B. Left – Relative mRNA expression levels for SOX2 and SOX9 in HNO223 cells comparing control (Luci) vs. knockdown (shSOX2). The relative gene expression of the genes of interest was calculated after normalization for the expression of a housekeeping gene (GAPDH) using  $2^{-\Delta CT}$  method. Right – WB analysis for both TFs in HNO223 cell lines. HNO223-Luci cells presented similar levels of both proteins, but silenced cells showed loss of SOX2 expression (HNO223-shSOX2#1). The relative protein expression for the genes of interest was calculated after normalization for the expression of Beta Actin. C. Cell morphology of HNO223-Luci (left panel) and HNO223-shSOX2 cells (right panel). Images were taken in Inverted Zeiss microscope Axiovert 100 using 20x objective. Statistical analysis was performed in GraphPad Prism 9.0, *p value* was considered significantly, when lower than 0,05 and the bars represent SEM. Three biological replicates were analyzed. Two-way ANOVA was used as statistical test. ns = not significant; \**p* < 0.05 and \*\**p* < 0.01.

To investigate the impact of ionizing irradiation on HNO223-shSOX2 cells and HNO223-Luci controls in 2D and 3D cell culture models, a clonogenic assay was performed with either a single dose ranging from 2-8 Gy or a fractionated protocol consisting of a daily dose of 2 Gy for a 5-days period (total dose of 10 Gy), followed by a recovery time of 5 days. In the 2D cell model, the HNO223-shSOX2 cells showed a significantly higher clonogenic survival after a single dose irradiation as compared to HNO223-Luci controls cells as well as two other cell lines (FaDu and Cal27; Figure 25A). However, a highly heterogenous plating efficiency was observed in 2D, which was particularly low for HNO223-Luci cells (Figure 25B).

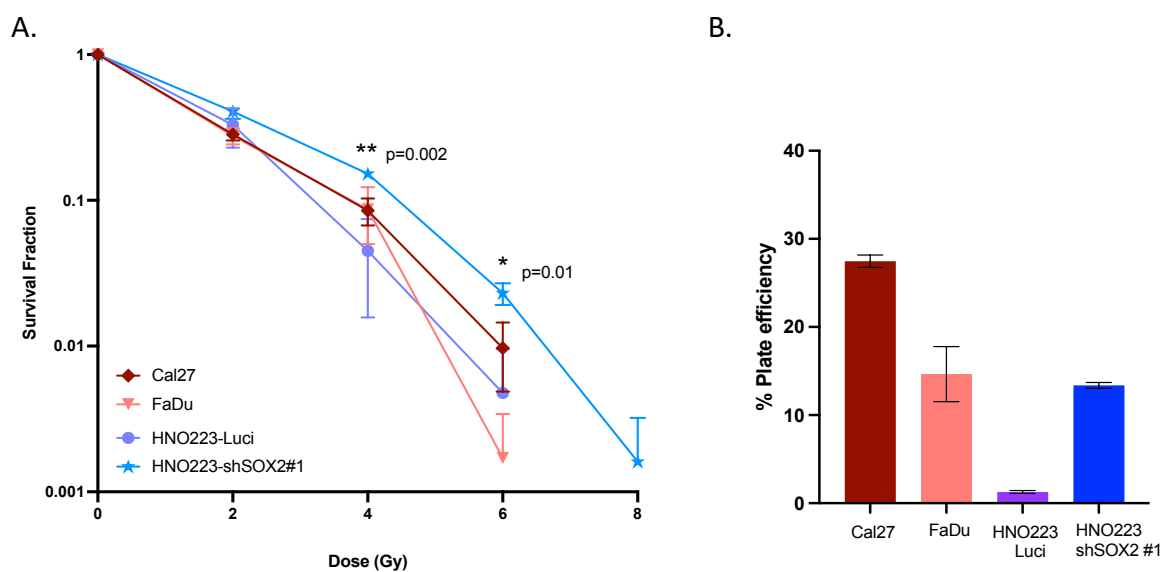


Figure 25. SOX2 Silencing in a HNSCC Cell Lines and Response to irradiation. A. 2D clonogenic assay for a single dose of irradiation of HNO223 cell lines, Cal27 and FaDu cells. B. Plating efficiency of the HNO223 cell lines, Cal27 and FaDu. Two-way ANOVA statistical analysis was performed in GraphPad Prism 9.0, *p* value was considered significantly, when lower than 0,05 and the bars represent SEM. Three biological replicates were analyzed.

To better explore the effect of ionizing irradiation on HNO223 cells, 2D and 3D clonogenic assay were performed with either in a single dose of irradiation or in a fractionated dose of irradiation (Figure 26 A and D). The data confirmed in the 2D and 3D clonogenic assays a higher survival probability of HNO223-shSOX2 cells as compared to HNO-Luci controls in both protocols of irradiation (Figure 26B-C and E-F). Since the HNO223-Luci controls showed lower plating efficiency in 2D, the parental HNO223 cell line was included as an additional control (HNO223 wild-type) for the 3D clonogenic assays.

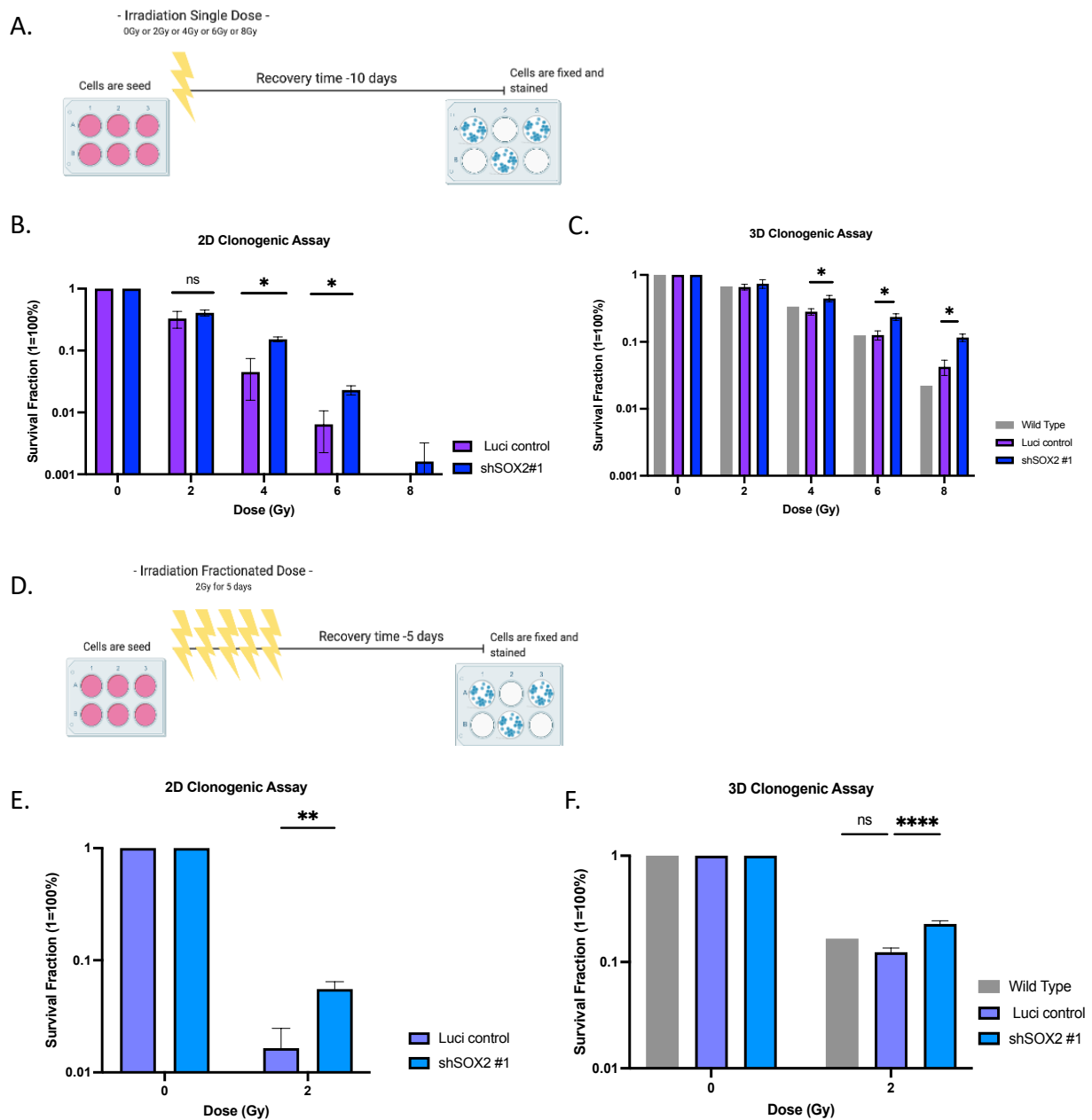


Figure 26. Irradiation Effect on HNO223 Cell Lines. A. Schematic representation of the single dose irradiation treatment schedule ranging from 2Gy to 8Gy. B. 2D assay for single dose. C. 3D assay for a single dose. D. Schematic representation of the irradiation treatment schedule with a fractionated dose of 2Gy daily for a 5 days period followed by a recovery time of 5 days. E. 2D assay for a fractionated dose. F. 3D assay for a fractionated dose. Two-way ANOVA statistical analysis was performed in GraphPad Prism 9.0, *p* value was considered significantly, when lower than 0,05 and the bars represent SEM. Three biological replicates were analyzed. ns, not significant; \**p* < 0.05; \*\**p* < 0.01 and \*\*\**p* < 0.001. For the HO223-WT the analysis was performed just in one biological replicate.

SOX2 and SOX9 serve as biomarkers for CSCs, which might explain variable survival upon treatment with irradiation. Hence, a Matrigel 3D assay was conducted to explore difference in tumor-initiating capacity between HNO223-shSOX2 and

HNO223-Luci controls. After 10 days of the cell seeding with equal cell concentration per well, the numbers of the colonies were counted and the area of each colony was measured using ImageJ software. The data demonstrated a significantly higher number of colonies for HNO223-shSOX2 cells (Figure 27A and C- left), but with a significantly smaller colony size when compared to HNO223-Luci control cells (Figure 27B and C - right).

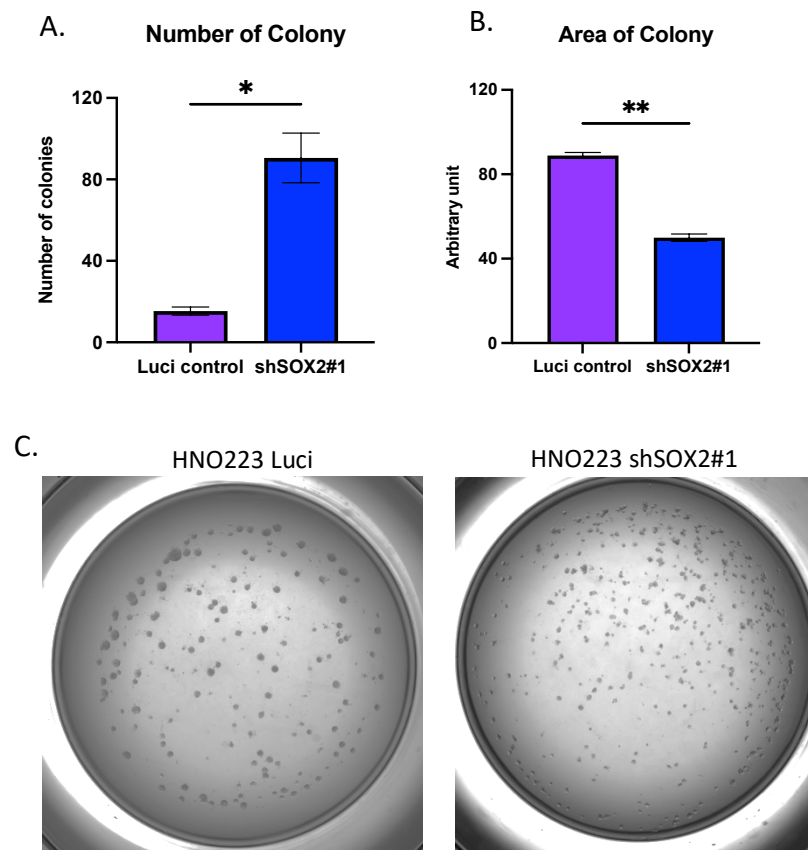


Figure 27. Tumor Initiating Capacity of HNO223 Cell Lines. A. The number of colonies in each well was counted and the average was represented in the graph. B. The colonies were considered bigger than 50  $\mu\text{m}$ , the area was measured per colonies and the average of areas in the well was represented in the graph. C. Representative pictures of a Matrigel 3D assay. Images were taken using a Nikon Eclipse Microscope 2x objective and image zoom details were made in ImageJ software. T-test statistical analysis was performed in GraphPad Prism 9.0, *p value* was considered significantly, when lower than 0.05 and the bars represent SEM. Three biological replicates were analyzed. \* $p < 0.05$  and \*\* $p < 0.01$ .

In summary, HNO223-shSOX2 cells presented a radioresistance phenotype, which might explain an unfavorable clinical prognosis of patients in the TCGA-HNSCC with lower SOX2 expression upon radiotherapy. The higher radioresistance might be linked to the tumor initiating capacity caused by SOX2 silencing.

### 4.3.3. Impact of Irradiation on SOX2 and SOX9 Expression

To investigate, whether ionizing irradiation affects SOX2 and/or SOX9 expression in the resistant and repopulating cancer cells at transcript or protein levels, HNO223 cells were subjected to the fractionated irradiation protocol (5 days with 2Gy daily), and cells were harvested in the end of recovery period of 5 days. qRT-PCR revealed an irradiation-induced reduction in SOX2 mRNA levels in both HNO223-shSOX2 and HNO223-Luci controls (Figure 28A). However, SOX2 mRNA levels were not affected by fractionated irradiation of Cal27 and FaDu cells. An irradiation-induced reduction was also evident at protein level for all cancer cell lines with detectable SOX2 protein expression (Figure 28C). In contrast, no statistically significant difference was detected for SOX9 mRNA expression in any of the investigated cell lines (Figure 28B), but an increase in SOX9 protein levels was evident after irradiation in all studied cell lines (Figure 28D). Taken together, ionizing irradiation decreased SOX2 transcript levels in both HNO223 cell lines and protein levels for the four cancer cell lines, whereas SOX9 protein levels increased despite no major changes on transcript levels in all cell lines.

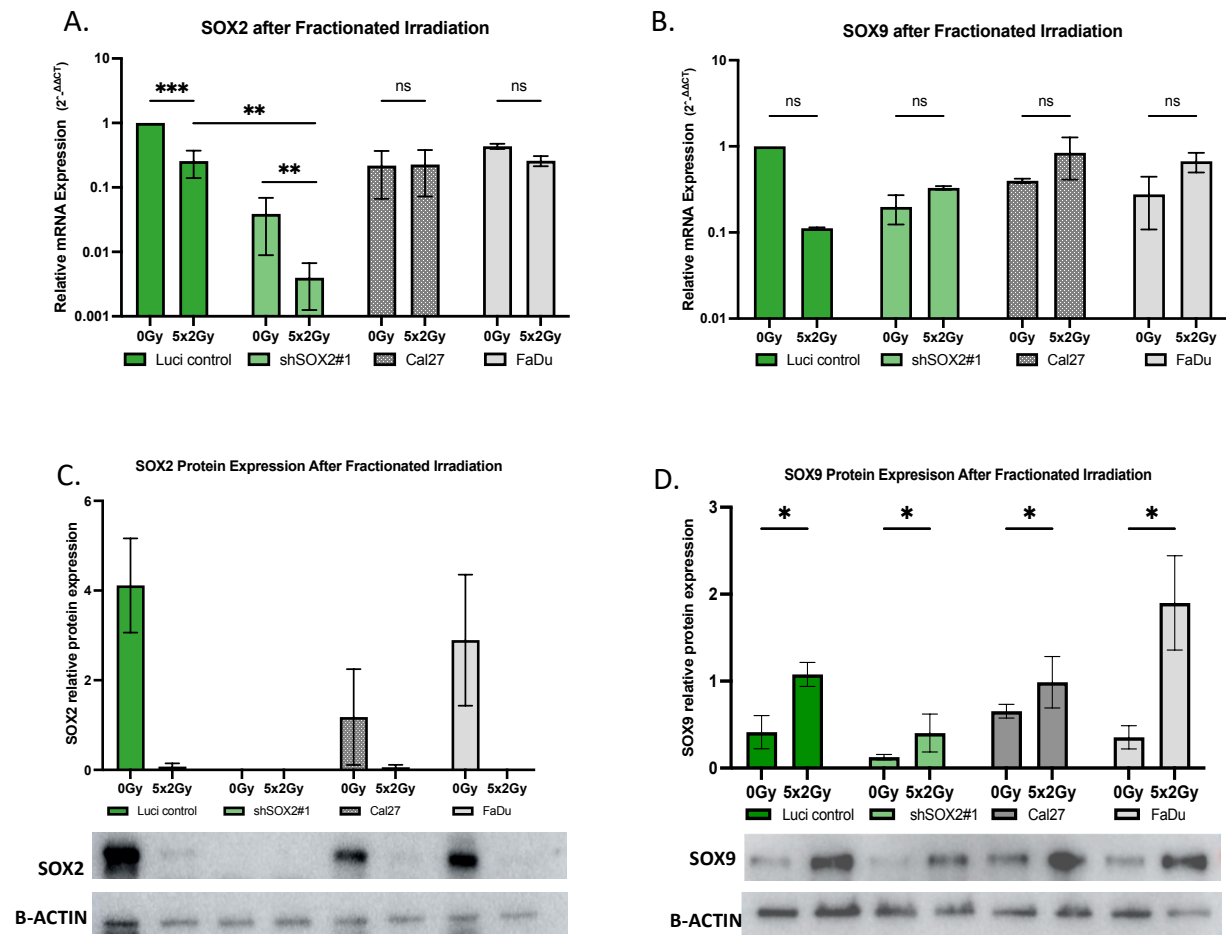


Figure 28. SOX2 and SOX9 Expression after fractionated irradiation. A. qRT-PCR for SOX2 mRNA expression in HNO223-Luci, HNO-shSOX2, Cal27 and FaDu cell lines after a fractionated irradiation protocol. B. qRT-PCR for SOX9 mRNA expression in HNO223-Luci, HNO223-shSOX2, Cal27 and FaDu cell lines after fractionated irradiation. The relative gene expression of the genes of interest was calculated after normalization for the expression of a housekeeping genes (GAPDH and ACTB) using  $2^{-\Delta\Delta CT}$  method. C. WB analysis for SOX2 protein in HNO223-Luci, HNO223-shSOX2, Cal27 and FaDu cell lines after a fractionated irradiation protocol. D. WB analysis for SOX9 in HNO223, Cal27 and FaDu cell lines after fractionated irradiation. The relative protein expression for the genes of interest was calculated after normalization for the expression of Beta Actin. Statistical analysis was performed in GraphPad Prism 9, *p value* was considered significantly, when lower than 0.05 and the bars represent SEM. ns = not significant; \**p* < 0.05; \*\**p* < 0.01 and \*\*\**p* < 0.001. Two-Ways ANOVA was used as statistical test. Three biological replicates were analyzed.

Finally, the expression of key genes of the TGF- $\beta$  signaling pathway, which were part of the 69-gene signature (Supplementary Table 1) and were highly expressed in tumors of the TCGA-HNSC cohort with SOX2<sup>Low</sup>SOX9<sup>High</sup> phenotype were analyzed in the cell lines with or without ionizing irradiation. qRT-PCR revealed an increase in INHBA gene expression after fractionated irradiation in all cell lines studied (Figure 29A). However, no irradiation-induced change was observed for SERPINE1 gene expression (Figure 29B).

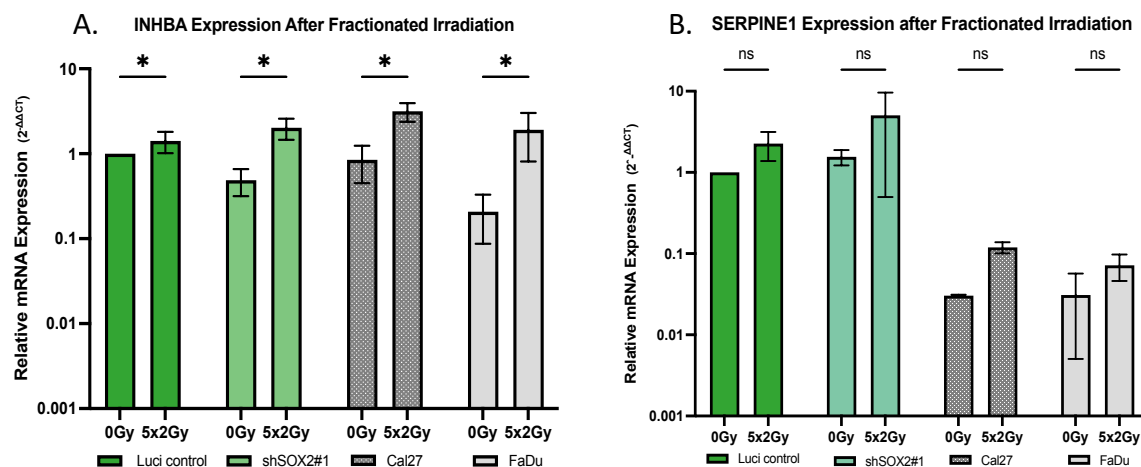


Figure 29: INHBA and SERPINE1 Expression after Fractionated Irradiation. A. qRT-PCR for INHBA mRNA expression in HNO223-Luci, HNO223-shSOX2, Cal27 and FaDu cell lines after fractionated irradiation. B. qRT-PCR for SERPINE1 mRNA expression in HNO223-Luci, HNO223-shSOX2, Cal27 and FaDu cell lines after fractionated irradiation. The relative gene expression of the genes of interest was calculated after normalization for the expression of a housekeeping genes (GAPDH and ACTB) using  $2^{-\Delta\Delta CT}$  method. Statistical analysis was performed in GraphPad Prism 9.0, *p value* was considered

significantly, when lower than 0.05 and the bars represent SEM. ns = not significant; \*p < 0.05. Two-Ways ANOVA was used as statistical test. Three biological replicates were analyzed.

SOX2 protein levels presented irradiation-induced reduction in all cell lines studied. In contrast, an increase in SOX9 protein expression was detected after fractionated irradiation without major changes in mRNA expression. IHNBA, one candidate gene which is associated with worse prognosis and radioresistance in HNSCC<sup>120,121</sup> was highly expressed after irradiation. These data supported the bioinformatic predictions, where patients with low SOX2 but higher INHBA expression presented a higher resistance to radiotherapy. Moreover, these findings indicate a potential role of irradiation on the TGF- $\beta$  signaling pathway in HNSCC cell lines with inverse SOX2 and SOX9 expression. However, future experimental studies are required to decipher in more detail the underlying mechanism.

#### **4.4. Impact of Altered SOX9 Expression in Cellular Models**

Presented data indicated that HNO223-shSOX2 cells presented a higher tumor-initiation capacity and radioresistance. However, SOX2 silencing was not associated with a change in SOX9 levels. Hence, HNO223 cells with SOX9 silencing were generated to investigate its impact on SOX2 expression and its able to restore the sensitivity to irradiation. Briefly, to generate the SOX9 knockdown in HNO223 cell line the *SMART Inducible Lentiviral shRNA vector*<sup>®</sup> was used according to the manufacturer's instructions. Three different sequences of SOX9 targets (#1, #2 and #3) and a non-targeting sequence were used. The transfection with viral particles were performed at the DKFZ Genomics and Proteomics Core Facility. The selection of positively transfected cells was established by subsequent puromycin selection and expansion for further experiments. The induction of SOX9 silencing was performed by treatment with 1  $\mu$ g/ml of Doxycycline for at least 72 hours prior to start experimental protocol.

In the first step, the consequence of SOX9 silencing on changes in cell morphology was explored by microscopic inspection after doxycycline treatment. However, HNO223-shSOX9 cells kept the epithelial morphology and no obvious differences were detected as compared to the parental wild type (WT) or non-targeted (NT) control cells (Figure 30A). As a second step, the levels of SOX9 protein expression were assessed by WB analysis. All specific targeting sequences after treatment with doxycycline were able to reduce SOX9 protein levels in the HNO223-shSOX9 cells as compared to WT controls, while no major differences were evident for NT controls (Figure 30B). As a third step, expression of well-established EMT markers in HNO223-shSOX9 cells, WT and NT controls were evaluated by WB analysis (Figure 30C, n=1).

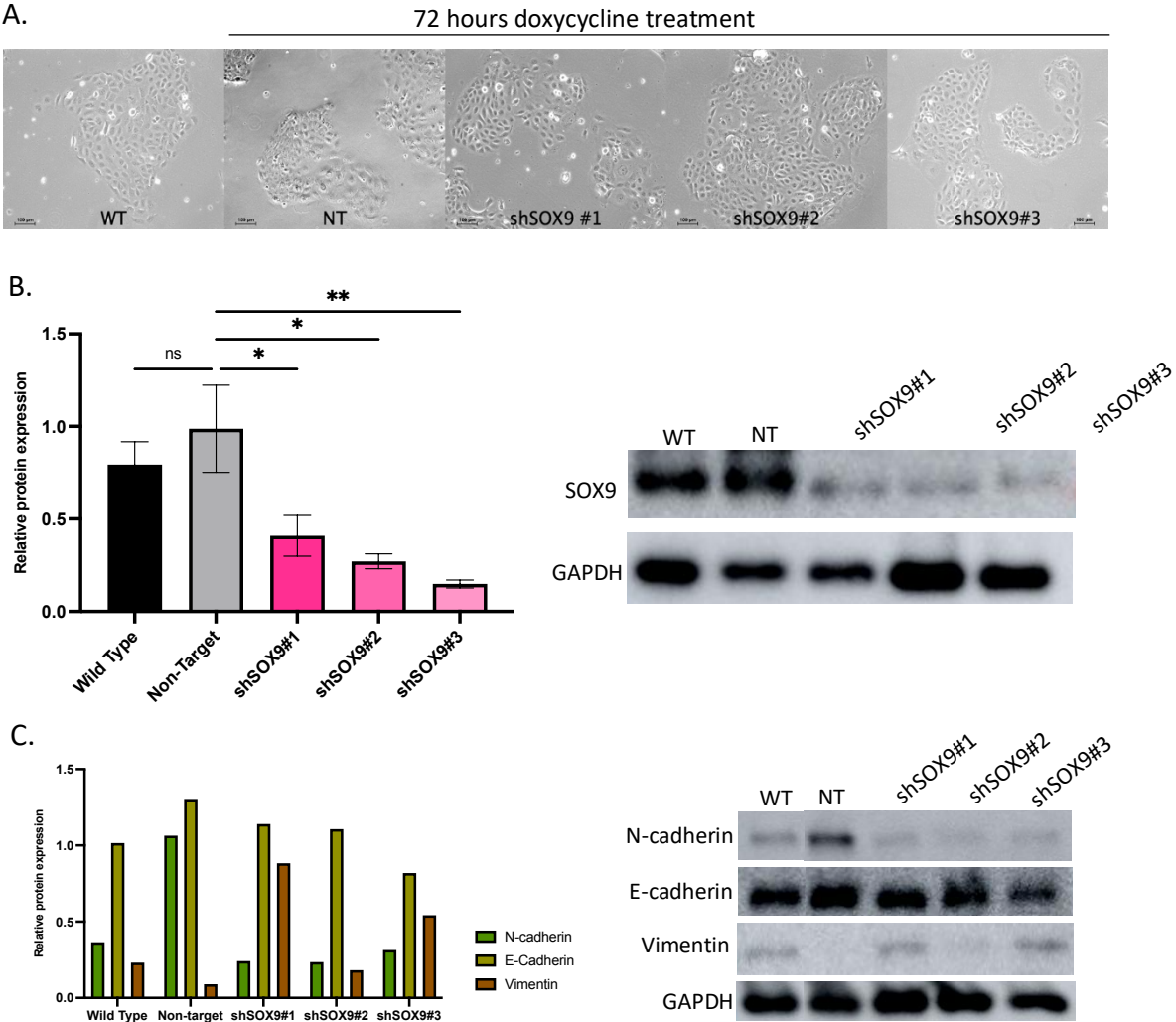


Figure 30. SOX9 Silencing in a HNSCC Cell Line. A. Phase contrast images of cell morphology for HNO223 parental wild type (WT), HNO223 non-targeted (NT) cells and the HNO223-shSOX9 #1, #2 and #3 clones. Doxycycline treatment was performed in NT and clones #1, #2 and #3 for at least 72



hours prior data acquisition. Images were taken in Inverted Zeiss microscope Axiovert 100 using 20x objective. B. WB analysis for SOX9 in HNO223-WT, HNO223-NT, HNO223-shSOX9#1, HNO223-shSOX9#2 and HNO223-SOX9#3 after doxycycline treatment. The relative protein expression for the gene of interest was calculated after normalization for the expression of GAPDH. HNO223 non-targeted cells presented similar levels of SOX9 protein as compared to the parental wild type (WT) cells. HNO223-shSOX9 clones presented reduction of SOX9 protein expression as compared to non-targeted cells. C. WB analysis of well-established EMT markers in HNO223-shSOX9 cell lines, WT and NT controls (n=1). The relative protein expression for the genes of interest was calculated after normalization for the expression of GAPDH. Statistical analysis was performed in GraphPad Prism 9.0, *p value* was considered significantly, when lower than 0.05 and the bars represent SEM. ns = not significant; \**p* < 0.05 and \*\**p* < 0.01. Three biological replicates were analyzed. One-Way ANOVA was used as a statistical test.

To address the question, whether SOX9 silencing promotes changes in SOX2 expression, SOX2 protein levels were assessed by WB analysis, which demonstrated similar levels of SOX2 in all three shSOX9 clones as well as controls (Figure 31).

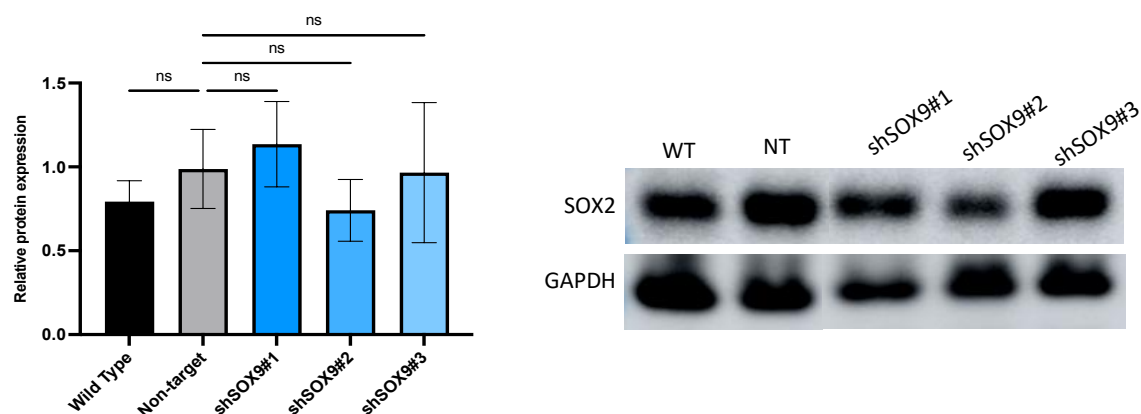


Figure 31. SOX2 Expression in SOX9 Silenced HNSCC Cell Lines. WB analysis for SOX2 in HNO223-shSOX9 cell lines after doxycycline treatment. No change in SOX2 protein expression was detected in silenced HNO-shSOX9 clones as compared to non-targeted cells. Also, the non-targeted did not present differences in SOX2 expression as compared to parental wild type (WT) cells. The relative protein expression for the gene of interest was calculated after normalization for the expression of GAPDH. Statistical analysis was performed in GraphPad Prism 9.0, *p value* was considered significantly, when lower than 0.05 and the bars represent SEM. ns = not significant. Three biological replicates were analyzed. One-Way ANOVA was used as a statistical test.

In breast and lung cancer it has been proposed that SOX9-Slug axis play a central role in regulating cancer stem cell properties, and also was correlated to poor clinical outcome<sup>122,123</sup>. Spearman's correlation revealed an inverse association between SLUG and SOX2 (Figure 32A), while a modest positive associated was observed for SOX9 (Figure 32B). To address the question, whether SOX9 silencing

promote changes in SLUG expression, SLUG protein levels were assessed by WB analysis. Initial and preliminary results, demonstrated reduction in Slug protein levels in two of three HNO223-shSX9 clones (#2 and #3), as compared to HNO223 non-targeted (NT). Parental wild type (WT) and HNO223-NT did not presented difference (Figure 32C). This initial association between SLUG and SOX9 was result of one experiment and need further experimental investigation.

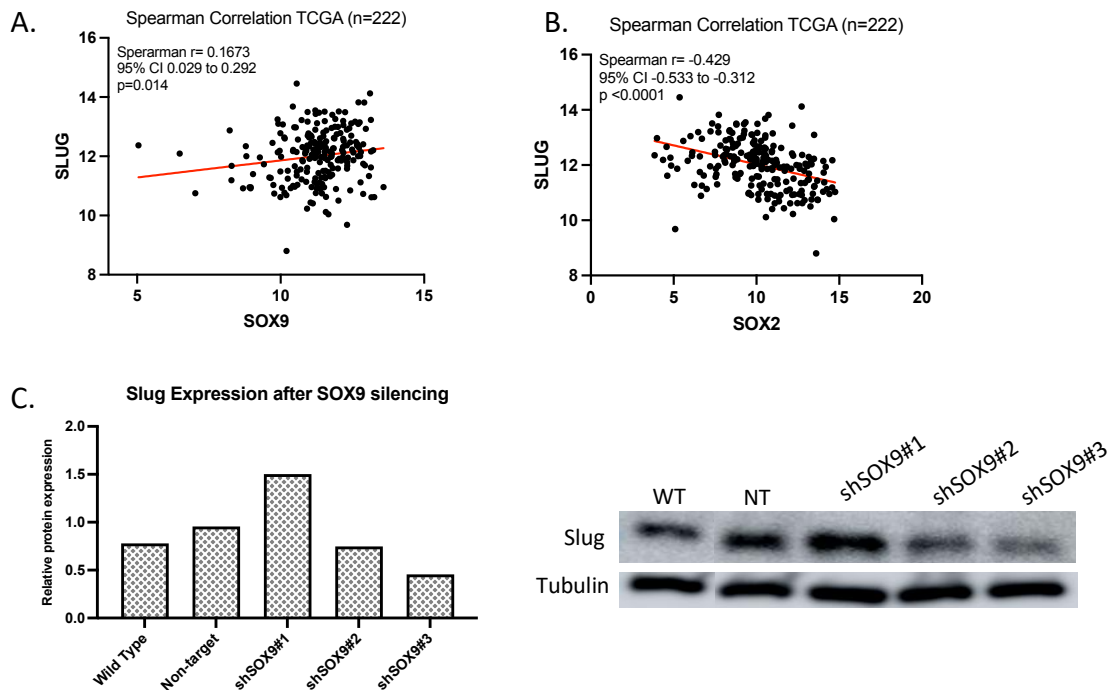


Figure 32. Slug Expression in SOX9 Silenced HNSCC Cell Lines. A. Spearman's correlation coefficient was calculated between SLUG and SOX9 gene expression values (mRNA level, RSEM). Dot plot represents HPV negative status patients treated with radiotherapy in TCGA-HNSCC (n=222). B. Spearman's correlation coefficient was calculated between SLUG and SOX2 gene expression values (mRNA level, RSEM). Dot plot represents HPV negative status patients treated with radiotherapy in TCGA-HNSCC (n=222). C. WB analysis of Slug protein expression in HNO223-shSOX9 and control cell lines. Reduction in Slug protein levels was observed in clone #2 and #3 as compared to non-targeted control. No difference was observed between HNO223-NT and HNO223-WT cells. The relative protein expression for the gene of interest was calculated after normalization for the expression of GAPDH. One biological replicate was analyzed. No statistical test was applied.

In summary, these data confirm successful establishment of three HNO223 clones with SOX9 silencing (shSOX9), which remain the epithelial cell morphology and high expression of E-cadherin, a classical marker of epithelial cells. Moreover, SOX9 silencing did not affect SOX2 protein expression. The exploratory association between SOX9 and Slug was observed to present a reduction in SLUG levels in two HNO223-

shSOX9 clones. Although, the mRNA expression levels, radiation properties, migratory behavior and predicted gene set need be further investigated in the HNO223-shSOX9.

#### 4.5. Downregulation of SOX2 and Immune Response

To further investigate the putative relation of SOX2 silencing with the predicted 69-gene signature and the modulation of the tumor immune microenvironment, the transcriptional profiles of HNO223-shSOX2 versus HNO223 Luci<sup>57</sup> were re-analyzed. Downregulation of genes related to the Antigen Presentation Machinery (APM) (Table 4) were observed in HNO223-shSOX2 as compared to HNO223-Luci controls. A recent study reported a selective pressure of immune surveillance on the modulation of SOX2 and SOX9 expression to enable tumor cell adaptation and immune escape<sup>68</sup>. While HLA class I classical (HLA-A, -B, and -C) are responsible to present antigenic peptide to cytotoxic T cells, the non-classical HLA-E and HLA-F mediate inhibitory and/or activation signaling to NK cells<sup>124</sup>. HLA-E can bind to NKG2A/B promoting an inhibitory signal to suppress NK cell cytotoxicity or it can bind to NKG2C resulting in an activating signal<sup>125</sup>. Bukur *et al.*, have shown that alterations in classical versus non-classical HLA proteins or other components of the APM provide tumor cells with the ability to evade the immune response promoting tumor growth and escape from immune surveillance<sup>126</sup>.

Gene	Log2FC	Adj.p-value	Location	Type(s)
<b>B2M</b>	-1,185	5,21E-02	Plasma Membrane	transmembrane receptor
<b>HLA-A</b>	-1,009	4,81E-02	Plasma Membrane	Other
<b>HLA-B</b>	-1,281	3,57E-02	Plasma Membrane	transmembrane receptor
<b>HLA-C</b>	-1,168	4,47E-02	Plasma Membrane	Other
<b>HLA-E</b>	-0,886	5,40E-02	Plasma Membrane	transmembrane receptor
<b>HLA-F</b>	-0,998	4,96E-02	Plasma Membrane	transmembrane receptor
<b>PSMB8</b>	-1,079	3,43E-02	Cytoplasm	Peptidase
<b>PSMB9</b>	-1,540	4,43E-02	Cytoplasm	Peptidase
<b>TAPBP</b>	-1,029	4,20E-02	Cytoplasm	Transporter

Table 4. List of Genes related to the Antigen Presentation Pathway after SOX2 Silencing as Assessed by Ingenuity Pathway Analysis. *Extracted from Jou, A. Thesis, 2014*<sup>127</sup>.

Antigen processing and presentation is a critical step for immune detection of tumor cells. Alterations in genes which are involved in these processes seems to be affected by transcriptional changes and altered regulation of HLA expression have

been reported for HNSCC<sup>128-130</sup>. To investigate the pathways involved in immune cells recognition, GSVA scores were computed for hallmark gene sets download from the Molecular Signature Data Base (MSigDB). Comparison of the groups G1 with SOX2<sup>Low</sup>SOX9<sup>High</sup> and G2 with SOX2<sup>High</sup>SOX9<sup>Low</sup> demonstrated a higher GSVA score for the APC-related gene set and suggested that other genes might be clinically relevant, when SOX2 is downregulated.

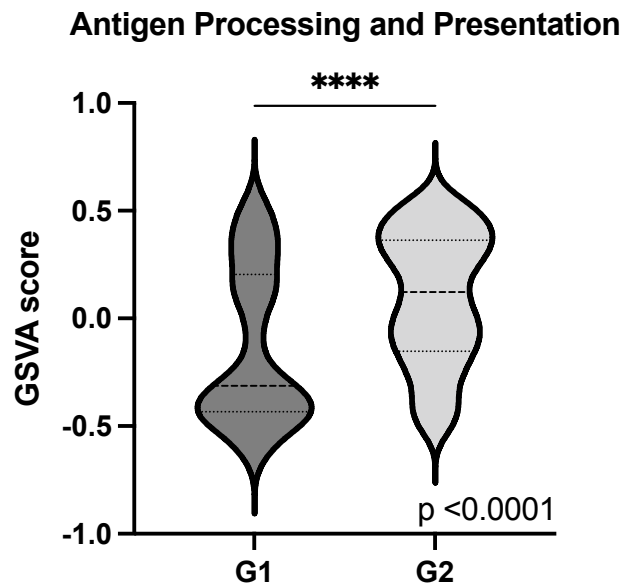


Figure 33. Antigen Presentation Machinery. GSVA score of the HPV negative status patients of the TCGA-HNSCC cohort (n=222) with radiotherapy. Statistical analysis was performed in R software and graph prepared in GraphPad Prism 9.0. *p* value was considered statistically significant, when lower than 0.05.

To further explore whether HLA-E reduction serves as a potential mechanism for SOX2 low cancer cells to evade the cytotoxicity of Natural Killer (NK) cells, HLA-E expression was investigated in HNO223 cells by WB analysis. However, no significant difference was detected between HNO223-shSOX2 cells and HNO223-Luci controls (Figure 34A). Flow Cytometry (FACs) analysis was performed to evaluate potential differences in HLA-E surface expression, but again no significant difference was evident between HNO223-shSOX2 cells and HNO223-Luci controls (Figure 34B).

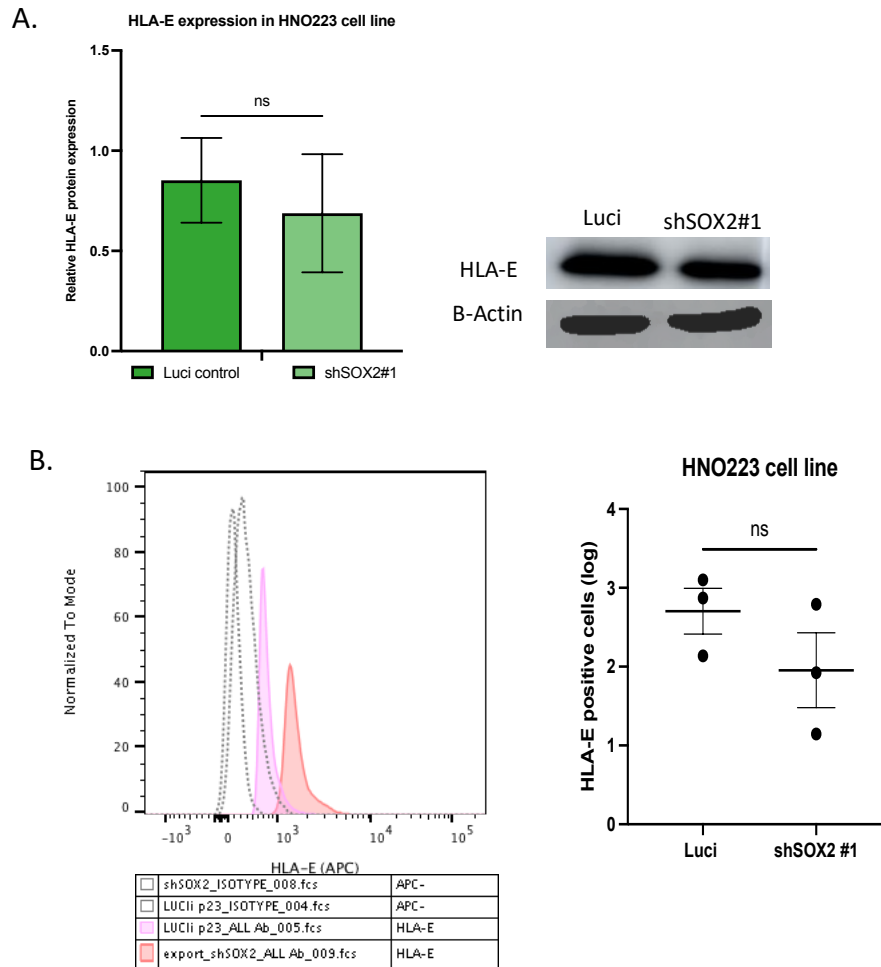


Figure 34: HLA-E Expression in HNO223 Cell Lines. A. Western blotting analysis for HLA-E expression in HNO223 cell lines. The relative protein expression for the gene of interest was calculated after normalization for the expression of Beta Actin. B. Flow cytometry analysis of HLA-E expression on the cellular surface of HNO223 cell lines. Pink curve represents the HNO223-Luci controls and the red curves the HNO223-shSOX2 cells. Dashed lines represented the isotype control of each cell line. T-test statistical analysis was performed in GraphPad Prism 9.0, p value was considered significantly, when lower than 0.05, and the bars represent SEM. Three biological replicates were analyzed. FACS was analyzed with FlowJo software. ns = not significant.

Studies demonstrated that CSCs can downregulate MHC molecules leading to inhibition of immune response<sup>128,131,132</sup>. As HNO223-shSOX2 cells presented a higher tumor-infiltration capacity, which was correlated with CSCs properties, the hypothesis that these cells could evade tumor immunity by overexpression HLA-E after irradiation was explored by a co-culture with NK cells. To define the base line of cytotoxic killing the NK cells were co-cultured with either HNO223 cell lines, SCC25 or Cal27 cells for a period of 16 hours. The NK cells were purified from buffy coat of healthy blood samples and stimulated overnight with IL-2. However, only a minor cytotoxicity (1-2%

of cellular killing) were observed under co-culture conditions (Figure 35A). Next, the impact of fractionated irradiation on HLA-E expression was determined for HNO223 cell lines, FaDu and Cal27 cells by WB analysis. A minor increase in HLA-E protein expression was detected after irradiation as compared to non-irradiated HNO223 cell lines, which did not reach statistical significance and was independent of SOX2 expression (Figure 35B). In contrast, HLA-E protein expression was reduced after irradiation for Cal27 and FaDu cell lines, but again the difference was not statistically significant (Figure 35B). In conclusion these data did not support a regulation of HLA-E by SOX2 with or without ionizing irradiation in cancer cell lines *in vitro*.

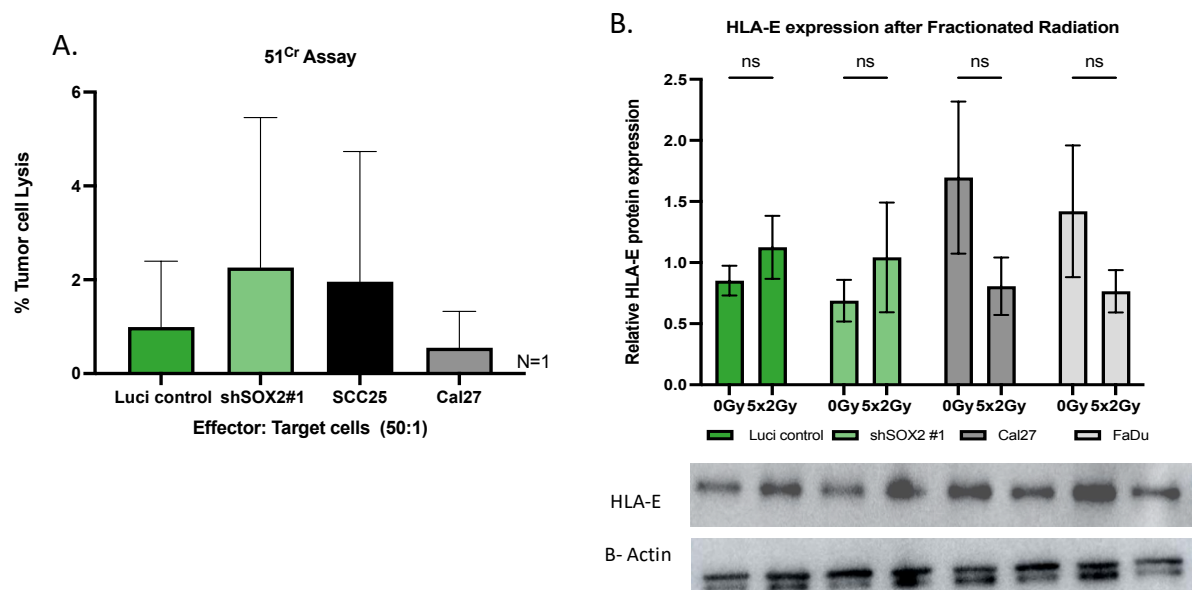
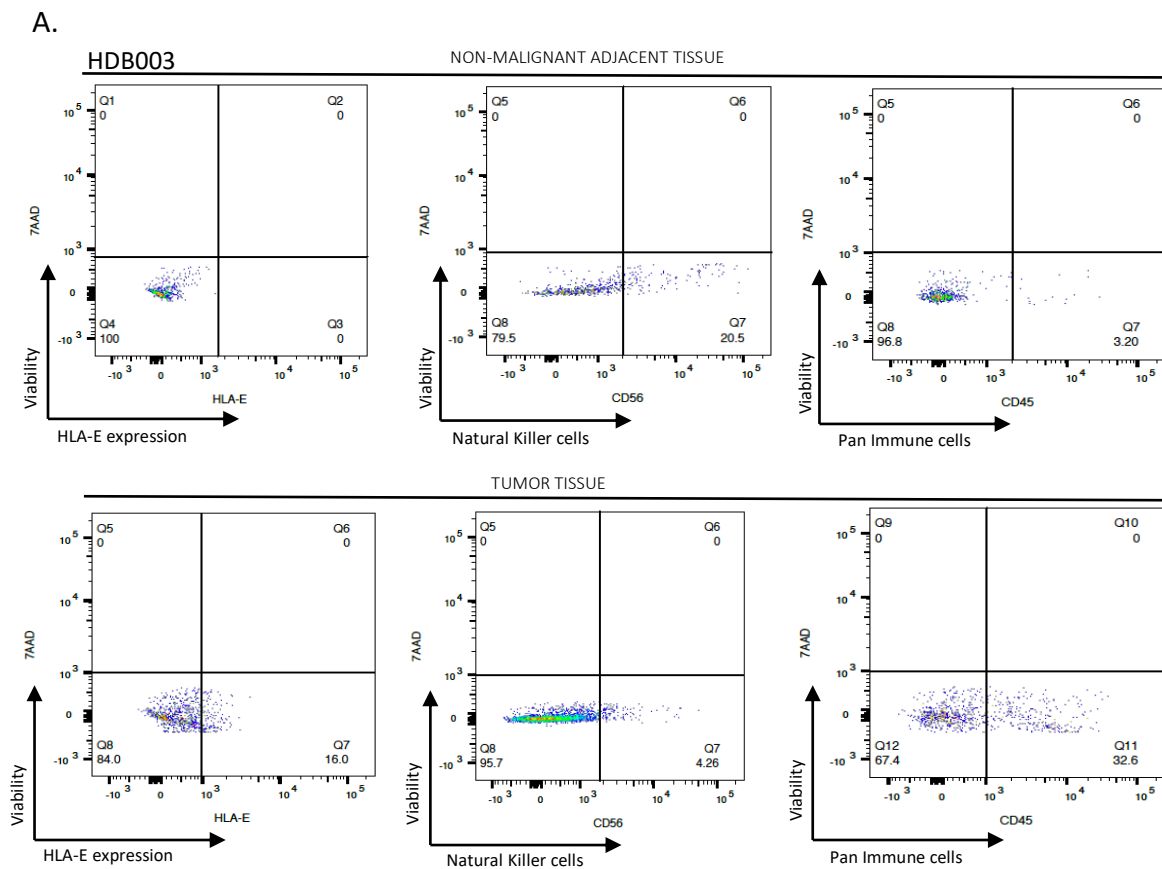


Figure 35. NK-related Cytotoxicity and HLA-E Expression in HNSCC Cell Lines under Irradiation. A. Chromium-51 assay for HNSCC cell lines in co-culture with isolated NK cells from buffy coat. One biological assay was performed due high toxicity of the reagent and restriction to radioactivity. B. Western blotting analysis for HLA-E expression in HNO223, Cal27 and FaDu cell lines after a fractionated irradiation. The relative protein expression for the gene of interest was calculated after normalization for the expression of Beta Actin. Statistical analysis was performed in GraphPad Prism 9, p value was considered significantly, when lower than 0.05 and the bars represent SEM. Three biological replicates were analyzed. Two-Ways ANOVA was used as statistical test. FACs was analyzed with FlowJo software.

Even with the support of transcriptional data, the mechanism of HLA-E action on the mutual interaction of HNSCC cells and NK cells was not evident in the current set-up of experimental protocols, indication a more complex and context dependent regulation mediated by the tumor microenvironment.

## 4.6. Single Cell Profiling with Tumor Samples

To further explore the potential relation between the inverse SOX2 and SOX9 expression in cancer cells and the composition of the tumor immune microenvironment, single cell profiling was performed. The initial aim was to evaluate the expression of HLA-E on tumor cells and the presence of immune cells in fresh tumor samples. The presence of immune cells was evaluated by Flow Cytometry analysis utilizing antibodies for CD45 (Pan Immune Cell) and CD56 (NK cells) as marker proteins. The viability dye 7-AAD was used to exclude the dead cell population. The full panel of markers was applied on cell suspension from two tumor samples (HDB003 and HDB004) and matched adjacent non-tumor controls. In both cases an increased amount of CD45-positive and CD56-positive immune cells was detected in the tumor tissue as compared with the adjacent non-tumor area. A lower fraction of HLA-E-positive cancer cells was detected in the tumor tissue as compared to the non-tumor adjacent area (Figure 36A and B). Moreover, for the sample HDB003 non-malignant adjacent tissue was available, which showed lower number of immune cells, when compared to the matched tumor tissue.



B. HDB004

TUMOR TISSUE

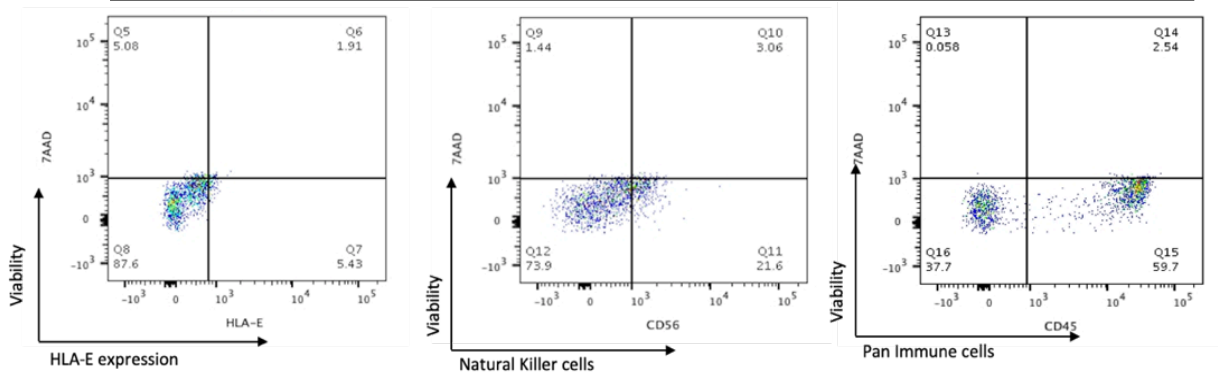


Figure 36: Tumor Dissociation Single Cell Analysis. A. The first row of the panel represented analysis for non-malignant adjacent tumor with the markers 7AAD (viability of the cells) and HLA-E followed by NK cells (CD56) and pan immune cells (CD45). The second row represents the analysis in the tumor tissue with the same panel of markers. B. Represented the analysis made in a different patient sample for the tumor tissue where was applied the panel of markers. Analysis were performed with FlowJo software.

In parallel, formalin fixed and paraffin embedded (FFPE) tumor samples were stained for SOX2 and SOX9 expression by IHC, and revealed a heterogeneous expression pattern for both TFs (Table 5). In addition, the presence of immune cell infiltration was accessed by Hematoxylin and Eosin (HE) staining and was evaluated as follows: 0-25%: Mild, 26-75%: Moderate and 76%-100%: Intense. The sample (HDB001) with lower and homogenous expression of both TFs presented a higher percentage of immune cells at the invasive front and tumor core. In contrast, the sample (HDB004) with moderate expression of both TFs at the invasive front, but low expression for the tumor core also presented a decreasing amount of immune cell infiltration at these sites. However, the size of this cohort was rather low to draw any final conclusion.

	Invasion Front			Tumor Core		
	SOX2	SOX9	Immune Infiltration	SOX2	SOX9	Immune Infiltration
HDB001	Low	Low	Intense	Low	Low	Intense
HDB002	Low	Moderate	Mild	Low	Moderate	Moderate
HDB003	Moderate	Low	Moderate	Moderate	Low	Mild
HDB004	Moderate	Moderate	Intense	Low	Low	Moderate

Table 5: Tumor Samples Collection. Descriptive analysis of the immunohistochemistry staining for SOX2 and SOX9 in samples from MKG – Heidelberg University Hospital. The score of percentage for positive cells are describing in 4 degrees (0-25%, 26-50%, 51-75% and 76-100%) and the intensity of immune cells are classified in 3 degrees (low, mild and intense).



To gain insight into the composition of the cytokine milieu, particularly in the tumor microenvironment, fresh tumor samples were subjected to a multiplex cytokine assay. The experiment was performed in collaboration with Dr. Alexander Rölle on two tumor samples (HDB002 and HDB004) and their matched non-malignant tissue. The results showed a similar profile of cytokines for both non-malignant tissues (Figure 37). In contrast, a difference in cytokine expression was evident for HDB002 and HDB004 samples (Figure 37). The heatmap demonstrated a difference in cytokine expressions for Interleukin-1 receptor antagonist (IL-1ra), Tumor necrosis factor alpha (TNF- $\alpha$ ), Macrophage inflammatory protein 1-beta (MIP1b or CCL4), Interferon gamma (IFN- $\gamma$ ) and chemokine ligand 1 (CXCL1 or GRO-a) (Figure 37). These cytokines decreased in protein levels in tumor tissue as compared to non-malignant samples (Figure 37). Interesting to note that these group of cytokines have been described as regulators of the immune and inflammatory response in tumors<sup>133,134</sup>. However, the limitation in sample size makes it difficult to draw a final conclusion about the complex network. Additional research and samples are needed to dissect potential associations between SOX2 and SOX9 expression with the immune system.

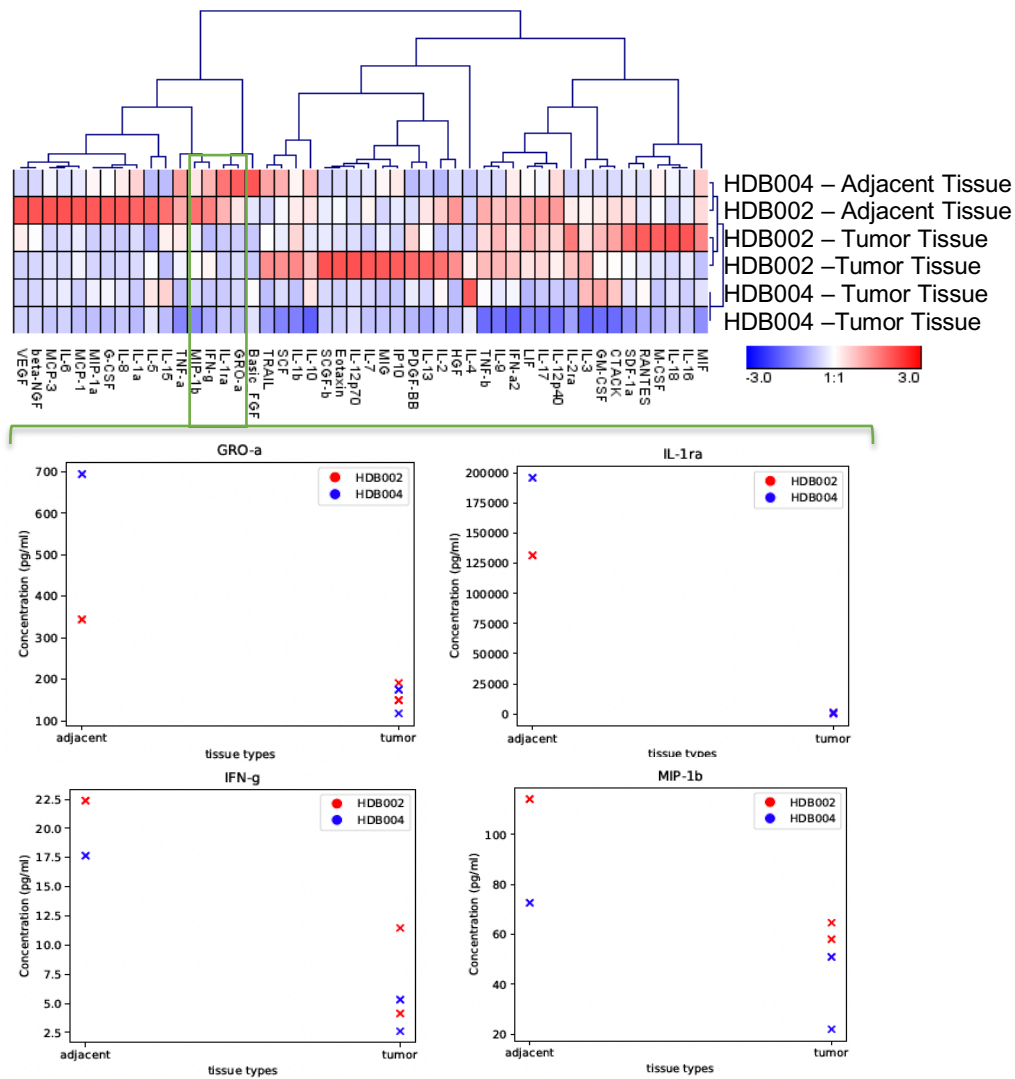


Figure 37: Cytokine Profile in Tumor Samples. Multiplex cytokine assay showed the similar profile between two samples of adjacent non-tumor area and opposite expression with the tumor samples.

In conclusion, presented data provide experimental evidence that a low SOX2 and high SOX9 expression serves as an unfavorable prognostic biomarker for HPV-negative status HNSCC with radiotherapy. Bioinformatics analysis established a signature gene set associated with these patients and defined TGF- $\beta$  signaling as a top canonical pathway. *In vitro* studies indicated a radioresistant phenotype of cancer cells upon SOX2 silencing and an up-regulation of the TGF- $\beta$  signaling-related candidate gene INHBA after irradiation. The single cell dissociation paves the way to investigate the communication of tumor cells with immune cells of the tumor microenvironment. The cytokine assay indicates a pro-inflammatory profile which might be linked to gene expression in the tumor microenvironment and related to high SOX9 expression in stromal cells.

## 5. Discussion

### 5.1. Computational Analysis of SOX2 and SOX9 as Potential Biomarkers in HNSCC

In recent years, the use of computational methods in personalized medicine allowed improvements in diagnosis and guidance towards a more precise and efficient therapy<sup>135,136</sup>. One of the most challenging issues was to define an ideal cut-off point considering gene expression without disregarding any patient in the cohort. Many software tools and statistical methods have been described to define the optimal cut-off point in gene expression, such as: CutOff Finder<sup>137</sup>, X-Tile<sup>138</sup>, Optimal Cutpoints<sup>139</sup> and Maxstat<sup>140</sup>. Among these tools, the most well accepted methodology is the maximally selected rank algorithm which defines a single cut-off point for a certain gene<sup>140</sup>. Here, two different analytical tools applying this strategy were used to establish a valid stratification for patients grouping according to inverse SOX2 and SOX9 expression in HNSCC tumor samples.

Initially, identifying an individual cut-off point for each gene using Maxstat reveals a higher degree of diversity into four sub-classes of patients for the TCGA-HNSCC cohort without clinical significance. These results could be explained by the differences in the distribution of SOX2 and SOX9 gene expression. Lin *et al.* reported that an epigenetic switch between SOX2 and SOX9 TFs modulates the plasticity of lung cancer cells<sup>65</sup>. The recent review of the hallmarks of cancer illustrated the importance of cellular plasticity, where cancer cells change from a differentiation state to an undifferentiated state. This plasticity allows cancer cells to reprogram their gene expression in the context to tumor progression and therapy resistance<sup>15</sup>. Thus, this simplified mathematical models are unlikely to capture the molecular complexity in the regulation and function of most TFs, including SOX2 and SOX9.

A secondary strategy was chosen with the aim to train a more accurate model, where gene expression of both TFs was integrated to define an ideal cut-off point. A ratio between SOX2 and SOX9 was obtained and subsequently Maxstat was applied. The second approach identified patients with a SOX2<sup>Low</sup>SOX9<sup>High</sup> phenotype and a tendency toward unfavorable clinical prognosis as compared to SOX2<sup>High</sup>SOX9<sup>Low</sup> counterparts. HNSCC is a multifactorial disease and using clinical parameters underscore accuracy to survival analysis, where patients with HPV negative status and

receiving radiotherapy presented a worse clinical prognosis. Thus, the inverse SOX2 and SOX9 expression pattern might help in defining the clinical prognosis of HNSCC patients, but other confounding factors also play important roles.

Changes in gene expression might reflect mutual interactions between cancer cells and stromal cells of the TME, causing alterations in cellular plasticity, metastasis and resistance to treatment<sup>66,68,141</sup>. A main hypothesis of this study was that the inverse regulation of these TFs might cause and/or reflect the consequences of altered gene expression in TME, and thereby regulates tissue/cell morphology. Therefore, a gene set signature was defined and TGF- $\beta$  signaling was predicted as top up-regulated pathway. INHBA, SERPINE1 and THSB1 genes were part of the gene set signature and highly expressed in the group of patients with a SOX2<sup>Low</sup>SOX9<sup>High</sup> phenotype. The TGF- $\beta$  signaling is a key process in tumorigenesis and regulates mechanisms either to suppress or promote tumor growth depending on the context<sup>76,142</sup>. Zhang *et al.* have shown that TGF- $\beta$  secreted by tumor associated macrophages induces expression of SOX9 leading to an EMT phenotype in lung cancer cells, which was linked to tumor proliferation, migration and invasion<sup>143</sup>. Another study with lung cancer cells, demonstrated the involvement of TGF- $\beta$  signaling in downregulation of SOX2 inducing EMT and promoting a change in cell morphology accomplished by a resistance to treatment<sup>56</sup>. In line with these reports, a high expression patterns of the gene set signature identified a subset of patient in the TCGA cohort with HPV- negative HNSCC of the TCGA cohort with poor clinical outcome after radiotherapy.

The presented results indicated that SOX9 protein levels and INHBA expression are upregulated after radiation therapy, while SOX2 is downregulated. It is important to note that secretion of soluble molecules in the TME may activate or repress many cascades in the adjacent tumor tissue and is one way to regulate cancer cell profile<sup>78</sup>. In HNSCC, expression of INHBA increases cellular migration, invasion and unfavorable clinical prognosis<sup>144</sup>. Members of the Activin receptor-like kinase (ALK) superfamily share ligands with TGF- $\beta$  signaling promoting downstream effects on phosphorylation of SMAD2/3 or SMAD1/5/8 to promote a complex formation with SMAD4 that translocates to the nucleus resulting in modulation of gene expression<sup>56,142,145</sup>.

Furthermore, integrin signaling, wound healing signaling, actin cytoskeleton signaling, Rho family pathways and other molecular mechanisms of cellular communication were enriched in the proposed gene set signature (Supplementary Table 4). Loomans *et al.*, demonstrated in squamous cell carcinoma an increased aggressiveness through modifications in cell-cell adhesion proteins caused by loss of epithelial Activin receptor type IB (ACVR1B). Loss of ACVR1B was inversely correlated with high stromal expression of INHBA, suggesting that this receptor modulates the tumor suppressor activity of INHBA<sup>146</sup>. Tsai *et al.*, reported that EGFR is activated by INHBA in OSCC cells via the non-canonical PI3K/SPI pathway, and high expression of EGFR was significantly associated with poor clinical prognosis<sup>103</sup>. In a recent study, a risk model based in inverse SOX2/SOX9 expression also identified EGFR signaling as a candidate for HNSCC patients with SOX2 low and SOX9 high expression pattern<sup>147</sup>.

Taken together, omics analyses presented in this study support the hypothesis that the inverse SOX2/SOX9 expression is involved in tumor progression and resistance to treatment. The latter issue is of particular importance for HPV-negative HNSCC, which are treated with radiotherapy. The functional importance of this inverse association might be linked to a gene signature, which is either the cause or the consequence of the interactions with stromal cells of the tumor environment via TGF- $\beta$  signaling. Presented data indicate that an increase in SOX9 levels and INHBA expression is linked with a radioresistant phenotype in cancer cell lines with low SOX2 levels. However, further investigations are required to elucidate the involvement of TGF- $\beta$  signaling, SMAD activation/repression and the involvement of gene signature with failure to radiation treatment.

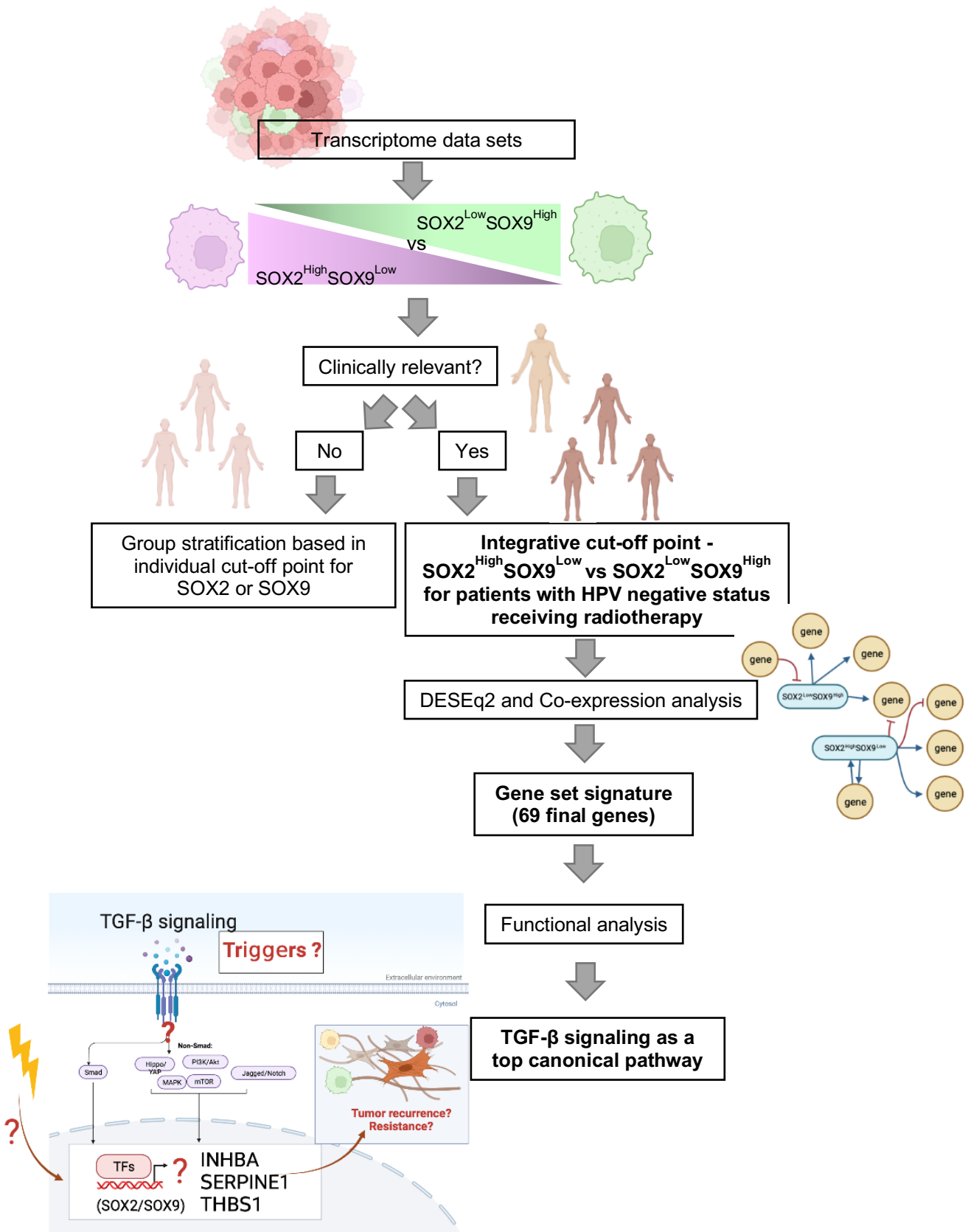


Figure 38. Schematic Diagram of the Role of Inverse SOX2 and SOX9 Expression Predicted by *in Silico* Analysis in HNSCC.

## 5.2. The Role of SOX2 and SOX9 Expression in Tumor Tissue and *in vitro*

The epithelial-mesenchymal transition (EMT) is a well-elucidated process in which epithelial cells gain a mesenchymal phenotype and this transformation correlates with a higher tumor progression and resistance to treatment<sup>65,148,149</sup>. In this study, six well-established HNSCC cell lines were analyzed and demonstrated a heterogeneous pattern of SOX2 and SOX9 expression. Interestingly, four of the six cell lines share similar levels of SOX2/SOX9 expression pattern. The difference in SOX2 and SOX9 levels was associated with a change from an epithelial into a mesenchymal phenotype arguing for a possible link between changes in TF gene expression and EMT<sup>65</sup>. Morphological inspections revealed that cells that presented epithelial phenotype share similar levels of both TFs, while cells with inverse expression presented mesenchymal morphology (Supplementary picture 1). In addition, five OSCC cell lines were analyzed of which four shared similar levels of both TFs and an epithelial phenotype, while one cell line presented high levels of SOX9 being correlated with a mesenchymal morphology. Bayo and Jou *et al.* demonstrated that silencing of SOX2 in HNSCC cell lines up-regulated the EMT marker vimentin, promoted a mesenchymal-like phenotype and cellular migration<sup>57</sup>. Here, after SOX9 silencing in identical cell lines it was observed that SOX9 was not sufficient to promote changes in cellular morphology and/or changes in E-cadherin levels.

Domenici *et al.*, have shown in breast cancer that a signaling network loop based on SOX2 and SOX9 regulation leads to activation of WNT signaling<sup>150</sup>. Garros-Regulez *et al.*, illustrated in glioma cells that exist a regulatory loop between SOX2/SOX9 occurs at post-transcriptional levels and might be a relevant molecular node to tumor maintenance and progression<sup>151</sup>. However, the presented data of cellular culture models with either SOX2 or SOX9 silencing does not supported direct SOX9 regulation by SOX2 or vice versa. However, in line with omics analysis a preliminary data indicates effect of SOX9 silencing on SLUG expression in HNSCC. Lin *et al.*, also showed in lung cancer cells that SOX9 silencing attenuates the SLUG expression resulting in impaired cellular migration and invasiveness<sup>65</sup>. Although, as a limitation of the study the silenced protocols were performed in one cellular model. Generation of additional HNSCC cell lines with SOX9 silencing will enable the characterization of additional alterations in the mutational landscape of HNSCC that modulate SOX9 function in this process.

In tumor samples of the HIPO-HNC cohort IHC staining demonstrated a spatial difference in SOX2 and SOX9 expression patterns with a positive correlation at the invasive front and an inverse correlation at the tumor core. The "go or grow" hypothesis postulates that cancer cells either migrate or proliferate<sup>152,153</sup>. Hence, these data suggest that SOX2 may contribute to tumor growth, while SOX9 would be involved in tumor cell migration and thereby modulates temporal and spatial differences of TFs expression during tumor development. Interesting to note that cancer cells at the invasive front presented a positive staining for both TFs, though the staining intensity differed from individual cancer cell. This was particularly evident for cancer cells at the border of the invasive areas, which presented higher expression for SOX9 but lower SOX2 expression. Baumeister *et al.* showed a low SOX2 expression and gain in vimentin expression in cancer cells at the invasive front<sup>58</sup>. Haga *et al.* demonstrated that CAFs at the invasion front of OSCC up-regulate SOX9 expression through secretion of TGF- $\beta$ 1 which promotes migration and invasiveness of OSCC cells<sup>154</sup>. Hence, high SOX9 expression could serve as surrogate marker for a TGF- $\beta$ -enriched TME, which might also explain the presence of SOX9-positive stromal cells. Riemenschnitter *et al.*, demonstrated that high SOX9 expression in tumor stromal cells correlates with lower overall survival in breast cancer patients treated with neoadjuvant therapy<sup>155</sup>. The tumor microenvironment is composed of CAFs, blood and lymph vessels, immune cells, growth factors, cytokines and they are in constant interplay with cancer cells<sup>18,23,75,143</sup>. The role of SOX9 expression in diverse stroma cells of the TME is not explored yet and these findings provide new insights into potential role of SOX9 in the mutual crosstalk of cancer cells with the microenvironment.

Numerous studies have supported the assumption that SOX2 expression is associated with a tumor-initiation capacity<sup>52,149,156</sup>. However, Sharma *et al.*, have demonstrated that in OSCC cells resistance to cisplatin does not depend on SOX2 levels or its tumor-initiation capacity, suggesting the involvement of another stem cell factor promoting the proliferation and growth of tumor cells<sup>66</sup>. Indeed, in a 3D Matrigel assay SOX2-silencing in a cancer cell line with concomitant SOX9 expression present a higher tumor-initiating capacity compared with control cells. These colonies presented a reduced size compared to controls which might be due to the altered migratory behavior of the SOX2-silenced cells. Interestingly, Garcia *et al.* reported that in colorectal adenocarcinoma cell lines, higher SOX9 expression was associated with CSC properties and metastasis. The cell lines with high SOX9 expression exhibited a



higher number of tumor sphere formation as compared to controls<sup>61</sup>. A more careful inspection in adequate pre-clinical models is warranted in order to address the question, whether SOX2 or SOX9 lead to the tumor initiation capacity in HNSCC or if the presence of both is necessary to promote this phenotype. Several studies support a coregulation of both transcription factors in a context dependent manner. In breast cancer cells resistance to tamoxifen is increased by upregulation of SOX2 which in turn induces upregulation of SOX9 expression<sup>150,157</sup>. Domenici *et al.*, reported that SOX2 activates WNT signaling and this process requires SOX9 expression. Moreover, SOX9 is crucial for maintaining CSCs in breast tumor tissue and resistance to tamoxifen<sup>150</sup>. A limitation of the study, the 3D Matrigel assay was not performed on the HNO 223-shSOX9 cell lines at time of this doctoral thesis. As a speculative result, it is expected a higher number of HNO223-shSOX9 colonies compared to control, and with a large size compared to HNO223-shSOX2 cell line.

Altogether, the results presented in this study demonstrate a spatial difference in SOX2 and SOX9 expression for primary HNSCC. While prominent SOX2 expression was only detected in cancer cells, SOX9 expression was also detected in stromal cells of the TME. These data not only suggest a possible role of the TME in the regulation of an inverse SOX2-SOX9 expression pattern in cancer cells, but also indicate the presence of a SOX9-related gene-regulatory network in stromal cells, which modulates characteristic trails of the TME. This spatial difference and the difference in cellular expression might explain the difficulty in elucidating the contribution of both TFs during metastasis and resistance to treatment in HNSCC, but also other solid tumors. Hence, more experimental data from adequate pre-clinical models are urgently needed to decipher in more detail the involved regulatory pathway and key players.

### **5.3. Radiosensitivity in HNSCC**

Radiotherapy is the first line of treatment in many tumor entities, including HNSCC and can be combined with other modalities to improve the clinical outcome of cancer patients<sup>100</sup>. A growing body of evidence supports the assumption that acquisition of CSC properties is associated with a radioresistant phenotype<sup>52,158-160</sup>. Transcriptional regulators such as Nanog, Oct4 and ALDH have been related to radioresistance and SOX2 expression in cancer cells<sup>52,160-163</sup>. However, bioinformatic analysis of this study demonstrated that HPV-negative HNSCC with a

SOX2<sup>Low</sup>SOX9<sup>High</sup> expression pattern present a worse survival probability under radiotherapy. Chung *et al.*, suggested that SOX2 enhances the effects of irradiation in HNSCC cell lines, and improves the prognosis of patients which might benefit from radiotherapy<sup>164</sup>. This assumption is further supported by experimental data of this study derived from *in vitro* cell cultures. First, the survival of OSCC cells upon irradiation was associated with relatively low SOX2 but high SOX9 expression. Second, the acquisition of a radioresistant phenotype was evident in SOX2 silenced OSCC cell line. Finally, fractionated irradiation revealed a further reduction in SOX2 levels, while SOX9 protein expression was increased. These data indicate a protective role of high SOX9 expression in cancer cells during radiotherapy, which is supported by Roche *et al.*<sup>101</sup>. They demonstrated in a mouse model an accelerated proliferation but reduced radioresistance of intestinal stem cells with a SOX9 knockout as compared to controls. Sharma *et al.*, proposed that stressor factors could increase SOX9 expression causing treatment resistance<sup>66</sup>. Other studies have reported that SOX9 acts as key factor for CSC, EMT and drug resistance<sup>165-168</sup>. It is also worth noting that Vlashi *et al.*, reported an induction of dedifferentiation in HPV-negative HNSCC cell lines by irradiation<sup>160</sup>, which could resemble the mesenchymal-like phenotype upon SOX2 silencing<sup>57</sup>. It is likely that the impact of SOX2 and SOX9 on the response of cancer cells to irradiation is highly context dependent and critically depends on other epigenetic and genetic alterations. Accordingly, a more comprehensive analysis in future studies, including a larger set of cancer cell lines and adequate *in vitro* as well as *in vivo* pre-clinical models is needed to unravel the complexity and context dependent function of both TFs during irradiation.

The tumor microenvironment plays an important role in cancer cell plasticity and regulation of the CSC state, and soluble factors secreted by stromal cells have the potential to dedifferentiate cancer cells into a CSC phenotype<sup>100</sup>. Presented data of this study show an irradiation-induced upregulation of SOX9 and INHBA, a soluble factor related to TGF- $\beta$  signaling. De Martino *et al.*, demonstrated that INHBA contributes in immune evasion of irradiated cancer cells by enhancing regulatory T cells infiltration into tumors after radiotherapy<sup>169</sup>. These data are in line with other studies demonstrating an increase in INHBA expression upon irradiation<sup>120,169,170</sup>. Hence, candidate genes of newly identified gene set signature related to inverse SOX2/SOX9 expression has not only innovative potential to identify cancer cell intrinsic modes of radioresistance, but could also predicted potential mechanisms of immune

downstream targets or upstream regulators of SOX9. As an example, Pal *et al.*, demonstrated that THBS1 is a tumor-specific ECM protein, which is expressed by OSCC cells. THBS1 is stimulated by TGF- $\beta$  signaling and initiates a response cascade that promotes tumor cell migration and activates the integrin signaling network<sup>177</sup>. As mentioned above, INHBA is another candidate that increases cancer cell proliferation and invasion, and its expression serves as an unfavorable risk factor for worse clinical outcome<sup>120,178</sup>. However, the tumor-related function of INHBA is context-dependent and its downregulation in epithelial cancer cells has been linked to tumor progression<sup>144</sup>. A recent study demonstrated that INHBA is overexpressed in tumor samples as compared to normal oral mucosa. This overexpression is negatively correlated with miR145 expression, leads to an invasive behavior of OSCC cancer cells and is associated with a poor prognosis<sup>179</sup>. Yu *et al.*, proposed that miR145 targets the SOX9/ADAM17 axis suppressing tumor initiation, growth and metastasis. They also observed induced secretion of IL-6 in the presence of high SOX9/ADAM17 activity, which results in EMT and cancer stemness<sup>180</sup>. Altogether, the presented data suggested one associated between SOX2/SOX9 and gene regulation which can be promoted by TME, and also present potential for target therapy.

### **5.5. SOX2 and SOX9 Expression and Immune Modulatory Relation**

In the past years, the first-line treatment of recurrent or metastatic HNSCC has been improved with the approval of immunotherapies<sup>69,79,181</sup>. However, many HNSCC has been described as cold tumors, where tumor and stromal cells promote an immunosuppressive microenvironment. The immunosuppressive microenvironment enable escape from immune recognition and impairs activation of cytotoxic immune cells<sup>129</sup>. Quin *et al.*, demonstrated the presence of Treg cells can stimulate an immunosuppressive response by inhibitory cytokines, such as IL-10<sup>79</sup>. As discussed previously, De Martino *et al.*, suggested the secretion of INHBA by irradiated tumor cells increases Tregs infiltration and promotes immune evasion. The analysis of dissociated single cells from patient samples in this study revealed a higher percentage of immune cells in tumor samples as compared to matched samples from adjacent areas. However, only two general markers of immune cells were applied, a more detailed characterization of the quantity and quality of distinct immune cell populations and a larger sample size is required to draw final conclusion, in particular in the context of the inverse SOX2/SOX9 expression.

Activation of the cytotoxic immune response depends on the antigen presentation by HLA molecules<sup>128</sup>. Transcriptome data of a previous study indicated that SOX2 silencing results in downregulation of several HLAs and might affect the antigen presentation machinery in HNO223 cells. On the other hand, primary tumors with a SOX2<sup>Low</sup>SOX9<sup>High</sup> expression pattern of TCGA-HNSCC presented higher scores for antigen presentation and processing. Conflicting data could be explained by differences in SOX9 levels or the potential impact of the TME. To further address a potential role of SOX2 expression in the escape of cancer cells from cytotoxic immune cells several *in vitro* experiments with a focus on the HLA-E/NK axis were conducted in this study. Recent reports demonstrated that blocking NKG2A increases CD8 T cells and NK cells activity<sup>182-184</sup>. NKG2A is an inhibitory receptor, which blocks the effect of T cells and NK cells upon interaction with the non-classical HLA-E. However, presented data of this study did not confirm the SOX2-dependent regulation of HLA-E at protein levels under normal growth condition or irradiation *in vitro*, and more adequate models are needed to further address this issue in future experiments.

Other components of the tumor mass comprise cytokines, growth factors, and extracellular matrix (ECM), which modulate the communication in the TME. Therefore, patient material was analyzed with a cytokine assay to explore differences in inflammatory markers, such as IL-1 and IL-6, in matched samples from tumor tissue and adjacent areas. The release of these particular cytokines in the TME has been related to impaired maturation of antigen-presenting cells (APCs), which promotes the growth of tumor cells<sup>79,185</sup>. Preliminary data of this study indicates a possible association between the high SOX9 expression in stromal cells with a pro-inflammatory cytokine profile in the adjacent area, but again a larger sample size is needed to draw final conclusions.

Radiotherapy can modulate the tumor immune microenvironment to facilitate the recruitment and activation of the immune system<sup>105</sup>. However, radiotherapy also increases TGF- $\beta$  expression, which regulates the balance between and cytotoxic T cell response versus a Treg phenotype<sup>105</sup>. Knitz *et al.*, demonstrated that reprogramming Treg cells in association with irradiation improves the response in radioresistance tumors<sup>181</sup>.

Taken together, preliminary results presented in this work explore the potential relation between SOX2 and SOX9 and the immune system. Limitations, such as sample size, in vitro set-ups, and relation with the TME made it difficult to draw a final conclusion of a direct relation between SOX expression and how the tumor cells can promote immune escape. Further research needs to be conducted to dissect the action mechanisms of this pathway and its relation to radiation treatment.

## 6. Conclusion and Perspectives

Established therapies for HNSCC are either rather toxic or present a low curative response rate. Accordingly, more efficient and less toxic therapies represent an urgent medical need to improve patient survival and their quality of life. Reliable prognostic and predictive biomarkers with the potential to identify patients at a higher risk for tumor relapse will be extremely beneficial for treatment decisions in order to reach this ambiguous task. Presented data of this study confirm that SOX2 and SOX9 act in a dose-dependent manner, indicating a more complex scenario of a gene regulatory network, which contributes to the pathogenesis and treatment response of HNSCC. These findings and newly established models provide a valuable background for future experimental studies and clinical trials to further explore molecular principles of the inverse SOX2/SOX9 regulation in HNSCC, but also beyond in other solid cancers. Expected deliverables will not only improve our knowledge on basic principles in the modulation of the tumor immune microenvironment, cancer cell plasticity and metastasis, but in a long-term perspective could guide the therapy decision-making process at the clinics. In this context, ongoing clinical trials are evaluating the role of TGF- $\beta$  signaling inhibition in combination with immune checkpoint blockade (NCT04247282, NCT04220775), or assess the efficacy of Activin A on patient quality of life (NCT03162042). It is intriguing to consider that the combination of TGF- $\beta$  and INHBA inhibition could improve the response to treatment, especially under radiotherapy.

This work sheds light on the potential association between SOX2/SOX9-related gene regulation and TGF- $\beta$  signaling, and suggests a mutual interaction of the TME with SOX2 and/or SOX9 expressing cancer cells, promoting treatment resistance. These findings need further investigated in future experimental studies with adequate pre-clinical models mimicking the complex nature of human cancers, and more comprehensive clinical studies with larger sample size. Worth noting that the interaction of transcription factors and cancer immune response under irradiation treatment is not well-explored in the literature. Hence, dissecting the molecular mechanism controlling the functional cross-talk between TGF- $\beta$  signaling, SOX-related gene-regulatory networks, and immune evasion could be an interesting focus for new strategies of patient stratification and establishing less toxic therapies.

## 7. Summary

Head and neck squamous cell carcinoma (HNSCC) is a highly heterogeneous disease with a rather diverse patient outcome. An epigenetic switch between SOX2 and SOX9 has been associated with modulation of cancer cells plasticity. Moreover, numerous studies support an association of altered expression for both transcription factors with the capacity to evade the immune surveillance, to develop treatment resistance and to promote metastasis. However, the molecular mechanisms involved in SOX2/SOX9 regulation and their clinical impact are still largely unknown. Main hypothesis of this study was that the stem cell-related transcription factors SOX2 and SOX9 modulate a gene-regulatory network and thereby influence the treatment response of cancer cells. The aims were to elucidate the role of the SOX2/SOX9 relationship and its heterogeneous expression in tumors, and how this potentially influences the tumor microenvironment including radiotherapy, which finally impacts patient response to therapy. SOX2 and SOX9 expression was investigated in established cancer cell lines, which showed a heterogeneous expression pattern for both transcriptional factors. Immunohistochemical staining confirmed a heterogeneous and inverse SOX2/SOX9 expression pattern in most tumor tissues of patients from the Heidelberg Center for Personalized Oncology (HIPO) HNC cohort. Bioinformatics analysis of transcriptome data from The Cancer Genome Atlas for Head and Neck Squamous Cell Carcinoma (TCGA-HNSCC) unraveled candidate genes that might regulate or are the consequence of inverse SOX2/SOX9 expression. Ingenuity Pathway Analysis for selected candidate genes predicted TGF- $\beta$  signaling as one of the top canonical pathways related to inverse SOX2/SOX9 expression. The analysis of clinical data from TCGA-HNSCC confirmed that HPV-negative HNSCC with a high SOX2 and low SOX9 expression pattern exhibit an improved overall survival, when radiotherapy is part of the treatment. In line with the clinical data, SOX2 silencing in the OSCC cell line HNO223 reduced their survival capabilities after a single dose as well as fractionated dose of irradiation. A more detailed molecular characterization of established HNSCC cell lines revealed a trend toward reduced SOX2 expression but increased SOX9 protein levels for irradiated cancer cells as compared to controls. An increase of INHBA expression was also evident after irradiation supporting the predicted model of the bioinformatic analysis (positive association between TGF- $\beta$  signaling and a SOX2<sup>low</sup>SOX9<sup>high</sup> phenotype).

To investigate whether SOX2 modulates the susceptibility of cancer cells to immune cell attack after irradiation, co-culture assays were established with peripheral blood mononuclear cells that demonstrated the involvement of a more complex network in the regulation of immune effectors. To investigate the putative relation of SOX2/SOX9 expression in tumor cells with the tumor immune milieu in patient samples, single cells dissociated from primary tumors were investigated for marker expression of pan-leukocytes and Natural Killer cells. The analysis revealed an increased occurrence of immune cells in the tumor sample as compared to matched non-malignant adjacent tissue. Immunohistochemistry staining for SOX2 and SOX9 confirmed their heterogenous expression pattern also in those matched pathological samples. Further investigation on the relation between SOX2/SOX9 and the immune milieu for patient tumor material need be performed in order to correlate above findings with the treatment outcome. The recent findings shown that SOX2 and SOX9 predicted a signature gene set which can modulate radioresistance and the TGF- $\beta$  signaling is one of the top canonical pathways acting in the overexpression of candidate genes. These results pave the way for a further investigation where these transcription factors could be used as predictive outcome of head and neck squamous cell carcinoma patients.



## Zusammenfassung

Das Plattenepithelkarzinom des Kopfes und Halses (HNSCC) ist eine sehr heterogene Erkrankung mit einem sehr unterschiedlichen Ausgang für die Patienten. Epigenetische Regulationen der Transkriptionsfaktoren SOX2 und SOX9 wurden mit der Modulierung der Plastizität von Krebszellen in Verbindung gebracht. Darüber hinaus belegen zahlreiche Studien einen Zusammenhang zwischen einer veränderten Expression beider Transkriptionsfaktoren und der Fähigkeit, sich der Immunüberwachung zu entziehen, Behandlungsresistenz zu entwickeln und die Metastasierung zu fördern. Die molekularen Mechanismen, die an der Regulierung von SOX2/SOX9 beteiligt sind und ihre klinischen Auswirkungen sind jedoch noch weitgehend unbekannt. Die Haupthypothese dieser Studie war, dass die stammzellverwandten Transkriptionsfaktoren SOX2 und SOX9 ein genregulatorisches Netzwerk modulieren und dadurch die Behandlungsreaktion von Krebszellen beeinflussen. Ziel war es, die Rolle der SOX2/SOX9-Beziehung und ihrer heterogenen Expression in Tumoren aufzuklären und herauszufinden, wie dies möglicherweise die Mikroumgebung des Tumors einschließlich der Strahlentherapie beeinflusst, was sich letztlich auf das Ansprechen der Patienten auf die Therapie auswirkt. Die Expression von SOX2 und SOX9 wurde in etablierten Krebszelllinien untersucht, die ein heterogenes Expressionsmuster für beide Transkriptionsfaktoren aufwiesen. Immunhistochemische Färbungen bestätigten ein heterogenes und inverses SOX2/SOX9-Expressionsmuster in den meisten Tumorgeweben von Patienten aus der HNC-Kohorte des Heidelberger Zentrums für Personalisierte Onkologie (HIPO). Eine bioinformatische Analyse von Transkriptomdaten aus dem The Cancer Genome Atlas for Head and Neck Squamous Cell Carcinoma (TCGA-HNSCC) brachte Kandidatengene ans Licht, die die inverse SOX2/SOX9-Expression regulieren oder deren Folge sein könnten. Die Ingenuity Pathway Analysis für ausgewählte Kandidatengene ergab, dass die TGF- $\beta$ -Signalübertragung einer der wichtigsten kanonischen Signalwege ist, die mit der inversen SOX2/SOX9-Expression zusammenhängen. Die Analyse der klinischen Daten von TCGA-HNSCC bestätigte, dass HPV-negative HNSCC mit einem hohen SOX2- und einem niedrigen SOX9-Expressionsmuster ein verbessertes Gesamtüberleben aufweisen, wenn eine Strahlentherapie Teil der Behandlung ist. In Übereinstimmung mit den klinischen Daten verringerte die Ausschaltung von SOX2 in der OSCC-Zelllinie HNO223 deren Überlebensfähigkeit sowohl nach einer einmaligen als auch nach fraktionierter

Bestrahlung. Eine detailliertere molekulare Charakterisierung etablierter HNSCC-Zelllinien zeigte einen Trend zu einer reduzierten SOX2-Expression, aber erhöhte SOX9-Proteinspiegel bei bestrahlten Krebszellen im Vergleich zu den Kontrollen. Eine Zunahme der INHBA-Expression war nach der Bestrahlung ebenfalls zu beobachten, was das von der bioinformatischen Analyse vorhergesagte Modell unterstützt (positive Assoziation zwischen TGF-Signalen und einem SOX2-niedrigen/SOX9-hohen Phänotyp).

Um zu untersuchen, ob SOX2 die Anfälligkeit von Krebszellen für den Angriff von Immunzellen nach der Bestrahlung moduliert, wurden Co-Kulturtests mit mononukleären Zellen aus peripherem Blut durchgeführt, die die Beteiligung eines komplexeren Netzwerks an der Regulierung von Immuneffektoren zeigten. Um den mutmaßlichen Zusammenhang zwischen der SOX2/SOX9-Expression in Tumorzellen und dem Tumorummunmilieu in Patientenproben zu untersuchen, wurden einzelne, aus Primärtumoren dissoziierte Zellen auf die Expression von Markern von Pan-Leukozyten und natürlichen Killerzellen untersucht. Die Analyse ergab ein erhöhtes Vorkommen von Immunzellen in der Tumorprobe im Vergleich zu abgestimmtem nicht-malignem Nachbargewebe. Die immunhistochemische Färbung für SOX2 und SOX9 bestätigte deren heterogenes Expressionsmuster auch in diesen angepassten pathologischen Proben. Weitere Untersuchungen über den Zusammenhang zwischen SOX2/SOX9 und dem Immunmilieu des Tumormaterials von Patienten müssen durchgeführt werden, um die oben genannten Ergebnisse mit dem Behandlungsergebnis in Beziehung zu setzen. Die jüngsten Ergebnisse zeigen, dass SOX2 und SOX9 eine Reihe von Signaturgenen vorhersagen, die die Radioresistenz modulieren können und dass der TGF- $\beta$ -Signalweg einer der wichtigsten kanonischen Signalwege ist, die bei der Überexpression der Kandidatengene eine Rolle spielen. Diese Ergebnisse ebnen den Weg für weitere Untersuchungen, bei denen diese Transkriptionsfaktoren zur Vorhersage von Ergebnissen bei Patienten mit Plattenepithelkarzinomen des Kopfes und Halses verwendet werden könnten.

## 8. Reference

1. Ferlay, J. et al. Cancer statistics for the year 2020: An overview. *International journal of cancer* (2021).
2. American Society of Clinical Oncology (ASCO). Head and Neck Cancer - Statistics. (2022) <<https://www.cancer.net/cancer-types/head-and-neck-cancer/statistics>>. Accessed 2022 20.05.
3. Johnson, DE. et al. Head and neck squamous cell carcinoma. *Nature reviews. Disease primers* 6(2020).
4. Leemans, C.R., et al. The molecular biology of head and neck cancer. *Nat Rev Cancer* 11, 9-22 (2011).
5. Cogliano, V.J. et al. Preventable Exposures Associated With Human Cancers. *JNCI: Journal of the National Cancer Institute* 103, 1827-1839 (2022).
6. Bhatia, A. & Burtneess, B. Human Papillomavirus-Associated Oropharyngeal Cancer: Defining Risk Groups and Clinical Trials. *J Clin Oncol* 33, 3243-50 (2015).
7. Leemans, C.R., Snijders, P.J.F. & Brakenhoff, R.H. The molecular landscape of head and neck cancer. *Nat Rev Cancer* 18, 269-282 (2018).
8. Shamseddine, AA et al. Tumor Immunity and Immunotherapy for HPV-Related Cancers. *Cancer discovery* 11(2021).
9. Seliger, B. et al. Immune Escape Mechanisms and Their Clinical Relevance in Head and Neck Squamous Cell Carcinoma. *International journal of molecular sciences* 21(2020).
10. Aggarwal, N. et al. Human Papillomavirus Infection in Head and Neck Squamous Cell Carcinomas: Transcriptional Triggers and Changed Disease Patterns. *Frontiers in cellular and infection microbiology* 10(2020).
11. Kostareli, E. et al. HPV-related methylation signature predicts survival in oropharyngeal squamous cell carcinomas. *J Clin Invest* 123, 2488-501 (2013).
12. Gillison, M.L., et al. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 33, 3235-42 (2015).
13. Stransky, N. et al. The mutational landscape of head and neck squamous cell carcinoma. *Science (New York, N.Y.)* 333(2011).
14. Pai, S.I. & Westra, W.H. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu Rev Pathol* 4, 49-70 (2009).
15. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer discovery* 12(2022).
16. Hanahan, D & Weinberg, RA. Hallmarks of cancer: the next generation. *Cell* 144(2011).
17. Zack, T.I. et al. Pan-cancer patterns of somatic copy number alteration. *Nat Genet* 45, 1134-40 (2013).
18. Koontongkaew, S. The tumor microenvironment contribution to development, growth, invasion and metastasis of head and neck squamous cell carcinomas. *J Cancer* 4, 66-83 (2013).
19. Athwal, VS. et al. SOX9 regulated matrix proteins are increased in patients serum and correlate with severity of liver fibrosis. *Scientific reports* 8(2018).
20. Dongre, A. & Weinberg, R.A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nature Reviews Molecular Cell Biology* 20, 69-84 (2018).

21. Tsai, J.H. & Yang, J. Epithelial–mesenchymal plasticity in carcinoma metastasis. *Genes Dev* 27, 2192-206 (2013).
22. Nieto, M.A., Huang, R.Y., Jackson, R.A. & Thiery, J.P. EMT: 2016. *Cell* 166, 21-45 (2016).
23. Puram, S.V. et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 171, 1611-1624.e24 (2017).
24. Morel, AP. et al. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS one* 3(2008).
25. Mani, SA. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133(2008).
26. Ferlay J, E.M. et al. Global Cancer Observatory: Cancer Today. Vol. 2022 (Lyon, France: International Agency for Research on Cancer., 2020).
27. Amin, MB. et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA: a cancer journal for clinicians* 67(2017).
28. Latchman, DS. Transcription factors: an overview. *The international journal of biochemistry & cell biology* 29(1997).
29. Simicevic, J. & Deplancke, B. Transcription factor proteomics-Tools, applications, and challenges. *Proteomics* 17(2017).
30. Papavassiliou, K.A. & Papavassiliou, A.G. Transcription Factor Drug Targets. *J Cell Biochem* 117, 2693-2696 (2016).
31. Zhou, T et al. Quantitative modeling of transcription factor binding specificities using DNA shape. *Proceedings of the National Academy of Sciences of the United States of America* 112(2015).
32. Spitz, F & Furlong, EE. Transcription factors: from enhancer binding to developmental control. *Nature reviews. Genetics* 13(2012).
33. Kadonaga, JT. Studying transcription factor function in the genome at molecular resolution. *Trends in genetics : TIG* 37(2021).
34. Kadonaga JT. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell* 116(2004).
35. Butler JE & Kadonaga JT . The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & development* 16(2002).
36. Pennacchio, LA. et al. Enhancers: five essential questions. *Nature reviews. Genetics* 14(2013).
37. Wittkopp PJ & Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature reviews. Genetics* 13(2011).
38. Mobley, A.S. Induced Pluripotent Stem Cells. in *Neural Stem Cells and Adult Neurogenesis* (ed. Press, A.) 67-94 (2019).
39. Lambert, SA. et al. The Human Transcription Factors. *Cell* 172(2018).
40. Wegner, M. All purpose Sox: The many roles of Sox proteins in gene expression. *Int J Biochem Cell Biol* 42, 381-90 (2010).
41. Angelozzi, M & Lefebvre, V. SOXopathies: Growing Family of Developmental Disorders Due to SOX Mutations. *Trends in genetics : TIG* 35(2019).

42. Grimm, D. et al. The role of SOX family members in solid tumours and metastasis. *Seminars in cancer biology* 67(2020).
43. Gasca, S. et al. A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of SOX9 during sexual determination. *Proceedings of the National Academy of Sciences of the United States of America* 99(2002).
44. Südbeck, P & Scherer, G. Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9. *The Journal of biological chemistry* 272(1997).
45. Castillo, S.D. & Sanchez-Cespedes, M. The SOX family of genes in cancer development: biological relevance and opportunities for therapy. <http://dx.doi.org/10.1517/14728222.2012.709239> (2012).
46. Kamachi, Y & Kondoh, H. Sox proteins: regulators of cell fate specification and differentiation. *Development (Cambridge, England)* 140(2013).
47. Dehshahri, A. et al. Editing SOX Genes by CRISPR-Cas: Current Insights and Future Perspectives. *International Journal of Molecular Sciences* 22, 11321 (2021).
48. Angelozzi, M. & Lefebvre, V. SOXopathies: Growing Family of Developmental Disorders Due to SOX Mutations. *Trends Genet* 35, 658-671 (2019).
49. Stevanovic, M. et al. The cDNA sequence and chromosomal location of the human SOX2 gene. *Mammalian genome : official journal of the International Mammalian Genome Society* 5(1994).
50. Collignon, J. et al. A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development (Cambridge, England)* 122(1996).
51. Zhang, S. et al. Functional characterization of SOX2 as an anticancer target. *Signal Transduction and Targeted Therapy* 5, 1-17 (2020).
52. Mamun, M.A. et al. SOX2 in Cancer Stemness: Tumor Malignancy and Therapeutic Potentials. *J Mol Cell Biol* (2018).
53. Schröck, A. et al. Sex Determining Region Y-Box 2 (SOX2) Amplification Is an Independent Indicator of Disease Recurrence in Sinonasal Cancer. in *PLoS One*, Vol. 8 (2013).
54. Sarkar, A. et al. Sox2 Suppresses Gastric Tumorigenesis in Mice. *Cell reports* 16(2016).
55. Rodriguez, D. et al. The Central Contributions of Breast Cancer Stem Cells in Developing Resistance to Endocrine Therapy in Estrogen Receptor (ER)-Positive Breast Cancer. *Cancers (Basel)* 11(2019).
56. Kuo, M.H. et al. Crosstalk between SOX2 and TGF- $\beta$  signaling regulates EGFR-TKI tolerance and lung cancer dissemination. (2020).
57. Bayo, P & Jou, A. et al. Loss of SOX2 expression induces cell motility via vimentin up-regulation and is an unfavorable risk factor for survival of head and neck squamous cell carcinoma. *Mol Oncol* 9, 1704-19 (2015).
58. Baumeister, P. et al. High Expression of EpCAM and Sox2 is a Positive Prognosticator of Clinical Outcome for Head and Neck Carcinoma. *Scientific reports* 8(2018).
59. Symon, A & Harley, V SOX9: A genomic view of tissue specific expression and action. *The international journal of biochemistry & cell biology* 87(2017).
60. M, P., SK, T. & BK, B. SOX9: An emerging driving factor from cancer progression to drug resistance. *Biochimica et biophysica acta. Reviews on cancer* 1875(2021).

61. Aguilar-Medina, M. et al. SOX9 Stem-Cell Factor: Clinical and Functional Relevance in Cancer. *J Oncol* 2019, 6754040 (2019).
62. Castillo, SD. et al The SOX family of genes in cancer development: biological relevance and opportunities for therapy. *Expert opinion on therapeutic targets* 16(2012).
63. Huang, JQ. et al. SOX9 drives the epithelial-mesenchymal transition in non-small-cell lung cancer through the Wnt/ $\beta$ -catenin pathway. *Journal of translational medicine* 17(2019).
64. Li, T. et al. TGF- $\beta$ 1-SOX9 axis-inducible COL10A1 promotes invasion and metastasis in gastric cancer via epithelial-to-mesenchymal transition. *Cell death & disease* 9(2018).
65. Lin, S.C. et al. Epigenetic Switch between SOX2 and SOX9 Regulates Cancer Cell Plasticity. *Cancer Res* 76, 7036-7048 (2016).
66. Sharma, A. et al. Longitudinal single-cell RNA sequencing of patient-derived primary cells reveals drug-induced infidelity in stem cell hierarchy. *Nat Commun* 9, 4931 (2018).
67. Malladi, S. et al. Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* 165, 45-60 (2016).
68. Laughney, A.M. et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nature Medicine* 26, 259-269 (2020).
69. Bhat, A.A. et al. Tumor microenvironment: an evil nexus promoting aggressive head and neck squamous cell carcinoma and avenue for targeted therapy. *Signal Transduction and Targeted Therapy* 6, 1-15 (2021).
70. Wu, T. & Dai, Y. Tumor microenvironment and therapeutic response. *Cancer letters* 387(2017).
71. Wondergem, N.E. et al. The Immune Microenvironment in Head and Neck Squamous Cell Carcinoma: on Subsets and Subsites. *Current Oncology Reports* 22, 1-14 (2020).
72. Lechien, J.R. et al. HPV Involvement in the Tumor Microenvironment and Immune Treatment in Head and Neck Squamous Cell Carcinomas. *Cancers* 12, 1060 (2020).
73. Mito, I. et al. Comprehensive analysis of immune cell enrichment in the tumor microenvironment of head and neck squamous cell carcinoma. *Scientific Reports* 11, 1-9 (2021).
74. Abd Hamid, M. et al. Enriched HLA-E and CD94/NKG2A Interaction Limits Antitumor CD8 + Tumor-Infiltrating T Lymphocyte Responses. *Cancer immunology research* 7(2019).
75. Markwell, S.M. & Weed, S.A. Tumor and Stromal-Based Contributions to Head and Neck Squamous Cell Carcinoma Invasion. in *Cancers (Basel)*, Vol. 7 382-406 (2015).
76. Derynck, R., Turley, S.J. & Akhurst, R.J. TGF $\beta$  biology in cancer progression and immunotherapy. *Nature Reviews Clinical Oncology*, 1-26 (2020).
77. Harper J & Sainson RC. Regulation of the anti-tumour immune response by cancer-associated fibroblasts. *Seminars in cancer biology* 25(2014).
78. Wang, G. et al. Tumor microenvironment in head and neck squamous cell carcinoma: Functions and regulatory mechanisms. *Cancer letters* 507(2021).
79. Qin, Y. et al Tumor microenvironment and immune-related therapies of head and neck squamous cell carcinoma. *Molecular therapy oncolytics* 20(2021).
80. Elkashty, OA. et al Head and neck cancer management and cancer stem cells implication. *The Saudi dental journal* 31(2019).

81. Linge, A. et al. Independent validation of tumour volume, cancer stem cell markers and hypoxia-associated gene expressions for HNSCC after primary radiochemotherapy. *Clinical and translational radiation oncology* 16(2019).
82. Wiechec, E. et al. Hypoxia Mediates Differential Response to Anti-EGFR Therapy in HNSCC Cells. *International journal of molecular sciences* 18(2017).
83. Göttgens, EL. et al. HPV, hypoxia and radiation response in head and neck cancer. *The British journal of radiology* 92(2019).
84. Sepich-Poore, GD. et al. The microbiome and human cancer. *Science (New York, N.Y.)* 371(2021).
85. Giraldo, NA. et al. The clinical role of the TME in solid cancer. *British journal of cancer* 120(2019).
86. Kirthi Koushik, AS. et al. Principles of radiation oncology: a beams eye view for a surgeon. *Indian journal of surgical oncology* 4(2013).
87. Baskar, R. et al. Cancer and radiation therapy: current advances and future directions. *International journal of medical sciences* 9(2012).
88. Halperin, E. et al. *Perez and Brady's principles and practice of radiation oncology*, (2013).
89. Gunderson, T. *Clinical Radiation Oncology*, (Elsevier, Philadelphia, 2012).
90. Hutchinson, MND. et al. Radiation resistance in head and neck squamous cell carcinoma: dire need for an appropriate sensitizer. *Oncogene* 39(2020).
91. Gatz, SA. et al. p53 in recombination and repair. *Cell death and differentiation* 13(2006).
92. Helton, ES. et al. p53 modulation of the DNA damage response. *Journal of cellular biochemistry* 100(2007).
93. Wu, CH. et al. Cellular senescence is an important mechanism of tumor regression upon c-Myc inactivation. *Proceedings of the National Academy of Sciences of the United States of America* 104(2007).
94. Pieters, R.S. et al. (eds.). *Principles of Radiation Oncology*, (University of Massachusetts Medical School, Worcester, MA, 2017).
95. Anderson, G. et al. An Updated Review on Head and Neck Cancer Treatment with Radiation Therapy. *Cancers* 13, 4912 (2021).
96. Seiwert, TY & Cohen EE. State-of-the-art management of locally advanced head and neck cancer. *British journal of cancer* 92(2005).
97. von der Grun, J. et al. Targeted Therapies and Immune-Checkpoint Inhibition in Head and Neck Squamous Cell Carcinoma: Where Do We Stand Today and Where to Go? *Cancers (Basel)* 11(2019).
98. Zhou, G & Liu, Z. TP53 Mutations in Head and Neck Squamous Cell Carcinoma and Their Impact on Disease Progression and Treatment Response. *Journal of cellular biochemistry* 117(2016).
99. Ghisolfi, L. et al. Ionizing radiation induces stemness in cancer cells. *PloS one* 7(2012).
100. Olivares-Urbano, MA. et al. CSC Radioresistance: A Therapeutic Challenge to Improve Radiotherapy Effectiveness in Cancer. *Cells* 9(2020).
101. Roche, KC. et al. SOX9 maintains reserve stem cells and preserves radioresistance in mouse small intestine. *Gastroenterology* 149(2015).

102. Yang, L. et al. Inhibition of PI3K/AKT Signaling Pathway Radiosensitizes Pancreatic Cancer Cells with ARID1A Deficiency in Vitro. *Journal of Cancer* 9(2018).
103. Tsai, CN. et al. Activin A regulates the epidermal growth factor receptor promoter by activating the PI3K/SP1 pathway in oral squamous cell carcinoma cells. *Scientific reports* 9(2019).
104. YU, CC. et al. Targeting the PI3K/AKT/mTOR signaling pathway as an effectively radiosensitizing strategy for treating human oral squamous cell carcinoma in vitro and in vivo. *Oncotarget* 8(2017).
105. Sharabi, A.B., Lim, M., DeWeese, T.L. & Drake, C.G. Radiation and checkpoint blockade immunotherapy: radiosensitisation and potential mechanisms of synergy. *Lancet Oncol* 16, e498-509 (2015).
106. Powell, SF. et al. Safety and Efficacy of Pembrolizumab With Chemoradiotherapy in Locally Advanced Head and Neck Squamous Cell Carcinoma: A Phase IB Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 38(2020).
107. Ferris, RL. et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *The New England journal of medicine* 375(2016).
108. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 517(2015).
109. Schmitt, K. et al. Somatic mutations and promotor methylation of the ryanodine receptor 2 is a common event in the pathogenesis of head and neck cancer. *International journal of cancer* 145(2019).
110. Cerami, E et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* 2(2012).
111. Gao, J. et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling* 6(2013).
112. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15, 1-21 (2014).
113. Metsalu, T. et al. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic acids research* 43(2015).
114. Bankhead, P. et al. QuPath: Open source software for digital pathology image analysis. *Scientific reports* 7(2017).
115. Franken, N.A.P., Rodermond, H.M., Stap, J., Haveman, J. & van Bree, C. Clonogenic assay of cells in vitro. *Nature Protocols* 1, 2315-2319 (2006).
116. The Cancer Genome Atlas Program - National Cancer Institute. (2018).
117. Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res* 43, W566-70 (2015).
118. Schmitt, K. et al. Somatic mutations and promotor methylation of the ryanodine receptor 2 is a common event in the pathogenesis of head and neck cancer. *Int J Cancer* (2019).
119. Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* 569(2019).
120. Wu, Z. et al. Expression and gene regulation network of INHBA in Head and neck squamous cell carcinoma based on data mining. *Scientific Reports* 9, 1-11 (2019).



121. Schmidt, S. et al. Development and Validation of a Gene Signature for Patients with Head and Neck Carcinomas Treated by Postoperative Radio(chemo)therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* 24(2018).
122. Luanpitpong, S. et al. SLUG is required for SOX9 stabilization and functions to promote cancer stem cells and metastasis in human lung carcinoma. *Oncogene* 35, 2824-2833 (2015).
123. Guo, W. et al. Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* 148(2012).
124. Halenius, A. et al. Classical and non-classical MHC I molecule manipulation by human cytomegalovirus: so many targets—but how many arrows in the quiver? *Cellular & Molecular Immunology* 12, 139-153 (2014).
125. Kamiya, T. et al. Blocking expression of inhibitory receptor NKG2A overcomes tumor resistance to NK cells. (2019).
126. Bukur, J. et al. The role of classical and non-classical HLA class I antigens in human tumors. *Seminars in cancer biology* 22(2012).
127. Rados, P.V. & Inchausti, A.J. Análise funcional do fator de transcrição SOX2 na patogênese de carcinoma espinocelular de cabeça e pescoço. Tese, (2014). Available><https://lume.ufrgs.br/handle/10183/102520?locale-attribute=en>
128. René, C. et al. Expression of classical HLA class I molecules: regulation and clinical impacts: Julia Bodmer Award Review 2015. *HLA* 87(2016).
129. Farlow, J.L. et al. Immune deserts in head and neck squamous cell carcinoma: A review of challenges and opportunities for modulating the tumor immune microenvironment. *Oral oncology* 120(2021).
130. Bando, N. et al. HLA class I antigen and transporter associated with antigen processing downregulation in metastatic lesions of head and neck squamous cell carcinoma as a marker of poor prognosis. *Oncology reports* 23(2010).
131. Tsuchiya H & Shiota G. Immune evasion by cancer stem cells. *Regenerative therapy* 17(2021).
132. Chen, J.L. et al. Persistently elevated soluble MHC class I polypeptide-related sequence A and transforming growth factor- $\beta$ 1 levels are poor prognostic factors in head and neck squamous cell carcinoma after definitive chemoradiotherapy. *PloS one* 13(2018).
133. Cooper, M.A. et al. Interleukin-1beta costimulates interferon-gamma production by human natural killer cells. *Eur J Immunol* 31, 792-801 (2001).
134. Kartikasari, A.E.R. et al. Tumor-Induced Inflammatory Cytokines and the Emerging Diagnostic Devices for Cancer Detection and Prognosis | *Frontiers in Oncology*. (2022).
135. Wang, Y.-C. et al. Computational Genomics in the Era of Precision Medicine: Applications to Variant Analysis and Gene Therapy. *Journal of Personalized Medicine* 12, 175 (2022).
136. Nakagawa, H. et al. Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer science* 109(2018).
137. Jan Budczisz, F.K. et al. Cutoff Finder: A Comprehensive and Straightforward Web Application Enabling Rapid Biomarker Cutoff Optimization. (2012).
138. Camp, R.L. et al. X-Tile | *Clinical Cancer Research* | American Association for Cancer Research. 7252–7259 (2004).
139. López-Ratón, M. et al. OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests. <https://www.jstatsoft.org/index.php/jss> (2014).

140. Torsten Hothorn, B.L. On the exact distribution of maximally selected rank statistic. *Computational Statistics & Data Analysis* 43, Pages 121-137 (2003).
141. Hanahan, D. & Coussens, Lisa M. Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment. *Cancer Cell* 21, 309-322 (2012).
142. Ikushima H & Miyazono K. TGFbeta signalling: a complex web in cancer progression. *Nature reviews. Cancer* 10(2010).
143. Zhang, S. et al. Tumor-associated macrophages promote tumor metastasis via the TGF- $\beta$ /SOX9 axis in non-small cell lung cancer. *8* (2017).
144. Loomans, HA. et al. Intertwining of Activin A and TGF $\beta$  Signaling: Dual Roles in Cancer Progression and Cancer Cell Invasion. *Cancers* 7(2014).
145. Grady, W.M. Transforming Growth Factor- $\beta$ , Smads, and Cancer | *Clinical Cancer Research* | American Association for Cancer Research. *Clin Cancer Res* 11, 3151-3154 (2005).
146. Loomans, HA. et al. Loss of ACVR1B leads to increased squamous cell carcinoma aggressiveness through alterations in cell-cell and cell-matrix adhesion proteins. *American journal of cancer research* 7(2017).
147. Khorani, K. et al. Establishment of a Plasticity-Associated Risk Model Based on a SOX2- and SOX9-Related Gene Set in Head and Neck Squamous Cell Carcinoma. *Molecular cancer research : MCR* 19(2021).
148. Hay, E.D. An overview of epithelio-mesenchymal transformation. *Acta Anat (Basel)* 154, 8-20 (1995).
149. Chen, C. et al. Evidence for epithelial-mesenchymal transition in cancer stem cells of head and neck squamous cell carcinoma. *PLoS One* 6, e16466 (2011).
150. Domenici, G. et al. A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. *Oncogene* 38(2019).
151. Garros-Regulez, L. et al. mTOR inhibition decreases SOX2-SOX9 mediated glioma stem cell activity and temozolomide resistance. *Expert opinion on therapeutic targets* 20(2016).
152. Garay, T. et al. Cell migration or cytokinesis and proliferation?--revisiting the "go or grow" hypothesis in cancer cells in vitro. *Experimental cell research* 319(2013).
153. Hatzikirou, H. et al. 'Go or grow': the key to the emergence of invasion in tumour progression? *Mathematical medicine and biology : a journal of the IMA* 29(2012).
154. Haga, K. et al. Crosstalk between oral squamous cell carcinoma cells and cancer-associated fibroblasts via the TGF- $\beta$ /SOX9 axis in cancer progression. *Translational oncology* 14(2021).
155. Riemenschnitter, C. et al. Stability and prognostic value of Slug, Sox9 and Sox10 expression in breast cancers treated with neoadjuvant chemotherapy. *SpringerPlus* 2(2013).
156. Boumahdi, S. et al. SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature* 511, 246-50 (2014).
157. Piva, M. et al. Sox2 promotes tamoxifen resistance in breast cancer cells. *EMBO molecular medicine* 6(2014).
158. Kurth, I. et al. Cancer stem cell related markers of radioresistance in head and neck squamous cell carcinoma. *Oncotarget* 6(2015).
159. Keysar, S.B. et al. Regulation of Head and Neck Squamous Cancer Stem Cells by PI3K and SOX2. *J Natl Cancer Inst* 109(2017).

160. Vlashi, E. et al. Radiation-Induced Dedifferentiation of Head and Neck Cancer Cells Into Cancer Stem Cells Depends on Human Papillomavirus Status. *International journal of radiation oncology, biology, physics* 94(2016).
161. Yang, L. et al. Targeting cancer stem cell pathways for cancer therapy. *Signal transduction and targeted therapy* 5(2020).
162. Huang, C. et al. ERK1/2-Nanog signaling pathway enhances CD44(+) cancer stem-like cell phenotypes and epithelial-to-mesenchymal transition in head and neck squamous cell carcinomas. *Cell death & disease* 11(2020).
163. Leong, HS. et al. Targeting cancer stem cell plasticity through modulation of epidermal growth factor and insulin-like growth factor receptor signaling in head and neck squamous cell cancer. *Stem cells translational medicine* 3(2014).
164. Chung, JH. et al. SOX2 activation predicts prognosis in patients with head and neck squamous cell carcinoma. *Scientific reports* 8(2018).
165. Kawai, T. et al. SOX9 is a novel cancer stem cell marker surrogated by osteopontin in human hepatocellular carcinoma. *Scientific reports* 6(2016).
166. Xue, Y. et al. SOX9/FXYD3/Src Axis Is Critical for ER + Breast Cancer Stem Cell Function. *Molecular cancer research : MCR* 17(2019).
167. Feng, C. et al. SOX9/miR-130a/CTR1 axis modulates DDP-resistance of cervical cancer cell. *Cell cycle (Georgetown, Tex.)* 17(2018).
168. Xue, Y. et al. HDAC5-mediated deacetylation and nuclear localisation of SOX9 is critical for tamoxifen resistance in breast cancer. *British journal of cancer* 121(2019).
169. De Martino, M. et al. Activin A Promotes Regulatory T-cell-Mediated Immunosuppression in Irradiated Breast Cancer. *Cancer immunology research* 9(2021).
170. Carl, C. et al. Ionizing radiation induces a motile phenotype in human carcinoma cells in vitro through hyperactivation of the TGF-beta signaling pathway. *Cellular and molecular life sciences : CMLS* 73(2016).
171. Jana, S. et al. SOX9: The master regulator of cell fate in breast cancer. *Biochemical pharmacology* 174(2020).
172. Yu, Y. et al. miR-190 enhances endocrine therapy sensitivity by regulating SOX9 expression in breast cancer. *Journal of experimental & clinical cancer research : CR* 38(2019).
173. Zhao, Y. et al. MicroRNA-511 inhibits malignant behaviors of breast cancer by directly targeting SOX9 and regulating the PI3K/Akt pathway. *International journal of oncology* 53(2018).
174. Seymour, PA. et al. SOX9 is required for maintenance of the pancreatic progenitor cell pool. *Proceedings of the National Academy of Sciences of the United States of America* 104(2007).
175. Hirano, T. et al. TGF- $\beta$ 1 induces N-cadherin expression by upregulating Sox9 expression and promoting its nuclear translocation in human oral squamous cell carcinoma cells. *Oncology letters* 20(2020).
176. Sumita, Y. et al. Cytoplasmic expression of SOX9 as a poor prognostic factor for oral squamous cell carcinoma. *Oncology reports* 40(2018).
177. Pal, SK. et al. THBS1 is induced by TGFB1 in the cancer stroma and promotes invasion of oral squamous cell carcinoma. *Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology* 45(2016).

178. Liu, M. et al. INHBA is a mediator of aggressive tumor behavior in HER2+ basal breast cancer. *Breast cancer research : BCR* 24(2022).
179. Bufalino, A. et al. Low miR-143/miR-145 Cluster Levels Induce Activin A Overexpression in Oral Squamous Cell Carcinomas, Which Contributes to Poor Prognosis. *PLoS one* 10(2015).
180. Yu, CC. et al. miR145 targets the SOX9/ADAM17 axis to inhibit tumor-initiating cells and IL-6-mediated paracrine effects in head and neck cancer. *Cancer research* 73(2013).
181. Knitz, M.W. et al. Targeting resistance to radiation-immunotherapy in cold HNSCCs by modulating the Treg-dendritic cell axis. (2021).
182. Rolle, A., Meyer, M., Calderazzo, S., Jager, D. & Momburg, F. Distinct HLA-E Peptide Complexes Modify Antibody-Driven Effector Functions of Adaptive NK Cells. *Cell Rep* 24, 1967-1976.e4 (2018).
183. van Hall, T. et al. Monalizumab: inhibiting the novel immune checkpoint NKG2A. *Journal for immunotherapy of cancer* 7(2019).
184. André, P. et al. Anti-NKG2A mAb Is a Checkpoint Inhibitor that Promotes Anti-tumor Immunity by Unleashing Both T and NK Cells. *Cell* 175(2018).
185. Lechner, MG. et al. Characterization of cytokine-induced myeloid-derived suppressor cells from normal human peripheral blood mononuclear cells. *Journal of immunology (Baltimore, Md. : 1950)* 185(2010).

## Supplements

Supplementary Table 1. Gene set Signature Based in Inverse SOX2/SOX9 Expression

<i>Gene Name</i>	<i>Gene Description</i>	<i>log2FC</i>	<i>lfcSE</i>	<i>padj</i>	<i>Spearman Sox2</i>	<i>Spearman Sox9</i>	<i>Group</i>
<i>ADAD2</i>	Adenosine Deaminase Domain Containing 2	-1,158	0,186	0,000	0,371	-0,223	G1
<i>AKR1C3</i>	Aldo-Keto Reductase Family 1 Member C3	-2,298	0,291	0,000	0,508	-0,208	G1
<i>ALDH3A1</i>	Aldehyde Dehydrogenase 3 Family Member A1	-2,712	0,317	0,000	0,671	-0,283	G1
<i>EVPLL</i>	Envoplakin Like	-1,598	0,284	0,000	0,414	-0,268	G1
<i>GCHFR</i>	GTP Cyclohydrolase	-1,202	0,166	0,000	0,401	-0,241	G1
<i>GSTA3</i>	Glutathione S-Transferase Alpha 3	-1,746	0,465	0,000	0,277	-0,224	G1
<i>GSTM2</i>	Glutathione S-Transferase Mu 2	-2,486	0,252	0,000	0,449	-0,240	G1
<i>HOGA1</i>	4-Hydroxy-2-Oxoglutarate Aldolase 1	-1,187	0,216	0,000	0,418	-0,236	G1
<i>MIR9-3HG</i>	MIR9-3 Host Gene	-1,195	0,246	0,000	0,352	-0,24	G1
<i>NKD2</i>	NKD Inhibitor of WNT Signaling Pathway 2	-1,297	0,184	0,000	0,330	-0,215	G1
<i>PDIA2</i>	Protein Disulfide Isomerase Family A Member 2	-1,247	0,413	0,002	0,206	-0,209	G1
<i>SERPINI1</i>	Serpin Family I Member 1	-2,606	0,235	0,000	0,402	-0,248	G1
<i>SLC25A48</i>	Solute Carrier Family 25 Member 48	-2,090	0,302	0,000	0,436	-0,215	G1
<i>TSGA10IP</i>	Testis Specific 10 Interacting Protein	-2,339	0,375	0,000	0,261	-0,264	G1
<i>UGT1A7</i>	UDP Glucuronosyltransferase Family 1 Member A7	-4,791	0,437	0,000	0,672	-0,202	G1
<i>UGT1A8</i>	UDP Glucuronosyltransferase Family 1 Member A8	-4,877	0,595	0,000	0,559	-0,219	G1
<i>ACTN1</i>	Actinin Alpha 1	1,112	0,124	0,000	-0,426	0,327	G2
<i>ADRA1B</i>	Adrenoceptor Alpha 1B	1,191	0,300	0,000	-0,262	0,319	G2
<i>ADTRP</i>	Androgen Dependent TFPI Regulating Protein	1,525	0,204	0,000	-0,378	0,209	G2
<i>ANTXR2</i>	ANTXR Cell Adhesion Molecule 2	1,166	0,172	0,000	-0,441	0,232	G2
<i>APCDD1L</i>	APC Down-Regulated 1 Like	1,458	0,264	0,000	-0,255	0,247	G2
<i>ARHGAP29</i>	Rho GTPase Activating Protein 29	1,007	0,171	0,000	-0,388	0,288	G2
<i>BCAR3</i>	BCAR3 Adaptor Protein, NSP Family Member	1,498	0,157	0,000	-0,550	0,277	G2
<i>CAMK2N1</i>	Calcium/Calmodulin Dependent Protein Kinase II Inhibitor 1	1,017	0,188	0,000	-0,428	0,279	G2
<i>CCBE1</i>	Collagen and Calcium Binding EGF Domains 1	1,952	0,265	0,000	-0,462	0,216	G2
<i>CDH3</i>	Cadherin 3	1,253	0,128	0,000	-0,470	0,233	G2
<i>CEBPB</i>	CCAAT Enhancer Binding Protein Beta	1,136	0,114	0,000	-0,464	0,240	G2
<i>CGB8</i>	Chorionic Gonadotropin Subunit Beta 8	2,811	0,420	0,000	-0,354	0,214	G2
<i>CHI3L2</i>	Chitinase 3 Like 2	1,402	0,260	0,000	-0,209	0,200	G2
<i>DCBLD1</i>	Discoidin, CUB and LCCL Domain Containing 1	1,291	0,134	0,000	-0,431	0,338	G2
<i>DUSP4</i>	Dual Specificity Phosphatase 4	1,085	0,194	0,000	-0,365	0,295	G2
<i>DUSP6</i>	Dual Specificity Phosphatase 6	1,338	0,150	0,000	-0,362	0,265	G2
<i>ELFN2</i>	Extracellular Leucine Rich Repeat and Fibronectin Type III Domain Containing 2	2,042	0,334	0,000	-0,263	0,244	G2
<i>EVA1A</i>	Eva-1 Homolog A, Regulator of Programmed Cell Death	1,371	0,222	0,000	-0,392	0,295	G2

<i>FHL2</i>	Four and A Half LIM Domains 2	1,333	0,122	0,000	-0,547	0,226	G2
<i>FOSL1</i>	FOS Like 1, AP-1 Transcription Factor Subunit	1,041	0,177	0,000	-0,337	0,214	G2
<i>FOXC2</i>	Fork head Box C2	2,064	0,249	0,000	-0,256	0,265	G2
<i>FRMD5</i>	FERM Domain Containing 5	1,008	0,259	0,000	-0,293	0,219	G2
<i>INHBA</i>	Inhibin Subunit Beta A	1,609	0,218	0,000	-0,391	0,296	G2
<i>INPP4B</i>	Inositol Polyphosphate-4-Phosphatase Type II B	1,417	0,179	0,000	-0,445	0,238	G2
<i>ITGB6</i>	Integrin Subunit Beta 6	1,392	0,188	0,000	-0,253	0,249	G2
<i>KIF12</i>	Kinesin Family Member 12	1,685	0,271	0,000	-0,458	0,220	G2
<i>L1CAM</i>	L1 Cell Adhesion Molecule	1,488	0,278	0,000	-0,413	0,224	G2
<i>LAMC2</i>	Laminin Subunit Gamma 2	1,285	0,204	0,000	-0,245	0,259	G2
<i>LGALS1</i>	Galectin 1	1,123	0,178	0,000	-0,465	0,242	G2
<i>LINC00452</i>	Long Intergenic Non-Protein Coding RNA 452	2,161	0,391	0,000	-0,203	0,233	G2
<i>LTBP1</i>	Latent Transforming Growth Factor Beta Binding Protein 1	1,178	0,174	0,000	-0,282	0,232	G2
<i>MANCR</i>	Mitotically Associated Long non-coding RNA	1,507	0,265	0,000	-0,316	0,21	G2
<i>NOX5</i>	NADPH Oxidase 5	2,188	0,226	0,000	-0,388	0,289	G2
<i>PCDHGC5</i>	Protocadherin Gamma Subfamily C, 5	2,029	0,324	0,000	-0,315	0,211	G2
<i>PMEPA1</i>	Prostate Transmembrane Protein, Androgen Induced 1	1,247	0,167	0,000	-0,336	0,244	G2
<i>PPP4R4</i>	Protein Phosphatase 4 Regulatory Subunit 4	1,416	0,242	0,000	-0,265	0,269	G2
<i>PRDM8</i>	PR/SET Domain 8	1,120	0,163	0,000	-0,357	0,270	G2
<i>RP1</i>	RP1 Axonemal Microtubule Associated	1,723	0,306	0,000	-0,249	0,252	G2
<i>RPSAP52</i>	Ribosomal Protein SA Pseudogene 52	1,927	0,273	0,000	-0,384	0,299	G2
<i>RTTN</i>	Rotatin	1,038	0,153	0,000	-0,356	0,215	G2
<i>SEC14L2</i>	SEC14 Like Lipid Binding 2	1,535	0,169	0,000	-0,433	0,392	G2
<i>SERPINB7</i>	Serpin Family B Member 7	1,747	0,278	0,000	-0,331	0,228	G2
<i>SERPINE1</i>	Serpin Family E Member 1	1,543	0,210	0,000	-0,361	0,376	G2
<i>SH3TC2</i>	SH3 Domain and Tetratricopeptide Repeats 2	1,219	0,189	0,000	-0,432	0,234	G2
<i>SRPX</i>	Sushi Repeat Containing Protein X-Linked	1,024	0,226	0,000	-0,290	0,231	G2
<i>TENM3</i>	Teneurin Transmembrane Protein 3	1,071	0,222	0,000	-0,313	0,231	G2
<i>THBS1</i>	Thrombospondin 1	1,120	0,212	0,000	-0,246	0,216	G2
<i>TINAGL1</i>	Tubulointerstitial Nephritis Antigen Like 1	1,253	0,177	0,000	-0,277	0,344	G2
<i>TMEM92</i>	Transmembrane Protein 92	1,716	0,187	0,000	-0,587	0,225	G2
<i>TNFRSF12A</i>	TNF Receptor Superfamily Member 12A	1,287	0,129	0,000	-0,497	0,290	G2
<i>TNS4</i>	Tensin 4	1,070	0,142	0,000	-0,460	0,204	G2
<i>TREM1</i>	Triggering Receptor Expressed on Myeloid Cells 1	1,069	0,245	0,000	-0,337	0,209	G2
<i>TRIML2</i>	Tripartite Motif Family Like 2	3,328	0,405	0,000	-0,499	0,205	G2

Supplementary Table 2. Differential Expressed Genes by Inverse SOX2/SOX9 Expression among G1 versus G2 of the TCGA-HNSCC Cohort

DEGs downregulated / DEGs up-regulated

Symbol	Log2FoldChange	Padj	Symbol	Log2FoldChange	Padj	Symbol	Log2FoldChange	Padj	Symbol	Log2FoldChange	Padj
UCN3	-7,643E+00	8,881E-07	PLAC8	-2,021E+00	5,587E-13	LOC105370792	-1,505E+00	5,436E-06	FBXO15	-1,204E+00	3,357E-07
PCSK2	-6,466E+00	1,256E-19	ADAM1B	-2,017E+00	6,861E-16	ITLN2	-1,505E+00	1,689E-06	ARFGEF3	-1,204E+00	2,653E-05
LINC01206	-5,460E+00	1,256E-50	C10orf90	-2,017E+00	3,248E-09	KIAA1671-AS1	-1,504E+00	1,413E-07	RGS22	-1,203E+00	1,249E-06
KRT20	-5,337E+00	2,569E-14	CLCNKA	-2,016E+00	2,568E-03	ENPP7P8	-1,504E+00	2,592E-06	CHRM2	-1,203E+00	1,626E-03
FAM167A-AS1	-5,309E+00	2,367E-32	ADH7	-2,015E+00	8,296E-08	HCG23	-1,503E+00	1,613E-06	KCNN1	-1,203E+00	1,389E-06
LINC01413	-5,048E+00	1,803E-03	GSTM5	-2,014E+00	1,062E-07	SLC13A2	-1,502E+00	1,323E-04	GCHFR	-1,202E+00	5,394E-12
GHRH	-4,972E+00	2,703E-12	PTH2R	-2,014E+00	1,033E-05	ENPP7P7	-1,501E+00	7,421E-04	SLC7A11	-1,201E+00	1,425E-05
GRM1	-4,946E+00	2,664E-19	GPHB5	-2,009E+00	7,683E-03	MYLIP	-1,501E+00	2,845E-22	LEF1	-1,201E+00	4,979E-12
ADCY8	-4,920E+00	1,150E-12	GPX2	-2,008E+00	9,567E-11	PAX2	-1,500E+00	2,265E-04	SOX13	-1,200E+00	1,232E-18
UGT1A8	-4,877E+00	9,240E-15	ABCA13	-2,004E+00	1,419E-12	RNY4P36	-1,499E+00	4,175E-04	LINC02809	-1,199E+00	5,641E-05
WIFI	-4,852E+00	1,218E-21	GPLDI	-2,004E+00	3,457E-15	M1AP	-1,499E+00	4,010E-05	FAM237A	-1,199E+00	1,327E-03
UGT1A7	-4,791E+00	1,487E-25	LINC02753	-2,003E+00	1,936E-04	RNA5SP490	-1,498E+00	8,200E-05	CLDN18	-1,198E+00	5,414E-06
MUC6	-4,735E+00	8,056E-04	LINC02232	-2,003E+00	1,430E-04	MT3	-1,498E+00	9,243E-04	SLC35G6	-1,198E+00	7,204E-07
PAH	-4,659E+00	1,178E-06	KLK15	-2,002E+00	1,963E-07	PCDHA8	-1,497E+00	7,637E-03	MXN1	-1,198E+00	1,437E-04
TAC3	-4,594E+00	2,187E-06	LOC100132215	-2,000E+00	1,146E-22	MLXIPL	-1,497E+00	5,720E-06	LOC105371050	-1,198E+00	2,458E-10
SHH	-4,586E+00	5,022E-40	SLC7A11-AS1	-1,998E+00	1,082E-08	NCR3LG1	-1,496E+00	7,392E-11	BCL2L15	-1,197E+00	1,673E-05
PAK5	-4,578E+00	3,176E-26	NDUFB4P11	-1,994E+00	3,020E-07	LTK	-1,495E+00	7,801E-10	BEST2	-1,195E+00	2,460E-03
FOXB2	-4,558E+00	1,704E-03	GSTA1	-1,991E+00	1,138E-13	PBX1-AS1	-1,491E+00	4,872E-03	MIR9-3HG	-1,195E+00	2,761E-06
SLC9A4	-4,532E+00	5,467E-36	RASL11A	-1,991E+00	1,069E-29	KRTAP17-1	-1,491E+00	2,005E-03	TMC2	-1,195E+00	9,771E-06
DBX1	-4,483E+00	1,578E-03	LDLRAD1	-1,990E+00	9,207E-07	SOSTDC1	-1,489E+00	1,319E-04	NRXN2	-1,194E+00	3,171E-06
LOC105374042	-4,448E+00	1,016E-13	TMEM246-AS1	-1,989E+00	1,162E-10	SLC13A4	-1,489E+00	2,055E-03	RNFT2	-1,193E+00	2,842E-06
ZIC3	-4,410E+00	4,627E-11	LINC01564	-1,987E+00	4,601E-16	QRFPR	-1,487E+00	3,022E-03	CARS1P2	-1,193E+00	2,889E-03
FAM153B	-4,392E+00	7,707E-31	OTX1	-1,987E+00	7,936E-32	PIP5K1B	-1,487E+00	7,594E-10	SLC44A4	-1,190E+00	2,809E-04
EPHA7	-4,384E+00	3,494E-09	IGHV1-45	-1,985E+00	1,866E-06	RPL31P43	-1,487E+00	8,880E-06	CLEC18B	-1,190E+00	4,524E-06
FGF19	-4,354E+00	1,787E-13	STUB1-DT	-1,985E+00	1,965E-11	RNU6-242P	-1,486E+00	4,719E-03	NUDT16-DT	-1,189E+00	7,837E-08
MPPED1	-4,343E+00	3,007E-26	SCIN	-1,985E+00	6,557E-17	SEMA6D	-1,485E+00	1,319E-08	NCMAP	-1,189E+00	4,180E-05
LOC101927870	-4,316E+00	1,440E-12	RND2	-1,983E+00	7,749E-15	TENT5D	-1,485E+00	4,628E-03	ROBO2	-1,188E+00	1,682E-04
HS3ST5	-4,181E+00	1,148E-20	TDH	-1,980E+00	1,973E-12	LINC01133	-1,483E+00	4,195E-07	CD200R1L	-1,188E+00	3,755E-05
GABRA5	-4,129E+00	9,816E-19	RBP4	-1,977E+00	7,952E-07	LOC339666	-1,481E+00	2,965E-11	LOC101927855	-1,187E+00	2,006E-05
SLC5A7	-4,124E+00	8,161E-14	ATRNL1	-1,973E+00	4,444E-10	LINC02671	-1,481E+00	1,657E-03	HOGA1	-1,187E+00	1,443E-07
DLEC1	-4,068E+00	1,026E-51	LOC349160	-1,972E+00	1,524E-06	SMOC2	-1,480E+00	5,594E-11	ZNF737	-1,186E+00	1,369E-07
DCC	-4,044E+00	1,599E-25	SLC6A4	-1,972E+00	1,161E-10	HCN1	-1,479E+00	5,428E-04	MUC17	-1,186E+00	3,062E-03
KLHDC8A	-3,999E+00	1,187E-45	RASSF9	-1,972E+00	3,271E-15	RN7SL399P	-1,479E+00	4,683E-06	UCHL1	-1,186E+00	3,281E-04
SOX2-OT	-3,987E+00	1,119E-54	GABRB2	-1,970E+00	3,194E-09	P2RX2	-1,479E+00	4,564E-10	SYT3	-1,186E+00	4,542E-05
NOTUM	-3,970E+00	3,894E-14	OR5BH1P	-1,970E+00	3,754E-06	SUSD4	-1,477E+00	3,228E-09	FUT6	-1,185E+00	7,341E-04

CPLX2	-3,961E+00	1,184E-26	IGSF11	-1,968E+00	9,654E-12	GSTT4	-1,476E+00	6,550E-05	LINC01474	-1,185E+00	4,990E-04
LOC105378137	-3,952E+00	1,105E-09	PAPPA2	-1,967E+00	3,301E-05	ZIC4	-1,472E+00	3,712E-04	PHYHD1	-1,184E+00	1,130E-04
NMUR2	-3,919E+00	2,808E-16	LOC100420939	-1,962E+00	6,523E-03	CPNE6	-1,472E+00	2,356E-05	ST6GAL2	-1,184E+00	5,089E-04
LOC101928441	-3,897E+00	1,692E-13	S100A5	-1,962E+00	8,688E-24	WNT16	-1,471E+00	5,850E-06	DISC1	-1,182E+00	1,315E-11
FGFBP2	-3,818E+00	6,454E-19	FAM71F1	-1,962E+00	1,915E-06	THRSP	-1,470E+00	1,207E-04	CHL1-AS1	-1,182E+00	2,248E-03
MAEL	-3,808E+00	2,639E-14	CFAP91	-1,961E+00	2,097E-10	LOC100506405	-1,469E+00	1,217E-03	MSH2-OT1	-1,182E+00	7,861E-05
SYT9	-3,719E+00	2,138E-25	TMEM271	-1,960E+00	3,849E-04	TTC9-DT	-1,469E+00	4,199E-07	ME1	-1,182E+00	5,503E-11
PCK1	-3,709E+00	7,621E-06	MSGN1	-1,957E+00	4,547E-15	OR2W6P	-1,468E+00	1,660E-05	WARS1P1	-1,181E+00	5,533E-03
LINC00615	-3,702E+00	2,496E-16	NLGN1	-1,952E+00	7,072E-10	ZDHHC11B	-1,467E+00	2,258E-07	MORN3	-1,179E+00	1,283E-05
LIM2	-3,640E+00	8,818E-10	INSM1	-1,951E+00	6,403E-07	PIPOX	-1,465E+00	6,904E-09	THSD4-AS1	-1,178E+00	6,575E-04
CYP1A1	-3,621E+00	1,059E-15	ABCA4	-1,949E+00	1,127E-09	LOC101927815	-1,465E+00	4,066E-06	AK7	-1,178E+00	2,228E-08
COLGALT2	-3,610E+00	8,100E-44	ALCB	-1,949E+00	2,861E-11	GULP1	-1,464E+00	2,826E-10	LINC00645	-1,178E+00	3,534E-04
SERPINI2	-3,524E+00	1,446E-12	PPP1R14D	-1,948E+00	2,306E-09	FGF14-AS2	-1,462E+00	3,134E-10	TULP1	-1,177E+00	2,713E-09
NEUROD2	-3,519E+00	9,131E-27	WDR87BP	-1,948E+00	6,202E-09	RPL21P63	-1,461E+00	1,097E-04	PANX2	-1,176E+00	2,919E-06
TRPM1	-3,502E+00	1,146E-22	FOLH1	-1,941E+00	2,032E-23	TMEM178A	-1,461E+00	2,156E-07	CYP2C19	-1,176E+00	3,340E-04
LINC02525	-3,498E+00	3,598E-06	RPS6KA6	-1,941E+00	4,195E-06	PHACTR3	-1,458E+00	1,882E-05	ZDHHC19	-1,176E+00	2,818E-05
LINC00851	-3,493E+00	1,813E-11	CHAT	-1,938E+00	1,572E-04	TBX2	-1,458E+00	4,868E-16	CRMP1	-1,174E+00	2,713E-11
OTOP1	-3,492E+00	2,287E-04	DCXR-DT	-1,936E+00	6,171E-12	ARPC1BP1	-1,455E+00	4,586E-04	LINC02071	-1,174E+00	7,570E-03
TFAP2D	-3,487E+00	1,757E-05	CT75	-1,935E+00	5,781E-18	ST18	-1,452E+00	3,396E-08	PNLDC1	-1,173E+00	4,805E-03
FRRS1L	-3,484E+00	7,558E-19	SST	-1,934E+00	2,072E-05	ARMC3	-1,451E+00	3,362E-07	NOS2	-1,173E+00	4,769E-11
ANKRD63	-3,472E+00	5,948E-13	ZBTB7C	-1,930E+00	8,654E-16	MUC20	-1,447E+00	1,333E-06	UPB1	-1,172E+00	1,715E-03
GPC3	-3,456E+00	2,724E-26	SYT2	-1,929E+00	1,156E-14	LINC00462	-1,447E+00	8,734E-08	MESP1	-1,170E+00	1,816E-05
MOBP	-3,441E+00	2,402E-22	TMEM210	-1,928E+00	9,883E-05	RCOR2	-1,444E+00	7,568E-10	OR6C70	-1,170E+00	9,791E-03
KCNJ3	-3,410E+00	1,678E-05	ZBBX	-1,922E+00	8,447E-05	C3orf70	-1,439E+00	5,295E-11	ZNF793	-1,169E+00	6,730E-06
SLC9A2	-3,407E+00	1,946E-31	UQCRBP2	-1,922E+00	1,641E-04	SCUBE2	-1,438E+00	4,599E-09	SERTAD4	-1,169E+00	4,085E-10
VIL1	-3,403E+00	1,805E-27	PIWIL2	-1,920E+00	3,598E-08	TAS1R3	-1,436E+00	1,103E-09	LINC00858	-1,169E+00	3,495E-03
P2RX3	-3,390E+00	6,005E-05	ARGFX	-1,920E+00	9,573E-05	GRAMD4P8	-1,436E+00	1,096E-03	ENPP7	-1,168E+00	5,873E-04
LOC100286922	-3,384E+00	2,119E-12	PLPPR1	-1,918E+00	1,356E-05	CLDN19	-1,435E+00	1,329E-03	YJEFN3	-1,166E+00	6,724E-07
FAM153A	-3,376E+00	1,366E-20	MEDI5P5	-1,909E+00	5,243E-05	RNU6-554P	-1,434E+00	2,023E-03	RASL10B	-1,165E+00	1,349E-06
NKX2-4	-3,371E+00	1,592E-08	LINC00051	-1,906E+00	2,821E-04	FZD7	-1,432E+00	1,767E-11	LINC01981	-1,164E+00	1,288E-04
ALDH1A1	-3,349E+00	1,032E-27	ENTPD8	-1,905E+00	8,341E-07	AMH	-1,430E+00	3,496E-06	TUBA3C	-1,164E+00	2,719E-03
LOC158434	-3,348E+00	1,737E-07	RNF183	-1,905E+00	4,710E-07	GPR15	-1,430E+00	1,948E-05	TBL1X	-1,163E+00	1,535E-14
CALHM1	-3,334E+00	7,044E-16	MYO3A	-1,904E+00	2,633E-06	C9orf24	-1,429E+00	2,395E-09	XXYLT1-AS1	-1,162E+00	2,722E-05
SCN9A	-3,327E+00	2,446E-26	RIMKLA	-1,903E+00	9,319E-09	ONECUT1	-1,428E+00	4,625E-06	ADAM7	-1,161E+00	4,044E-03
CYP26A1	-3,321E+00	2,254E-12	C1orf194	-1,903E+00	6,327E-07	MCF2L	-1,428E+00	5,084E-14	USP44	-1,160E+00	3,649E-07
GABRG3	-3,308E+00	1,635E-09	CYP4F11	-1,898E+00	2,611E-08	CFAP47	-1,428E+00	3,375E-04	PDE5A	-1,159E+00	1,520E-11
DPP10	-3,297E+00	3,087E-08	MRPL23-AS1	-1,897E+00	2,112E-04	RPL7L1P8	-1,427E+00	1,417E-05	ZNF610	-1,159E+00	1,687E-07
DSG4	-3,296E+00	2,292E-14	FAM3B	-1,896E+00	3,614E-07	RFPL3S	-1,426E+00	1,315E-09	LINC02145	-1,158E+00	9,840E-05
AKR1C4	-3,281E+00	8,334E-18	OXGR1	-1,896E+00	3,739E-08	LINC00677	-1,426E+00	2,536E-07	SLC16A11	-1,158E+00	3,426E-07
MAP4K1-AS1	-3,274E+00	8,031E-17	PCAT2	-1,894E+00	1,799E-07	VWCE	-1,425E+00	1,897E-14	ADAD2	-1,158E+00	2,574E-09
COCH	-3,272E+00	5,691E-28	NFE4	-1,893E+00	3,291E-10	ERC2	-1,425E+00	9,863E-11	LOC107985164	-1,157E+00	1,506E-04
FEZF1	-3,272E+00	6,397E-14	C11orf16	-1,890E+00	6,935E-09	LINC01697	-1,424E+00	4,382E-04	LOC101928844	-1,156E+00	2,132E-04
PLAAT5	-3,245E+00	6,555E-16	MIOX	-1,888E+00	1,337E-06	MFSD4A	-1,422E+00	1,174E-06	MIXL1	-1,156E+00	5,752E-06
SCGB1A1	-3,238E+00	2,699E-03	GRPR	-1,887E+00	1,532E-10	SFRP5	-1,422E+00	6,288E-04	SYT1	-1,156E+00	8,735E-06



RPL21P13	-3,223E+00	1,166E-12	LINC00092	-1,886E+00	3,419E-09	TRPM5	-1,420E+00	5,871E-04	NRN1	-1,156E+00	5,516E-06
PHOX2A	-3,215E+00	1,842E-04	WNK2	-1,885E+00	1,484E-09	LINC01605	-1,420E+00	1,350E-08	SYNE4	-1,154E+00	3,651E-04
LINC01895	-3,214E+00	9,238E-03	ASCL1	-1,885E+00	1,815E-06	HOXA9	-1,420E+00	1,700E-05	GABRR1	-1,154E+00	6,169E-05
DLGAP1	-3,206E+00	2,017E-27	PLCH1	-1,884E+00	1,431E-11	BCO1	-1,419E+00	7,418E-06	B4GALNT4	-1,153E+00	6,920E-05
UGT1A9	-3,195E+00	8,445E-12	UGT1A1	-1,883E+00	8,891E-10	PLCH1-AS2	-1,418E+00	3,792E-03	MC5R	-1,153E+00	3,483E-03
LY6G6F	-3,184E+00	7,118E-13	LINC01475	-1,883E+00	2,977E-04	PRSS30P	-1,417E+00	6,249E-03	HKDC1	-1,152E+00	3,379E-04
ZMAT4	-3,160E+00	1,149E-13	RSPH6A	-1,879E+00	7,955E-05	LRRTM4	-1,417E+00	6,414E-05	HUNK	-1,151E+00	1,376E-07
KIAA0319	-3,152E+00	1,334E-22	KRT71	-1,878E+00	1,221E-09	RGSS	-1,416E+00	2,262E-10	CGN	-1,150E+00	3,655E-09
CYP2A7	-3,137E+00	6,942E-08	FOLH1B	-1,878E+00	1,767E-08	LINC01087	-1,415E+00	6,537E-03	CBSL	-1,149E+00	2,843E-03
DNAJB3	-3,133E+00	1,330E-18	NPSR1	-1,877E+00	1,079E-09	RIPPLY2	-1,415E+00	5,055E-03	SRXN1	-1,148E+00	9,106E-09
PTPRT	-3,128E+00	1,398E-15	GRK7	-1,876E+00	1,523E-21	HOATZ	-1,414E+00	1,202E-08	RNF165	-1,148E+00	8,975E-06
LGSN	-3,113E+00	8,025E-11	RGS6	-1,875E+00	1,581E-12	LOC107985122	-1,412E+00	7,470E-04	RGS7BP	-1,147E+00	1,852E-03
LINC01250	-3,109E+00	3,193E-13	CEACAM16	-1,874E+00	2,490E-05	CBX2	-1,409E+00	4,511E-19	CHRNA7	-1,146E+00	8,820E-05
SLCO4C1	-3,106E+00	3,842E-20	EMILIN3	-1,874E+00	1,248E-12	ANKRD40CL	-1,407E+00	7,551E-04	CYP20	-1,146E+00	4,223E-05
SOX2	-3,104E+00	4,032E-35	LINC02043	-1,871E+00	3,674E-14	LINC02135	-1,402E+00	1,667E-03	SLIT2	-1,146E+00	1,082E-06
UGT1A3	-3,098E+00	2,343E-14	ZNF541	-1,871E+00	5,005E-09	TXNRD1	-1,399E+00	1,422E-12	PGD	-1,146E+00	1,019E-11
LINC02382	-3,096E+00	2,163E-04	MC4R	-1,869E+00	1,062E-07	OR56B4	-1,397E+00	2,325E-05	C20orf204	-1,145E+00	4,587E-08
GSTA9P	-3,094E+00	4,343E-10	TCTE1	-1,869E+00	6,790E-11	CCDC74A	-1,394E+00	2,913E-10	ZIC1	-1,144E+00	6,034E-03
CLDN20	-3,077E+00	9,160E-21	EWSAT1	-1,868E+00	1,391E-09	CABP1	-1,394E+00	2,068E-04	TEX26	-1,144E+00	4,242E-04
SLC38A11	-3,068E+00	2,203E-09	CDH12	-1,868E+00	7,946E-04	PSPC1P1	-1,393E+00	5,041E-08	LOC100421620	-1,144E+00	9,185E-04
SLC22A11	-3,067E+00	7,164E-14	SMIM10L2A	-1,865E+00	2,319E-13	MKRN7P	-1,389E+00	3,843E-03	CT83	-1,144E+00	6,197E-03
FGF13-AS1	-3,066E+00	7,181E-03	SLC16A14	-1,863E+00	2,006E-20	CYP2B7P	-1,388E+00	3,528E-05	DYDC2	-1,144E+00	3,381E-04
MUC2	-3,055E+00	5,064E-13	LOC101927708	-1,863E+00	2,070E-07	SLC5A4-AS1	-1,387E+00	6,531E-04	MALRD1	-1,143E+00	2,590E-05
SOX1-OT	-3,054E+00	9,259E-03	PTCSC2	-1,863E+00	3,413E-10	IYD	-1,387E+00	5,460E-05	ITPKA	-1,143E+00	1,699E-09
SHISA9	-3,051E+00	1,308E-08	CACNA1E	-1,863E+00	1,586E-13	WNT8B	-1,387E+00	9,681E-06	ALOX12	-1,143E+00	3,752E-05
WDR11-AS1	-3,048E+00	2,087E-07	GRIK5	-1,862E+00	3,774E-13	KCNS3	-1,386E+00	1,643E-14	SLC25A47	-1,143E+00	2,071E-05
LINC02393	-3,043E+00	2,306E-09	MIAT	-1,861E+00	1,109E-15	SHISA8	-1,386E+00	4,356E-05	GRIK2	-1,141E+00	1,927E-03
ARHGEF26	-3,040E+00	9,308E-31	C9orf152	-1,857E+00	1,761E-06	COL18A1-AS2	-1,386E+00	2,113E-04	PLCB4	-1,141E+00	4,427E-05
LINC00871	-3,034E+00	7,516E-22	PTCH1	-1,857E+00	9,599E-25	STK32B	-1,386E+00	6,402E-10	CALY	-1,140E+00	5,977E-04
UROCI	-3,032E+00	1,415E-04	BANCR	-1,857E+00	1,959E-07	LOC100421238	-1,385E+00	1,938E-03	RNU6-418P	-1,140E+00	5,910E-05
LOC101928596	-3,027E+00	1,422E-05	NR5A1	-1,854E+00	6,842E-04	CHST7	-1,384E+00	4,645E-17	KCNMB3	-1,138E+00	1,473E-09
CA4	-3,024E+00	9,760E-04	KCNMB2	-1,853E+00	1,152E-16	SYCP2L	-1,383E+00	2,134E-05	CNN2P6	-1,135E+00	2,048E-03
GSTA12P	-3,013E+00	8,200E-05	LRP1B	-1,853E+00	1,014E-07	MERTK	-1,382E+00	1,119E-12	GDPD1	-1,135E+00	2,388E-14
OR7A5	-3,010E+00	1,050E-20	SSTR5	-1,852E+00	1,196E-07	CELF2-AS1	-1,382E+00	1,338E-05	RNU6-720P	-1,135E+00	2,724E-04
LINC00661	-2,998E+00	7,347E-09	CYP2AB1P	-1,850E+00	1,336E-08	GPR62	-1,381E+00	5,609E-10	LOC100130760	-1,135E+00	9,429E-03
PCDH11X	-2,990E+00	3,807E-10	AMER2	-1,849E+00	1,084E-05	FREM1	-1,381E+00	3,655E-11	FAM222A	-1,135E+00	2,086E-14
TBX4	-2,988E+00	2,134E-18	UGT8	-1,848E+00	2,588E-07	MORN5	-1,380E+00	3,193E-03	NTN3	-1,135E+00	3,067E-05
LY6G6F-LY6G6D	-2,973E+00	2,389E-03	CCDC136	-1,846E+00	3,669E-20	CSMD1	-1,379E+00	3,168E-04	KLHL13	-1,133E+00	3,444E-11
AADAC	-2,955E+00	1,087E-15	SEMG1	-1,839E+00	1,187E-04	EPB41L1	-1,379E+00	3,960E-16	TET1	-1,132E+00	4,990E-08
BEND4	-2,929E+00	5,192E-14	PVALB	-1,839E+00	5,103E-08	SLC22A9	-1,379E+00	8,905E-06	HOXB-AS3	-1,132E+00	1,689E-04
LINC01994	-2,927E+00	2,877E-17	ADAM23	-1,837E+00	1,668E-10	ETNK2	-1,379E+00	6,111E-14	RALGPS1	-1,131E+00	1,844E-15
JAKMIP2	-2,918E+00	9,449E-28	LINC01695	-1,834E+00	1,419E-06	LINC02313	-1,379E+00	1,765E-04	RBM20	-1,130E+00	3,278E-06
KCNH8	-2,897E+00	3,457E-15	LOC107984827	-1,832E+00	1,308E-06	WNT2B	-1,378E+00	6,831E-10	SYPL1P2	-1,130E+00	4,390E-08
RPL36AP1	-2,896E+00	5,219E-11	TTC6	-1,829E+00	9,612E-07	CYP2T3P	-1,377E+00	2,453E-06	STAG3	-1,130E+00	9,936E-08

LINC02301	-2,884E+00	2,128E-06	PRKX	-1,826E+00	3,403E-43	MICU3	-1,377E+00	7,152E-10	WNK3	-1,130E+00	1,009E-05
CPA2	-2,884E+00	3,932E-20	GSTM4	-1,825E+00	1,126E-22	RPL27P5	-1,377E+00	1,528E-04	ETV1	-1,130E+00	6,414E-08
DDC	-2,873E+00	1,160E-12	TBX2-AS1	-1,822E+00	9,878E-16	PROX1-AS1	-1,376E+00	5,062E-06	TPRG1	-1,129E+00	1,964E-09
NELL1	-2,863E+00	5,392E-13	BPIFB1	-1,820E+00	2,348E-04	FBN2	-1,376E+00	4,063E-06	SEMA6A	-1,128E+00	6,327E-07
SLC26A5	-2,861E+00	7,268E-24	CES3	-1,819E+00	7,390E-14	ANKLE1	-1,376E+00	1,227E-09	DGCR5	-1,128E+00	8,666E-06
UPK1B	-2,858E+00	7,392E-09	NRCAM	-1,818E+00	9,527E-10	ZNF385C	-1,374E+00	1,926E-08	CCDC74B	-1,127E+00	1,796E-06
TDH-AS1	-2,843E+00	1,155E-19	CCDC177	-1,817E+00	5,778E-09	PBX1	-1,374E+00	5,890E-11	PLA2G4A	-1,126E+00	1,520E-06
CYP2A13	-2,829E+00	2,125E-07	APOBEC4	-1,815E+00	5,566E-04	TPO	-1,374E+00	7,725E-07	PINLYP	-1,125E+00	6,095E-08
RNA5SP111	-2,820E+00	2,138E-06	RNY4P19	-1,813E+00	2,872E-08	VSIG2	-1,374E+00	9,241E-06	LINC00954	-1,124E+00	2,355E-06
RFX4	-2,817E+00	2,811E-06	MYH11	-1,812E+00	4,434E-10	NPIPB9	-1,373E+00	8,363E-05	LRRC37A6P	-1,122E+00	9,937E-06
CYP4F3	-2,814E+00	1,558E-15	GLB1L2	-1,810E+00	3,356E-09	FNDC8	-1,373E+00	5,436E-10	LL22NC03	-1,121E+00	2,408E-03
DCAF12L2	-2,809E+00	1,257E-13	CLDN3	-1,809E+00	2,149E-05	CRHR2	-1,373E+00	3,590E-08	METTL27	-1,120E+00	6,872E-06
RNA5SP515	-2,806E+00	1,211E-04	ABCC2	-1,807E+00	1,604E-08	TROAP-AS1	-1,372E+00	2,824E-09	PPM1L	-1,120E+00	4,890E-09
LOC400622	-2,801E+00	2,733E-08	MUC13	-1,803E+00	2,382E-06	SNAP91	-1,372E+00	3,167E-04	HS3ST4	-1,119E+00	5,966E-03
LINC01342	-2,790E+00	3,291E-10	DMRT1	-1,803E+00	7,442E-06	NPR3	-1,372E+00	1,554E-06	ALCAM	-1,118E+00	5,377E-11
COLEC11	-2,789E+00	6,688E-24	LMX1A	-1,803E+00	1,122E-05	FXYP4	-1,371E+00	9,533E-05	GPRC5B	-1,118E+00	8,674E-08
LINC01414	-2,786E+00	5,862E-09	PCYT1B	-1,802E+00	4,064E-08	FGF12	-1,369E+00	1,545E-06	EPHX1	-1,118E+00	2,617E-10
ATP4A	-2,784E+00	1,221E-13	C1QTNF8	-1,801E+00	3,271E-03	MUC4	-1,369E+00	3,266E-04	SNORA14A	-1,117E+00	4,811E-04
GAL3ST2	-2,783E+00	2,174E-18	SOHLH1	-1,801E+00	4,759E-03	TSPAN12	-1,367E+00	2,308E-12	TRPV4	-1,117E+00	3,754E-11
GFRA3	-2,761E+00	3,671E-16	CCDC187	-1,800E+00	4,843E-06	PRRG3	-1,367E+00	1,526E-05	CRACD	-1,117E+00	4,960E-06
NMRAL2P	-2,753E+00	5,075E-19	FOXE1	-1,800E+00	6,261E-18	GRM5P1	-1,367E+00	2,710E-03	AKR1B10	-1,116E+00	1,274E-04
NRXN3	-2,742E+00	5,365E-24	ADD2	-1,797E+00	1,921E-07	PPM1H	-1,367E+00	3,789E-10	ACTG1P25	-1,114E+00	2,267E-06
RAB6B	-2,740E+00	5,631E-32	OR56A5	-1,797E+00	5,229E-04	FGF10-AS1	-1,367E+00	1,243E-03	AACSP1	-1,114E+00	5,037E-04
NTS	-2,731E+00	5,176E-17	LOC100422564	-1,792E+00	1,232E-03	ASRGL1	-1,367E+00	1,623E-12	CFAP100	-1,114E+00	1,461E-05
LINC01992	-2,731E+00	2,180E-06	SLC35D3	-1,792E+00	2,092E-06	SYCP2	-1,366E+00	7,669E-08	PGAP1	-1,113E+00	1,941E-16
FTCD	-2,720E+00	2,268E-06	COLCA2	-1,792E+00	7,515E-11	NOS1	-1,366E+00	4,610E-05	STPG3	-1,112E+00	1,692E-04
ALDH3A1	-2,712E+00	2,929E-16	H3P9	-1,791E+00	4,284E-04	VPS37D	-1,364E+00	4,796E-11	ATP2B1	-1,111E+00	1,023E-18
EPS8L3	-2,711E+00	4,123E-03	SLC5A4	-1,790E+00	4,234E-08	HTR3B	-1,360E+00	1,753E-03	SLC4A9	-1,110E+00	7,789E-07
MRAP2	-2,710E+00	2,440E-19	MYO5C	-1,785E+00	2,500E-15	BMP4	-1,359E+00	1,258E-03	MAP3K15	-1,109E+00	2,473E-07
RAB3B	-2,703E+00	6,521E-20	C10orf105	-1,785E+00	1,404E-07	DOC2A	-1,359E+00	1,016E-06	MSLN	-1,109E+00	3,128E-03
LINC01224	-2,703E+00	3,667E-09	LINC01249	-1,784E+00	2,638E-03	GCLC	-1,357E+00	3,505E-14	TAS2R31	-1,109E+00	5,042E-06
GPR149	-2,703E+00	3,886E-05	CLEC2L	-1,780E+00	5,167E-05	PRH1	-1,357E+00	4,319E-04	ANKDD1B	-1,108E+00	3,867E-09
FSTL5	-2,699E+00	1,090E-06	ELF5	-1,779E+00	3,486E-06	C2CD6	-1,356E+00	3,868E-06	LOC101928372	-1,108E+00	1,769E-03
PPIAP46	-2,697E+00	1,353E-04	TMEM221	-1,777E+00	3,934E-17	TBX5	-1,355E+00	2,238E-03	RETREG1-AS1	-1,107E+00	6,384E-06
LINC02261	-2,693E+00	5,516E-08	SPDYC	-1,776E+00	3,629E-04	LINC01979	-1,355E+00	5,018E-06	CHRM5	-1,106E+00	6,440E-05
RN7SL585P	-2,687E+00	7,577E-04	GLI2	-1,775E+00	1,097E-12	UGT2A1	-1,354E+00	6,548E-03	TNFRSF18	-1,106E+00	4,195E-07
WDR49	-2,686E+00	2,439E-13	MYEF2	-1,773E+00	2,263E-12	OR13A1	-1,352E+00	3,621E-05	SLC9A9	-1,106E+00	1,360E-09
JAKMIP3	-2,663E+00	3,048E-18	LOC100506071	-1,772E+00	1,643E-08	CARMIL3	-1,351E+00	2,163E-10	TEX41	-1,105E+00	4,451E-05
SOST	-2,661E+00	3,290E-14	MYB	-1,771E+00	3,590E-13	DPEP3	-1,351E+00	2,083E-03	AASS	-1,104E+00	2,427E-11
DIPK1C	-2,658E+00	2,141E-19	FAM216B	-1,769E+00	2,873E-06	ARHGAP24	-1,350E+00	4,993E-16	MIR554	-1,104E+00	7,415E-04
LOC101927815	-2,657E+00	1,778E-04	WFIKKN2	-1,768E+00	3,768E-03	GPR63	-1,350E+00	4,203E-07	SH3BGR2	-1,104E+00	5,201E-07
LOC100507336	-2,657E+00	1,985E-09	SOX21-AS1	-1,767E+00	8,099E-12	PTPRQ	-1,347E+00	1,580E-05	ZIC2	-1,103E+00	3,152E-07
LHFPL3	-2,645E+00	5,613E-08	PHYHIP	-1,764E+00	6,738E-14	SLC47A1	-1,347E+00	1,501E-07	ZBED3-AS1	-1,103E+00	6,984E-09
OR7E12P	-2,644E+00	4,136E-10	PRDM13	-1,763E+00	1,178E-05	TRIM16L	-1,347E+00	2,446E-09	KLRK1-AS1	-1,103E+00	8,028E-08

LOC730101	-2,642E+00	1,171E-24	FENDRR	-1,761E+00	1,727E-11	FGF17	-1,346E+00	1,926E-04	NR1I2	-1,102E+00	2,234E-04
LOC105369980	-2,642E+00	4,410E-09	PPPIR9A	-1,760E+00	4,367E-07	LOC101928495	-1,346E+00	2,922E-03	RSPH14	-1,102E+00	3,416E-05
UCP1	-2,640E+00	2,988E-07	LINC01484	-1,758E+00	4,839E-08	LCN1	-1,346E+00	7,102E-09	LNC-LBCS	-1,101E+00	7,764E-06
UICLM	-2,640E+00	8,254E-04	GSTM1	-1,755E+00	1,977E-04	ZNF334	-1,344E+00	2,715E-07	LOC100130550	-1,101E+00	2,288E-04
UMODL1	-2,637E+00	1,490E-17	CABYR	-1,754E+00	1,643E-14	MCIDAS	-1,344E+00	3,293E-12	DRC1	-1,101E+00	6,663E-06
SEZ6L	-2,630E+00	6,137E-13	LGI3	-1,752E+00	2,231E-07	LINC00885	-1,344E+00	6,058E-08	B3GALT2	-1,100E+00	9,375E-05
KCNH1	-2,626E+00	1,721E-15	SIAH3	-1,749E+00	4,527E-07	ZNF711	-1,343E+00	8,704E-09	LOC101927418	-1,100E+00	8,006E-03
KCNE3	-2,618E+00	2,182E-32	FAM184B	-1,748E+00	2,746E-12	TMTC1	-1,343E+00	1,352E-08	DNMBP-AS1	-1,100E+00	7,213E-08
RPS3AP46	-2,615E+00	3,750E-07	HS6ST3	-1,747E+00	9,203E-06	MANSC1	-1,343E+00	6,454E-13	LOC105375095	-1,100E+00	2,334E-05
SERPINI1	-2,606E+00	2,073E-26	MGAT3-AS1	-1,746E+00	3,333E-05	MTARC1	-1,342E+00	4,287E-10	NYAPI	-1,099E+00	1,191E-05
METTL24	-2,606E+00	8,935E-15	GSTA3	-1,746E+00	1,665E-04	IGHV3-19	-1,342E+00	5,908E-04	PPP1R3G	-1,099E+00	3,819E-08
DMRT3	-2,601E+00	6,861E-16	LPA	-1,745E+00	8,625E-07	KLHL14	-1,341E+00	3,652E-06	OCA2	-1,098E+00	3,157E-03
UGT1A6	-2,596E+00	9,774E-14	DACHI	-1,744E+00	4,103E-14	LINC02042	-1,341E+00	5,446E-03	ANXA10	-1,098E+00	3,573E-07
LINC01101	-2,579E+00	3,083E-06	TSPAN18	-1,744E+00	6,738E-15	PCLO	-1,341E+00	3,080E-05	RIMS4	-1,097E+00	5,584E-05
TPTEP1	-2,575E+00	2,412E-16	SNORA80B	-1,736E+00	9,458E-08	SCML2	-1,339E+00	6,325E-10	TSPY26P	-1,097E+00	5,095E-09
DIRAS2	-2,561E+00	1,319E-09	ATP13A5	-1,735E+00	3,023E-07	WNT1	-1,338E+00	3,918E-05	PTPRVP	-1,097E+00	5,262E-05
STOX1	-2,559E+00	1,832E-20	TRPM8	-1,731E+00	3,956E-10	CHRM3	-1,338E+00	8,619E-05	SAMMSON	-1,097E+00	7,204E-03
STRA8	-2,546E+00	6,246E-04	SLAIN1	-1,729E+00	3,101E-14	LINC00642	-1,337E+00	4,699E-04	SPRR3	-1,097E+00	3,138E-03
HCAR1	-2,544E+00	9,929E-11	GABRB3	-1,728E+00	9,492E-06	ADD3-AS1	-1,337E+00	6,577E-11	CA15P1	-1,097E+00	3,010E-04
ADCY10	-2,541E+00	2,141E-18	LINC00365	-1,727E+00	6,171E-08	TRIM2	-1,334E+00	3,614E-12	RGS9BP	-1,096E+00	1,720E-08
HMX2	-2,539E+00	1,178E-06	LINC02570	-1,725E+00	1,553E-06	EPS8	-1,331E+00	7,218E-11	LYG2	-1,094E+00	1,311E-04
KCNV1	-2,527E+00	2,077E-06	MFSD6L	-1,722E+00	7,335E-08	PPARG	-1,330E+00	3,256E-08	TAS2R2P	-1,092E+00	1,402E-03
GPRC5D	-2,522E+00	1,791E-27	LMO3	-1,717E+00	3,321E-05	MMP27	-1,328E+00	3,698E-05	GALR2	-1,091E+00	2,829E-05
SLC30A3	-2,506E+00	1,031E-09	PLAAT3	-1,716E+00	5,134E-13	EFCAB8	-1,328E+00	3,491E-03	MAP3K19	-1,089E+00	1,069E-04
TREML3P	-2,498E+00	3,788E-06	OTX2-AS1	-1,714E+00	1,836E-04	SHISA6	-1,328E+00	4,443E-04	SPIRE2	-1,086E+00	1,026E-07
LOC729654	-2,491E+00	8,724E-10	LOC439933	-1,713E+00	6,100E-06	CHRDL1	-1,328E+00	2,873E-04	C19orf84	-1,086E+00	3,152E-04
AKR1C1	-2,489E+00	1,246E-13	ZDHH11	-1,711E+00	8,940E-11	BMP7	-1,328E+00	5,560E-09	MEGF10	-1,085E+00	2,473E-04
GSTM2	-2,486E+00	4,160E-21	CRHBP	-1,707E+00	2,736E-12	CACNB4	-1,327E+00	2,128E-06	EPHB6	-1,084E+00	4,745E-06
ARHGFE26-AS1	-2,485E+00	1,120E-12	CAVIN2	-1,706E+00	2,681E-16	AQP4-AS1	-1,327E+00	6,326E-04	WIPF3	-1,084E+00	9,987E-07
LOC101927136	-2,481E+00	7,474E-06	TDRD1	-1,704E+00	7,501E-05	SLC29A4	-1,325E+00	3,496E-08	FOXD3-AS1	-1,083E+00	4,966E-05
LINC02389	-2,478E+00	2,623E-06	PGM5P3-AS1	-1,703E+00	3,205E-04	PGPEP1L	-1,321E+00	4,483E-04	C10orf62	-1,082E+00	2,245E-03
GPR160	-2,433E+00	1,309E-34	OSGIN1	-1,703E+00	6,489E-12	PRSS37	-1,321E+00	2,185E-05	BRAFP1	-1,082E+00	1,074E-03
STOML3	-2,430E+00	7,739E-04	ERLNC1	-1,701E+00	4,391E-08	HNF1A	-1,318E+00	1,149E-04	DCSTAMP	-1,080E+00	1,089E-03
LINC02008	-2,429E+00	6,796E-06	ZNF280B	-1,700E+00	1,988E-09	FAM83E	-1,317E+00	4,797E-08	PARD3-AS1	-1,079E+00	9,081E-07
ADRA2B	-2,424E+00	1,186E-15	CCDC190	-1,700E+00	8,808E-05	LINC01010	-1,316E+00	1,830E-10	LOC654780	-1,078E+00	2,347E-07
NTRK2	-2,415E+00	6,870E-12	ANKRD34B	-1,694E+00	1,375E-08	PRSS3P4	-1,316E+00	1,874E-03	GPR26	-1,078E+00	2,785E-03
KCNJ9	-2,414E+00	4,180E-08	GJB1	-1,692E+00	7,108E-05	HYDIN	-1,314E+00	2,069E-06	RPL23AP92	-1,078E+00	9,596E-05
PRR36	-2,408E+00	1,048E-20	FGFR2	-1,692E+00	7,705E-27	SCNN1A	-1,314E+00	7,909E-10	SLC25A21-AS1	-1,078E+00	1,314E-06
TEX15	-2,408E+00	2,683E-06	KRT42P	-1,691E+00	2,223E-12	ACVR2B-AS1	-1,313E+00	1,685E-14	BCHE	-1,078E+00	1,061E-03
OR7C1	-2,407E+00	4,115E-10	KCNG3	-1,690E+00	3,295E-07	TBC1D26	-1,313E+00	4,482E-04	NLGN3	-1,078E+00	5,765E-10
SERTM1	-2,406E+00	6,014E-03	OR5M11	-1,689E+00	1,836E-05	MYH14	-1,312E+00	1,242E-08	SIX4	-1,078E+00	1,767E-08
GPC5	-2,405E+00	3,718E-09	RGMA	-1,688E+00	4,303E-14	FGF14	-1,312E+00	3,143E-07	KRTAP5-10	-1,077E+00	4,888E-04
B3GALT5	-2,404E+00	1,079E-10	REG1A	-1,686E+00	1,896E-03	CYP2F1	-1,312E+00	7,973E-06	RPL7P28	-1,075E+00	2,091E-03
PTX4	-2,397E+00	5,444E-07	C13orf42	-1,686E+00	4,704E-05	ACSS3	-1,311E+00	5,739E-03	C2orf15	-1,075E+00	1,638E-10

LINC02875	-2,396E+00	3,671E-28	ERVH48-1	-1,684E+00	1,292E-06	KSR2	-1,310E+00	2,913E-07	ZDHHC15	-1,074E+00	3,083E-05
OR7E158P	-2,391E+00	8,829E-06	ZDHHC2	-1,683E+00	2,159E-14	CYP2S1	-1,310E+00	3,457E-12	RADIL	-1,074E+00	4,423E-08
JCADP1	-2,390E+00	4,179E-08	CACNA2D3-AS1	-1,683E+00	5,749E-05	CCDC85A	-1,310E+00	1,665E-05	LINC00898	-1,073E+00	1,394E-04
SRGAP3-AS2	-2,388E+00	1,545E-04	IRS4	-1,680E+00	1,489E-09	LOC105371022	-1,310E+00	1,535E-06	LOC283387	-1,073E+00	1,202E-03
FETUB	-2,385E+00	1,017E-09	KRT36	-1,679E+00	1,034E-06	CCDC185	-1,309E+00	7,622E-05	SPATA46	-1,072E+00	2,923E-04
LRRC4	-2,385E+00	1,270E-16	LOC105375787	-1,678E+00	1,614E-05	GRM7-AS1	-1,308E+00	4,525E-03	LOC100128007	-1,071E+00	1,527E-06
LOC105374989	-2,384E+00	5,301E-04	ZNF492	-1,672E+00	1,111E-03	WDR38	-1,308E+00	1,265E-06	NPC1L1	-1,071E+00	1,862E-04
TBX20	-2,380E+00	1,805E-04	HOXD12	-1,665E+00	6,230E-09	MYL6B-AS1	-1,307E+00	2,084E-05	MGAT3	-1,070E+00	4,260E-05
SYNPR-AS1	-2,376E+00	3,489E-07	TNS1-AS1	-1,662E+00	3,155E-04	IGSF10	-1,307E+00	9,871E-05	QRICH2	-1,068E+00	3,915E-09
CTCF	-2,374E+00	7,839E-13	RNF157-AS1	-1,660E+00	7,382E-10	TPRG1-AS2	-1,306E+00	1,424E-05	LINC01932	-1,067E+00	1,300E-07
LINC00391	-2,373E+00	1,514E-07	CCSER1	-1,657E+00	4,982E-09	IAPP	-1,306E+00	1,378E-03	GCLM	-1,066E+00	1,826E-07
CELP	-2,366E+00	2,284E-12	ERN 2,00	-1,653E+00	5,164E-05	EDAR	-1,306E+00	6,762E-07	MYRF	-1,066E+00	1,178E-06
ESRG	-2,365E+00	7,460E-06	KCNK2	-1,653E+00	8,332E-09	B4GALT4-AS1	-1,305E+00	1,238E-05	DUX4L26	-1,065E+00	2,624E-03
TSSC2	-2,365E+00	4,639E-17	LOC100130691	-1,651E+00	6,400E-23	NEU4	-1,304E+00	3,883E-05	APOD	-1,064E+00	1,372E-04
SHISA7	-2,364E+00	4,341E-03	SLC6A3	-1,650E+00	2,180E-06	TLL10-AS1	-1,303E+00	3,890E-07	LINC01891	-1,064E+00	4,042E-03
DUXA	-2,362E+00	3,194E-06	HERC2P6	-1,647E+00	1,970E-04	GSR	-1,302E+00	3,212E-14	RNU4-42P	-1,064E+00	4,410E-03
HAPLN2	-2,358E+00	7,419E-07	IL25	-1,642E+00	7,088E-07	ACRP1	-1,302E+00	1,559E-03	CAPNS1P1	-1,063E+00	1,469E-04
SOX5	-2,357E+00	9,432E-23	LINC00161	-1,642E+00	1,738E-06	AGT	-1,301E+00	1,575E-05	FXVD6	-1,063E+00	1,428E-05
TRPM3	-2,356E+00	7,562E-10	DACT2	-1,641E+00	4,495E-07	HS3ST6	-1,300E+00	1,190E-03	GPR143	-1,063E+00	4,036E-05
LINC02561	-2,352E+00	1,106E-12	TCP11	-1,640E+00	5,955E-06	LINC00605	-1,299E+00	8,074E-03	PLEKHD1	-1,063E+00	4,402E-07
OTX2	-2,351E+00	1,183E-10	KLHL23	-1,636E+00	7,661E-17	FOXD3	-1,297E+00	2,337E-06	LINC01873	-1,061E+00	6,170E-05
JAKMIP2-AS1	-2,350E+00	8,829E-06	MAP7D2	-1,633E+00	8,975E-06	NKD2	-1,297E+00	1,866E-11	C6orf15	-1,061E+00	7,264E-03
ALKAL1	-2,347E+00	3,169E-07	XKR5	-1,630E+00	7,855E-08	LHX3	-1,296E+00	2,093E-03	MBNL1-AS1	-1,061E+00	4,949E-09
BRDT	-2,347E+00	1,923E-04	NAALADL2	-1,629E+00	1,998E-11	CPSF1P1	-1,295E+00	8,738E-05	USP6	-1,061E+00	6,807E-09
SOX21	-2,342E+00	1,897E-14	GPR83	-1,627E+00	8,102E-11	LINC01995	-1,295E+00	2,179E-03	CACNA2D1	-1,060E+00	2,315E-05
IGFALS	-2,340E+00	1,176E-06	PNCK	-1,627E+00	3,832E-07	CDIPTOSP	-1,294E+00	8,403E-05	CROCCP1	-1,060E+00	1,372E-03
TSGA10IP	-2,339E+00	2,186E-09	ZNF730	-1,626E+00	3,200E-07	SLC2A12	-1,294E+00	5,230E-09	LOC101927468	-1,059E+00	2,398E-05
HIF3A	-2,338E+00	1,844E-11	LOC100420746	-1,626E+00	6,553E-04	KLHDC9	-1,292E+00	1,459E-07	MAT1A	-1,059E+00	1,037E-03
CYP26C1	-2,338E+00	6,402E-10	MTND2P40	-1,624E+00	8,295E-05	CLU	-1,290E+00	2,240E-07	ERICH5	-1,059E+00	1,840E-03
SMNP	-2,329E+00	2,983E-04	GCGR	-1,624E+00	4,097E-10	RPRM	-1,289E+00	1,809E-04	LOC105374758	-1,058E+00	4,160E-04
KIRREL3-AS3	-2,327E+00	9,319E-03	DLX6	-1,624E+00	7,546E-10	LINC00381	-1,289E+00	1,139E-03	VN1R83P	-1,058E+00	1,735E-06
GSTM3	-2,316E+00	3,540E-17	LCN12	-1,623E+00	6,655E-04	LPO	-1,288E+00	4,086E-05	LOC105369203	-1,058E+00	1,967E-03
NPIP6	-2,315E+00	7,117E-15	PCARE	-1,623E+00	1,892E-08	RPL13AP17	-1,288E+00	1,644E-03	VTCN1	-1,057E+00	2,889E-03
C16orf90	-2,311E+00	1,767E-11	TAF4	-1,620E+00	2,329E-04	ABCG5	-1,287E+00	3,823E-06	LONRF1	-1,056E+00	2,660E-17
CYP2A6	-2,305E+00	6,847E-08	COL21A1	-1,618E+00	1,070E-06	HNF1A-AS1	-1,287E+00	8,649E-05	CNTNAP2	-1,056E+00	1,289E-03
CPA1	-2,304E+00	5,359E-05	PRKX-AS1	-1,616E+00	6,206E-06	STMN4	-1,286E+00	1,362E-04	SLC5A5	-1,055E+00	1,363E-03
ELAPOR1	-2,302E+00	3,280E-14	MAOB	-1,614E+00	1,063E-07	PIGR	-1,285E+00	5,395E-03	OLFM1	-1,053E+00	2,603E-04
AKR1C3	-2,298E+00	3,674E-14	KLRG2	-1,612E+00	7,418E-13	RPL23AP66	-1,284E+00	6,193E-05	KRTCAP3	-1,053E+00	5,324E-06
LOC112267871	-2,297E+00	5,631E-09	ADARB2-AS1	-1,610E+00	2,088E-04	LYPD6	-1,284E+00	5,246E-06	GBA3	-1,053E+00	1,851E-06
LINC00557	-2,295E+00	1,696E-07	DGKB	-1,608E+00	2,079E-06	MIR600HG	-1,284E+00	4,458E-13	KIAA2012	-1,052E+00	4,923E-04
SBK1	-2,295E+00	3,665E-19	SYNGR4	-1,607E+00	5,528E-06	CD177	-1,284E+00	6,968E-04	MS4A8	-1,052E+00	1,630E-03
CACNA1B	-2,286E+00	6,883E-10	HHLA2	-1,607E+00	7,222E-06	CHMP1B2P	-1,283E+00	3,559E-03	SLC4A4	-1,051E+00	1,317E-04
EPHA6	-2,286E+00	4,290E-06	IGSF5	-1,607E+00	8,531E-05	SCTR	-1,283E+00	1,978E-07	CCNO-DT	-1,051E+00	2,688E-09
TMEM211	-2,284E+00	4,674E-06	RNU6-911P	-1,604E+00	1,398E-06	TACR2	-1,282E+00	1,494E-15	DNALI1	-1,050E+00	7,195E-08

GDA	-2,284E+00	2,520E-08	NODAL	-1,604E+00	1,490E-04	ABHD1	-1,282E+00	3,277E-09	TRIM51BP	-1,050E+00	6,584E-03
THSD7B	-2,281E+00	1,766E-10	SLC4A11	-1,603E+00	2,714E-20	BSG-AS1	-1,281E+00	1,647E-07	KRT8P15	-1,049E+00	4,059E-04
NEFH	-2,279E+00	1,493E-07	SH3GL2	-1,600E+00	4,101E-06	CLGN	-1,279E+00	5,307E-05	SOX6	-1,049E+00	1,915E-06
WSCD2	-2,274E+00	1,021E-08	NUTM1	-1,598E+00	6,711E-06	HPD	-1,277E+00	3,202E-03	PCDHAC1	-1,047E+00	1,857E-03
BDNF	-2,273E+00	2,516E-23	FAM47A	-1,598E+00	2,935E-03	LOC100129734	-1,276E+00	1,538E-04	GUSBP2	-1,047E+00	3,036E-07
SLC1A2	-2,271E+00	1,325E-06	CFAP65	-1,598E+00	4,875E-05	IGKV1D-42	-1,276E+00	2,374E-03	ARHGEF38	-1,047E+00	1,243E-05
KRT40	-2,271E+00	3,737E-05	EVPLL	-1,598E+00	6,169E-08	PIR	-1,276E+00	3,863E-13	OR2A1	-1,046E+00	2,968E-03
FGF14-IT1	-2,270E+00	1,687E-07	MUC22	-1,597E+00	3,926E-04	KRTAP5-1	-1,275E+00	2,161E-06	LINC02688	-1,046E+00	9,667E-04
IGFL4	-2,269E+00	1,011E-12	TGM6	-1,595E+00	2,657E-04	PDE3B	-1,275E+00	1,283E-10	PKHD1L1	-1,046E+00	1,490E-04
AKR1C7P	-2,267E+00	3,200E-15	FOXP2	-1,592E+00	1,552E-10	ABO	-1,270E+00	2,665E-05	NOS1AP	-1,045E+00	5,744E-07
FOXI3	-2,267E+00	2,764E-07	LINC01502	-1,591E+00	1,612E-03	AXDND1	-1,270E+00	2,261E-05	SINHCAPP3	-1,045E+00	4,410E-03
CA8	-2,266E+00	1,369E-06	OR6L2P	-1,590E+00	2,936E-06	IGF1	-1,269E+00	4,147E-06	GNAZ	-1,045E+00	2,948E-05
ADH1C	-2,266E+00	1,176E-06	CHODL	-1,588E+00	4,408E-11	KCNH2	-1,269E+00	9,815E-06	EMID1	-1,044E+00	5,413E-05
MROH2A	-2,264E+00	1,294E-07	CELA3A	-1,588E+00	7,496E-03	MAPK10	-1,268E+00	7,336E-09	IGFBP2	-1,042E+00	3,621E-05
LUZP2	-2,262E+00	8,952E-12	TNFSF18	-1,588E+00	6,653E-07	ENPP7P1	-1,268E+00	7,338E-03	CST9LP2	-1,042E+00	1,327E-03
DLGAP1-ASS	-2,261E+00	4,123E-03	CHRM3-AS1	-1,587E+00	1,884E-05	MTCO2P16	-1,267E+00	1,597E-03	TSPAN33	-1,041E+00	9,695E-08
GRIK3	-2,260E+00	5,180E-03	SLC30A10	-1,585E+00	8,651E-06	RPL36P7	-1,266E+00	1,422E-05	PCSK4	-1,041E+00	9,771E-06
GRIK4	-2,252E+00	2,748E-22	KRT24	-1,584E+00	4,268E-05	SAMD5	-1,265E+00	1,246E-07	HGFAC	-1,039E+00	3,900E-04
SAMD12	-2,249E+00	2,676E-17	PIANP	-1,583E+00	1,275E-07	RBP7	-1,265E+00	1,737E-07	DPY19L2	-1,039E+00	6,545E-07
FAXC	-2,248E+00	1,042E-22	KCNJ16	-1,582E+00	5,751E-05	DISC1FP1	-1,264E+00	1,397E-03	DIAPH2	-1,039E+00	6,070E-10
FAM153CP	-2,238E+00	1,259E-09	EFCC1	-1,580E+00	7,539E-17	KIAA1549	-1,264E+00	3,387E-09	MAST1	-1,039E+00	4,318E-06
TMEM116	-2,221E+00	1,639E-30	POU3F3	-1,580E+00	1,471E-04	WDR17	-1,263E+00	4,947E-05	LINC01018	-1,039E+00	1,560E-03
LOC101060000	-2,221E+00	1,401E-03	ANKRD62	-1,579E+00	4,116E-08	TCF15	-1,262E+00	7,719E-05	SLC8A1-AS1	-1,039E+00	8,557E-06
DMRT2	-2,218E+00	9,216E-13	NEUROG1	-1,579E+00	2,996E-03	IQCN	-1,261E+00	4,323E-09	TLCD4	-1,038E+00	2,491E-05
NEXMIF	-2,217E+00	4,215E-10	TUSC8	-1,577E+00	5,459E-03	KRT33A	-1,259E+00	5,076E-05	PAX6	-1,038E+00	3,143E-07
ERICH3	-2,216E+00	5,758E-08	NPSR1-AS1	-1,575E+00	3,087E-08	DMBT1	-1,258E+00	3,149E-04	LOC101929657	-1,038E+00	6,167E-04
NXPE2	-2,210E+00	8,420E-06	POMC	-1,575E+00	6,877E-07	FAM182A	-1,257E+00	4,543E-04	ZNF536	-1,037E+00	7,319E-07
LMAN1L	-2,207E+00	4,356E-05	UNC5CL	-1,575E+00	1,643E-14	SPDYE2	-1,256E+00	1,027E-07	MBOAT1	-1,037E+00	5,988E-09
VWA5B2	-2,204E+00	2,184E-15	CNNM1	-1,573E+00	7,766E-06	HRG	-1,255E+00	2,122E-04	GASK1A	-1,035E+00	2,600E-05
LINC01783	-2,201E+00	1,577E-07	IGFBPL1	-1,572E+00	8,399E-06	WFIKKN1	-1,255E+00	9,882E-09	CACNA1D	-1,034E+00	1,220E-05
C12orf40	-2,189E+00	1,338E-04	LINC00974	-1,571E+00	3,514E-04	CACNA2D3	-1,253E+00	1,387E-10	RBM11	-1,033E+00	1,393E-03
MIR559	-2,187E+00	1,915E-07	CBS	-1,570E+00	9,879E-08	DIO1	-1,253E+00	2,328E-07	CLCN2	-1,033E+00	1,849E-16
NECAB2	-2,183E+00	4,109E-10	LINC00485	-1,569E+00	1,669E-04	RHBDL3	-1,250E+00	1,450E-05	LOC728877	-1,033E+00	3,685E-03
CHST9	-2,182E+00	2,888E-05	FAM81B	-1,568E+00	8,105E-05	XK	-1,249E+00	3,457E-07	NPM2	-1,032E+00	4,723E-05
CTTNBP2	-2,176E+00	1,592E-15	TCEAL2	-1,567E+00	7,108E-06	CFTR	-1,249E+00	9,649E-04	ROPN1B	-1,032E+00	7,843E-04
LHFPL3-AS2	-2,172E+00	7,462E-12	KMT5AP3	-1,566E+00	1,502E-03	HID1	-1,249E+00	9,862E-10	TMEM139	-1,031E+00	1,350E-03
UGT1A5	-2,159E+00	1,535E-03	C6orf223	-1,566E+00	2,366E-07	PRR15L	-1,249E+00	1,655E-05	MIR3065	-1,030E+00	4,326E-03
NKX2-3	-2,156E+00	4,300E-09	FAM86GP	-1,562E+00	1,085E-14	ANKRD66	-1,247E+00	2,583E-03	ADCY5	-1,030E+00	2,173E-04
TUBAP13	-2,155E+00	3,244E-07	SPAG6	-1,562E+00	1,610E-05	PDIA2	-1,247E+00	1,796E-03	ODAD1	-1,030E+00	2,002E-05
MSH1	-2,148E+00	1,090E-12	GALNT9	-1,561E+00	1,527E-07	CSRNP3	-1,246E+00	2,722E-09	SYDE2	-1,029E+00	6,095E-08
CNTNAP5	-2,144E+00	6,630E-07	LGALS4	-1,561E+00	7,541E-06	POU6F2	-1,246E+00	2,098E-03	IGHV1OR21-1	-1,029E+00	9,556E-03
TUBB7P	-2,140E+00	2,822E-04	TLX2	-1,559E+00	1,353E-07	CHST4	-1,245E+00	6,029E-04	PGDP1	-1,028E+00	1,178E-03
NKAIN2	-2,140E+00	1,237E-11	FGF13	-1,558E+00	1,849E-08	LRRC9	-1,242E+00	1,215E-05	PRKG2	-1,028E+00	2,210E-08
TMEM151B	-2,138E+00	2,007E-08	LINC02105	-1,556E+00	1,039E-03	LINC01694	-1,241E+00	6,007E-04	IGHA2	-1,026E+00	5,664E-03

SNTN	-2,137E+00	5,775E-05	NLRP11	-1,555E+00	6,545E-07	CCDC166	-1,239E+00	2,678E-03	HHEX	-1,025E+00	6,513E-10
CLDN8	-2,137E+00	5,181E-06	VWDE	-1,555E+00	1,015E-04	TRPV6	-1,238E+00	7,951E-05	CRYBB3	-1,024E+00	2,942E-06
SPATA21	-2,137E+00	3,987E-05	KIF1A	-1,554E+00	2,470E-04	ACSS1	-1,238E+00	1,982E-15	HPSE2	-1,024E+00	1,284E-03
LONRF2	-2,133E+00	3,368E-09	GCNT2P1	-1,552E+00	3,245E-05	LOC101927040	-1,235E+00	8,273E-05	RN7SL812P	-1,024E+00	1,691E-04
BNIP3P9	-2,132E+00	1,860E-03	SH2D6	-1,549E+00	3,688E-10	SMIM24	-1,235E+00	7,857E-05	STOX2	-1,023E+00	1,098E-08
SLC35G3	-2,132E+00	2,129E-06	ABCC5	-1,549E+00	7,760E-21	RPS2P32	-1,234E+00	1,734E-08	ATP2B1-AS1	-1,023E+00	1,163E-13
LINC01964	-2,132E+00	3,671E-07	LOC100129774	-1,546E+00	1,306E-03	FAM117B	-1,233E+00	1,457E-22	CHP2	-1,022E+00	3,904E-03
FOXJ1	-2,130E+00	4,968E-06	LEMD1-AS1	-1,546E+00	3,278E-06	PRR18	-1,233E+00	4,614E-03	MLC1	-1,021E+00	9,222E-05
LEMD1-DT	-2,130E+00	2,767E-10	SSTR5-AS1	-1,545E+00	7,703E-05	LINC00939	-1,233E+00	1,818E-03	NFE2L2	-1,021E+00	3,656E-17
CFAP92	-2,128E+00	1,701E-17	CELA3B	-1,544E+00	1,235E-03	TMEM35A	-1,232E+00	3,750E-05	AMN	-1,020E+00	1,266E-05
SIGLEC11	-2,128E+00	5,499E-04	IGLCOR22-2	-1,543E+00	5,113E-04	HHIP	-1,231E+00	4,366E-04	RAMP2-AS1	-1,020E+00	8,673E-06
CA7	-2,125E+00	1,260E-06	INMT-MINDY4	-1,542E+00	9,869E-03	WASIR2	-1,231E+00	1,139E-04	MYT1	-1,020E+00	7,845E-05
MUC5AC	-2,122E+00	1,632E-06	PTCH2	-1,542E+00	3,296E-14	CCNYL7	-1,229E+00	1,022E-05	MDGA1	-1,020E+00	2,792E-06
CEL	-2,117E+00	3,856E-08	LOC100507334	-1,541E+00	4,321E-07	MEDI12L	-1,229E+00	4,493E-11	KRT74	-1,019E+00	1,419E-06
PCDH19	-2,116E+00	3,484E-09	BRD9P2	-1,539E+00	5,861E-07	NCAPD2P1	-1,226E+00	3,962E-06	SYCP1	-1,017E+00	7,252E-03
MYOC	-2,115E+00	7,406E-11	TEX35	-1,539E+00	2,838E-04	PTGER1	-1,226E+00	6,096E-05	CPSF4L	-1,017E+00	1,323E-03
TUBB2B	-2,114E+00	6,093E-11	DIPK2A	-1,539E+00	5,123E-27	KRTAP5-9	-1,225E+00	3,108E-05	SNCB	-1,016E+00	1,663E-04
FLJ36000	-2,113E+00	1,240E-04	LINC00626	-1,538E+00	6,192E-04	C4orf17	-1,225E+00	3,095E-03	LOC100996404	-1,016E+00	9,995E-04
DPP10-AS1	-2,111E+00	4,978E-07	TGFBR3	-1,538E+00	1,992E-15	C5orf60	-1,225E+00	2,402E-06	BTBD18	-1,014E+00	6,154E-06
USH1C	-2,111E+00	3,460E-06	STK33	-1,538E+00	1,081E-07	C8orf74	-1,224E+00	1,327E-03	ILDR1	-1,014E+00	2,515E-06
LINC01258	-2,109E+00	8,544E-07	C2orf73	-1,536E+00	3,559E-06	SND1-IT1	-1,223E+00	6,944E-07	CECR2	-1,013E+00	8,904E-04
GCNT2	-2,098E+00	2,738E-17	DYNC1H1	-1,533E+00	1,078E-08	DMRTA2	-1,223E+00	6,109E-03	MAP1B	-1,012E+00	3,589E-06
COLEC10	-2,095E+00	4,871E-12	STXBPSL	-1,532E+00	9,583E-07	DEUP1	-1,222E+00	3,186E-04	CFAP46	-1,012E+00	4,024E-04
ULBP1	-2,094E+00	6,596E-16	RHOXF1-AS1	-1,531E+00	2,383E-07	FSTL4	-1,221E+00	9,094E-06	CIART	-1,012E+00	2,146E-06
FEZF1-AS1	-2,093E+00	9,408E-08	NPPC	-1,531E+00	2,168E-12	GNG13	-1,221E+00	5,461E-05	SV2A	-1,012E+00	8,624E-06
FREM2	-2,092E+00	4,210E-06	ILDR2	-1,530E+00	2,558E-10	TOX3	-1,220E+00	8,683E-03	NMNAT3	-1,012E+00	2,167E-07
SLC35G1	-2,091E+00	9,772E-35	LOC541473	-1,529E+00	7,388E-07	PAWRP2	-1,220E+00	6,678E-03	GNAT1	-1,010E+00	4,176E-04
LINC01644	-2,090E+00	3,238E-07	NTF3	-1,527E+00	4,116E-08	ABCC1	-1,219E+00	9,238E-16	CFAP221	-1,009E+00	7,266E-04
SLC25A48	-2,090E+00	3,674E-11	FRAS1	-1,525E+00	9,343E-10	PAX9	-1,218E+00	1,811E-08	POTEE	-1,009E+00	9,750E-05
SBSPON	-2,088E+00	1,643E-14	C22orf15	-1,524E+00	6,265E-10	PGAP4	-1,217E+00	3,061E-09	MCCC1	-1,009E+00	4,153E-20
LINC01574	-2,086E+00	2,971E-03	CACNA1I	-1,524E+00	6,485E-08	ACBD7	-1,217E+00	2,833E-06	TG	-1,009E+00	2,087E-05
COLCA1	-2,085E+00	1,775E-10	PRR15	-1,522E+00	3,222E-08	CILK1	-1,216E+00	2,174E-18	PTN	-1,009E+00	1,394E-04
GLI1	-2,080E+00	3,048E-18	SLITRK5	-1,522E+00	1,836E-05	ST6GALNAC1	-1,216E+00	2,312E-06	SALL2	-1,008E+00	1,817E-05
SCUBE3	-2,078E+00	1,362E-21	FGF10	-1,521E+00	3,766E-05	AADACL2	-1,215E+00	5,876E-03	KCNE1B	-1,008E+00	3,287E-03
AKR1C2	-2,068E+00	4,257E-11	FOXA1	-1,521E+00	2,262E-06	IGLV7-46	-1,215E+00	3,571E-06	KRT8P3	-1,007E+00	2,273E-04
ISM2	-2,068E+00	9,606E-12	LOC101928266	-1,521E+00	1,133E-07	GACAT2	-1,215E+00	5,716E-03	EFHB	-1,007E+00	6,876E-06
CNKSR2	-2,063E+00	4,259E-14	SLC5A11	-1,518E+00	7,754E-06	TTC28	-1,215E+00	2,557E-11	ENTPD3	-1,007E+00	2,094E-07
LOC101929823	-2,059E+00	4,341E-06	RXFP4	-1,517E+00	2,740E-05	PPP1R13B-DT	-1,213E+00	2,571E-05	GAPDHP51	-1,006E+00	4,426E-03
CYP2G1P	-2,055E+00	6,427E-08	HPN-AS1	-1,517E+00	1,031E-04	CASKIN1	-1,213E+00	8,516E-06	ICAIL	-1,006E+00	6,774E-10
GLS2	-2,051E+00	3,208E-14	DPYSL5	-1,516E+00	4,157E-05	HTR3D	-1,213E+00	5,410E-03	COL6A5	-1,005E+00	2,498E-03
TMEM212	-2,051E+00	2,451E-04	IGFLIP2	-1,516E+00	2,503E-03	DMTN	-1,213E+00	2,808E-10	EFCAB10	-1,004E+00	1,102E-05
IGKV6-21	-2,049E+00	8,621E-03	SLC5A2	-1,515E+00	1,130E-12	STXBP6	-1,212E+00	3,804E-05	HOXB3	-1,004E+00	6,029E-05
SYTL5	-2,046E+00	3,349E-11	OR52L1	-1,515E+00	9,154E-03	TARID	-1,210E+00	7,736E-16	LOC100996442	-1,004E+00	7,645E-03
C1orf87	-2,045E+00	4,304E-04	SNX31	-1,515E+00	4,984E-12	HAP1	-1,210E+00	2,287E-07	FGFBP3	-1,003E+00	7,201E-08

PCDHB17P	-2,044E+00	1,721E-07	ODC1	-1,515E+00	2,044E-12	ODAD3	-1,208E+00	2,983E-06	VAX2	-1,003E+00	7,426E-05
LGR6	-2,043E+00	1,194E-13	ISLR2	-1,513E+00	6,931E-09	HEY1	-1,208E+00	6,698E-07	SCARA3	-1,003E+00	3,133E-07
RNA5SP442	-2,033E+00	5,537E-05	LOC100129434	-1,509E+00	1,708E-07	ESPNL	-1,207E+00	7,797E-05	LOC642484	-1,003E+00	3,266E-04
SLC27A2	-2,029E+00	3,080E-15	CP	-1,509E+00	3,163E-05	TLL10	-1,206E+00	4,231E-08	LTBP4	-1,003E+00	1,307E-14
HLF	-2,029E+00	4,571E-15	KRT19	-1,509E+00	1,144E-04	C1orf167	-1,206E+00	9,794E-06	GSTA4	-1,003E+00	3,573E-07
CES1	-2,026E+00	4,671E-06	TMPRSS11A	-1,508E+00	2,455E-05	EFNA2	-1,205E+00	2,557E-05	ATP13A4	-1,003E+00	1,608E-03
SLC4A8	-2,024E+00	2,816E-18	CACNG5	-1,508E+00	1,703E-03	DGKG	-1,204E+00	7,262E-07	TMEM158	-1,002E+00	9,408E-08
EPCAM	-2,024E+00	7,719E-19	TMEM163	-1,507E+00	9,901E-12	CLIC6	-1,204E+00	1,133E-06	LRRRC73	-1,001E+00	7,772E-07
BHMT	-2,021E+00	3,073E-09	TMEM270	-1,505E+00	4,937E-05	PHYHIPL	-1,204E+00	9,511E-07	IGHG4	-1,001E+00	5,116E-03

Symbol	Log2FoldChange	Padj	Symbol	Log2FoldChange	Padj	Symbol	Log2FoldChange	Padj	Symbol	Log2FoldChange	Padj
ZNF415P1	1,002E+00	3,456E-03	COBLP1	1,139E+00	7,853E-03	FHL2	1,333E+00	1,599E-25	RP1	1,723E+00	2,263E-08
LOC105371855	1,002E+00	9,852E-04	PADI2	1,139E+00	2,766E-04	LY6D	1,333E+00	1,429E-05	LOC100421664	1,724E+00	1,725E-03
PSORS1C1	1,002E+00	2,202E-07	FBLIM1	1,142E+00	5,553E-18	ANGPTL4	1,335E+00	5,792E-10	ANKRD1	1,732E+00	2,536E-07
LINC00520	1,002E+00	6,607E-03	MYBPH	1,142E+00	1,597E-05	LAMA3	1,336E+00	2,293E-09	RN7SKP11	1,736E+00	9,556E-03
SERPINE2	1,002E+00	2,073E-05	SERINC2	1,142E+00	2,317E-14	DUSP6	1,338E+00	6,954E-18	CSPG4	1,738E+00	1,298E-14
IRX1	1,005E+00	7,112E-03	SLC22A1	1,142E+00	1,026E-04	TAS2R62P	1,338E+00	2,341E-03	WNT6	1,738E+00	5,919E-03
PODNL1	1,005E+00	2,598E-08	TYMP	1,145E+00	8,666E-11	SMTNL2	1,340E+00	2,471E-03	TCHHL1	1,738E+00	1,726E-06
STRIP2	1,005E+00	1,471E-09	LINC02273	1,152E+00	1,040E-05	RGS20	1,342E+00	5,403E-11	NTSR1	1,741E+00	2,569E-14
AGPAT4	1,007E+00	7,261E-13	IFIT2	1,153E+00	2,844E-07	PAQR5	1,344E+00	1,203E-09	CITED4	1,743E+00	1,993E-14
ARHGAP29	1,007E+00	1,314E-08	CCNJL	1,156E+00	1,054E-07	FABP5	1,346E+00	1,872E-07	FAM83A-AS1	1,745E+00	1,755E-11
FRMD5	1,008E+00	2,920E-09	APLN	1,157E+00	2,312E-09	LINC02015	1,346E+00	1,453E-04	SERPINB7	1,747E+00	6,272E-10
KRT18P3	1,009E+00	2,081E-03	PDLIM7	1,158E+00	9,666E-19	CHRND	1,347E+00	9,154E-03	AGBL1-AS1	1,747E+00	5,384E-09
CXCR1	1,010E+00	3,756E-04	USP2	1,158E+00	1,377E-06	FXYD5	1,349E+00	6,223E-16	CST6	1,755E+00	6,036E-10
CARD18	1,011E+00	3,117E-03	PDCD1LG2	1,160E+00	3,598E-08	H3C14	1,355E+00	5,331E-03	SERPINA4	1,756E+00	2,203E-04
LINC01068	1,012E+00	4,107E-04	SERPIND1	1,160E+00	5,556E-03	ALPP	1,360E+00	1,991E-06	SFRP1	1,759E+00	2,470E-08
DPYD-AS1	1,012E+00	4,905E-04	SMIM2-IT1	1,160E+00	7,595E-03	BIRC7	1,366E+00	2,816E-06	TGFBI	1,760E+00	4,447E-15
HDAC9	1,012E+00	1,810E-08	THSD1	1,161E+00	3,765E-11	FABP12	1,367E+00	2,972E-04	TNPO1P2	1,775E+00	2,937E-04
NMNAT2	1,015E+00	2,850E-06	C11orf91	1,165E+00	7,966E-08	ADRB2	1,370E+00	1,639E-09	LINC02154	1,779E+00	2,373E-09
ZBTB32	1,015E+00	1,573E-04	FOLR3	1,165E+00	1,244E-08	EVA1A	1,371E+00	1,735E-09	CFAP300	1,783E+00	1,389E-03
KRT18P22	1,015E+00	2,798E-04	ANTXR2	1,166E+00	5,337E-11	PSAPL1	1,372E+00	1,524E-04	LINC02777	1,797E+00	6,405E-18
FAM214B	1,016E+00	1,493E-11	SLC28A3	1,169E+00	1,139E-07	CFAP58	1,373E+00	8,804E-07	F3	1,799E+00	3,051E-15
LRAT	1,017E+00	5,990E-03	DKK1	1,171E+00	8,820E-05	SLC6A17	1,377E+00	1,528E-06	WNT7A	1,809E+00	2,123E-12
OASL	1,017E+00	5,807E-05	ASIC2	1,172E+00	2,798E-03	CHRNA9	1,379E+00	5,634E-07	TLL1	1,810E+00	3,165E-12
CAMK2N1	1,017E+00	1,365E-07	GRIA3	1,175E+00	1,236E-05	SORCS2	1,381E+00	7,668E-10	SAA1	1,815E+00	8,323E-10
LY6E	1,017E+00	6,863E-08	LINC01479	1,177E+00	1,899E-03	KRT73	1,381E+00	1,663E-04	MIR31	1,819E+00	1,401E-03
IL20RB	1,020E+00	6,724E-07	LTBP1	1,178E+00	2,168E-10	KCNAB1-AS2	1,381E+00	2,900E-03	LHX1	1,827E+00	1,437E-08
NPY4R	1,020E+00	4,499E-04	TMEM229B	1,179E+00	8,577E-08	MYBPHL	1,383E+00	7,554E-04	S100A7	1,829E+00	8,238E-08
LMOD3	1,021E+00	1,291E-03	SAA2	1,182E+00	1,020E-04	PCDH7	1,383E+00	5,027E-09	VEGFC	1,834E+00	6,771E-16
S100P	1,022E+00	1,063E-03	PPY	1,184E+00	5,290E-04	CHRNG	1,387E+00	3,818E-04	LINC00707	1,843E+00	1,244E-08
CCL8	1,023E+00	5,796E-05	LINC02721	1,185E+00	2,016E-03	NLRP3	1,388E+00	6,414E-08	PSG9	1,858E+00	3,347E-08
SIK1B	1,024E+00	1,635E-07	LINC01117	1,186E+00	8,630E-09	SLC6A2	1,389E+00	2,831E-06	IL24	1,859E+00	1,050E-10
SRPX	1,024E+00	7,842E-06	GALNT6	1,186E+00	3,259E-10	ITGB6	1,392E+00	9,728E-13	LINC02454	1,861E+00	8,853E-14
ZG16	1,025E+00	4,946E-03	UTS2	1,186E+00	4,704E-04	ZDHHC20P1	1,392E+00	3,860E-06	EREG	1,865E+00	3,327E-08

PRMT8	1,025E+00	1,093E-05	RET	1,186E+00	1,546E-06	C6orf141	1,394E+00	1,214E-10	CT62	1,867E+00	3,048E-09
HOXC13-AS	1,026E+00	1,546E-05	CLMP	1,187E+00	6,800E-10	APOBEC2	1,394E+00	1,277E-03	SI00A7A	1,870E+00	7,725E-07
GZMA	1,027E+00	6,849E-05	ROPN1	1,188E+00	3,181E-06	COL22A1	1,394E+00	2,925E-05	MOGAT2	1,873E+00	4,189E-06
CD24	1,028E+00	1,010E-03	LOC100507516	1,188E+00	5,425E-06	CXCL2	1,395E+00	1,490E-08	CXCL10	1,880E+00	9,796E-09
TNFAIP6	1,028E+00	5,314E-03	MOB3B	1,189E+00	1,039E-11	SFTA1P	1,398E+00	1,291E-05	MYL3	1,887E+00	1,035E-05
NAGS	1,029E+00	2,103E-11	LINC01287	1,190E+00	6,715E-07	TMEM92-AS1	1,399E+00	9,680E-13	KRT16P5	1,890E+00	2,068E-11
SHANK2	1,029E+00	2,314E-04	ADRA1B	1,191E+00	4,688E-11	CSF2	1,401E+00	1,500E-05	SAA2-SAA4	1,898E+00	3,712E-09
CA2	1,029E+00	3,106E-05	RAET1G	1,191E+00	9,218E-07	FOXC2-AS1	1,401E+00	2,789E-04	RASD2	1,920E+00	1,658E-13
MATN1	1,030E+00	2,270E-08	CXCL1	1,192E+00	1,578E-05	CH3L2	1,402E+00	9,950E-08	RPSAP52	1,927E+00	4,803E-12
AQP3	1,031E+00	3,003E-04	CCDC68	1,193E+00	2,774E-05	MMP1	1,407E+00	3,617E-06	LINC01213	1,935E+00	2,405E-09
MT1M	1,031E+00	1,430E-05	MYPN	1,194E+00	3,045E-04	HS1BP3-IT1	1,409E+00	1,799E-07	C9	1,936E+00	6,700E-13
LOC284344	1,031E+00	3,136E-03	NGF	1,194E+00	6,136E-06	LOC101929470	1,412E+00	6,427E-08	IFNA1	1,942E+00	3,841E-07
C10orf67	1,033E+00	1,631E-05	FYB1	1,194E+00	6,099E-08	IRGM	1,412E+00	1,385E-06	LINC02029	1,946E+00	5,631E-11
FABP6	1,034E+00	5,940E-03	COL17A1	1,204E+00	2,474E-07	PPP4R4	1,416E+00	8,848E-09	CCBE1	1,952E+00	6,033E-13
RHOD	1,034E+00	1,750E-07	SERPINA5	1,206E+00	5,903E-05	MYOSLID	1,416E+00	2,587E-11	CYP4Z2P	1,962E+00	1,732E-04
LOC117779438	1,034E+00	1,396E-03	MIR193BHG	1,206E+00	2,031E-10	TNPO1P3	1,417E+00	5,440E-07	PNLIPRP3	1,963E+00	2,757E-07
RNU4-2	1,035E+00	4,624E-03	SERPINA3	1,206E+00	3,231E-04	INPP4B	1,417E+00	2,239E-14	KRT75	1,966E+00	1,491E-08
PGLYRP4	1,035E+00	1,378E-04	CAV1	1,208E+00	6,181E-11	HNRNPA1P33	1,420E+00	1,185E-06	LINC01629	1,969E+00	1,350E-06
TNNI1	1,036E+00	1,039E-07	LOC100418723	1,210E+00	4,180E-08	INHBA-AS1	1,422E+00	2,771E-09	LINC00460	1,977E+00	1,884E-11
PDPN	1,036E+00	3,051E-09	SH2D2A	1,211E+00	1,191E-12	TINAG	1,425E+00	1,872E-04	LINC01356	1,994E+00	2,312E-14
LHB	1,036E+00	1,194E-05	GTSF1	1,212E+00	1,060E-03	ADGRE1	1,429E+00	3,951E-07	HYAL4	2,006E+00	3,987E-10
MUCL1	1,037E+00	5,795E-06	BASP1	1,212E+00	2,305E-10	STARD13-AS	1,429E+00	3,597E-10	LHX1-DT	2,013E+00	1,145E-10
RTTN	1,038E+00	6,089E-11	NFE2	1,215E+00	2,299E-07	LINC02535	1,430E+00	8,722E-08	PCDHGC5	2,029E+00	5,457E-10
LINC01711	1,039E+00	1,397E-04	OR10V3P	1,215E+00	1,147E-03	SMOC1	1,432E+00	1,009E-14	NTNG1	2,030E+00	1,823E-05
FJX1	1,039E+00	1,603E-12	ATPIB4	1,215E+00	5,584E-05	CLEC2B	1,438E+00	6,082E-12	MMP28	2,041E+00	4,528E-19
FOSL1	1,041E+00	2,155E-08	P2RX6	1,216E+00	1,556E-06	GPR78	1,438E+00	2,588E-07	ELFN2	2,042E+00	1,416E-09
LINC01679	1,041E+00	9,554E-11	KRT16P4	1,218E+00	2,726E-06	RTP3	1,442E+00	6,971E-04	FOXC2	2,064E+00	7,736E-16
PPP1R27	1,042E+00	1,224E-05	CSRP3	1,218E+00	5,934E-06	BATF2	1,448E+00	1,766E-07	ARL14	2,064E+00	1,136E-10
GSDMA	1,049E+00	5,204E-04	SH3TC2	1,219E+00	4,109E-10	GPRACR	1,448E+00	2,989E-07	CNGB1	2,069E+00	3,356E-12
LINC01450	1,050E+00	1,908E-03	KRT14	1,220E+00	3,269E-06	NKILA	1,449E+00	3,137E-12	PSG4	2,072E+00	7,100E-10
LINC01920	1,050E+00	3,471E-03	LOC101929460	1,222E+00	1,069E-04	PLA2G4E-AS1	1,452E+00	8,263E-07	ALB	2,092E+00	9,450E-03
HLA-V	1,051E+00	6,087E-06	NPY4R2	1,226E+00	7,555E-05	LINC01555	1,457E+00	2,348E-04	LINC00862	2,114E+00	5,082E-04
CAV3	1,053E+00	8,109E-04	APCDD1L-DT	1,226E+00	2,084E-05	APCDD1L	1,458E+00	4,752E-08	SLC22A2	2,120E+00	9,593E-04
TPT1P5	1,053E+00	2,641E-06	IDO1	1,227E+00	1,011E-04	RYR3-DT	1,464E+00	3,565E-05	PLA2G2F	2,125E+00	7,856E-18
RPL9P16	1,055E+00	6,220E-03	NLRP10	1,228E+00	5,364E-04	KRT1	1,466E+00	2,809E-07	LINC01776	2,140E+00	6,216E-05
CPA4	1,055E+00	4,542E-05	MYBPC2	1,228E+00	1,593E-07	AREG	1,468E+00	6,975E-09	OR10Y1P	2,143E+00	8,279E-09
SENCR	1,055E+00	2,805E-07	GRM4	1,233E+00	1,532E-05	LINC00052	1,468E+00	4,642E-13	LINC00452	2,161E+00	4,589E-06
CYP2U1-AS1	1,055E+00	7,530E-05	TNNT1	1,234E+00	1,398E-06	LINC01704	1,469E+00	3,614E-07	CAPN6	2,186E+00	8,719E-03
GS1-204I12.4	1,056E+00	4,563E-04	GPR68	1,235E+00	2,309E-13	LOC574538	1,470E+00	5,759E-07	NOX5	2,188E+00	5,570E-21
SYPL2	1,059E+00	2,409E-03	ADAM21P1	1,235E+00	1,447E-03	LINC00636	1,474E+00	6,239E-07	AJAP1	2,193E+00	6,494E-13
RDH16	1,060E+00	3,533E-06	VNN3	1,237E+00	2,445E-05	LOC105376116	1,478E+00	3,119E-03	ACTBL2	2,220E+00	3,378E-14
ZPLD1	1,061E+00	1,663E-04	LINC01907	1,241E+00	4,674E-06	LINC00601	1,479E+00	3,128E-05	LINC00911	2,225E+00	7,860E-10
PICART1	1,061E+00	1,384E-10	NT5E	1,241E+00	3,996E-08	LINC02643	1,480E+00	8,763E-04	MMP10	2,230E+00	1,456E-10
LINC00943	1,066E+00	6,786E-07	LOC102724434	1,242E+00	2,976E-04	L1CAM	1,488E+00	3,949E-14	LINC01910	2,255E+00	3,532E-15



TREM1	1,069E+00	1,431E-05	PMEPA1	1,247E+00	5,518E-13	LOC105369187	1,488E+00	4,373E-10	DISP3	2,275E+00	1,067E-11
FAM87A	1,069E+00	2,950E-07	PLK2	1,248E+00	8,577E-15	C17orf50	1,492E+00	3,708E-06	MT2A	2,279E+00	2,923E-23
TNS4	1,070E+00	2,017E-11	LGALS9C	1,249E+00	2,071E-05	HEPHL1	1,495E+00	3,334E-06	IL13RA2	2,288E+00	1,856E-09
FABP5P3	1,070E+00	1,961E-05	LGALS9B	1,251E+00	5,481E-06	CD5L	1,496E+00	3,150E-03	HRAT17	2,316E+00	4,688E-04
SULF2	1,070E+00	7,240E-10	KCNJ12	1,251E+00	1,133E-06	ITGA3	1,497E+00	3,478E-22	FDCSP	2,325E+00	9,796E-04
TENM3	1,071E+00	1,963E-06	RNU4-1	1,252E+00	6,675E-04	BCAR3	1,498E+00	3,687E-20	SERPINA9	2,327E+00	3,845E-06
ATP2C2-AS1	1,071E+00	1,604E-08	TMEM239	1,252E+00	7,288E-04	MMP13	1,500E+00	1,094E-04	RN7SL356P	2,330E+00	3,273E-06
SERPINA6	1,071E+00	2,521E-04	SLC16A2	1,252E+00	6,493E-10	TEX29	1,505E+00	2,824E-09	LOC101927661	2,334E+00	1,613E-11
GJB6	1,072E+00	1,199E-04	TINAGL1	1,253E+00	1,243E-11	MANCR	1,507E+00	1,807E-08	KISS1	2,335E+00	1,754E-04
LINC02716	1,072E+00	5,747E-05	CDH3	1,253E+00	1,298E-20	MYL4	1,511E+00	4,655E-09	C20orf141	2,343E+00	3,674E-11
CASQ2	1,074E+00	5,412E-03	CLEC2A	1,253E+00	7,125E-04	CSF3	1,514E+00	3,187E-04	OR2L6P	2,346E+00	2,674E-04
CDA	1,074E+00	3,262E-05	KRT16	1,261E+00	3,727E-05	CAVIN4	1,515E+00	7,247E-06	MYH8	2,358E+00	6,607E-03
DDIT4L	1,075E+00	1,837E-04	SFN	1,264E+00	7,684E-09	RPS16P2	1,520E+00	2,805E-07	PSG1	2,416E+00	1,347E-08
COL16A1	1,080E+00	1,989E-11	CASP14	1,266E+00	1,000E-08	SH2D5	1,524E+00	1,231E-09	ZFP42	2,425E+00	8,168E-09
GJB2	1,081E+00	1,284E-05	AMIGO2	1,266E+00	5,123E-10	ADTRP	1,525E+00	4,230E-13	SLC22A3	2,445E+00	1,525E-20
LOC105377448	1,082E+00	2,765E-03	SYT7	1,267E+00	1,325E-08	LINC02577	1,527E+00	6,443E-09	XKR7	2,446E+00	1,635E-05
CYP27C1	1,082E+00	1,333E-06	MT1L	1,268E+00	3,747E-06	PIC SAR	1,529E+00	1,674E-05	ACTBP12	2,447E+00	3,069E-09
DUSP4	1,085E+00	5,474E-08	SCG5	1,270E+00	1,831E-08	LOC112267895	1,532E+00	9,920E-05	LNCOG	2,489E+00	4,828E-18
ADAM8	1,086E+00	2,058E-09	NHLH2	1,274E+00	1,884E-04	PSG7	1,535E+00	5,830E-04	LINC02251	2,496E+00	8,191E-09
TMSB10	1,086E+00	5,180E-08	KANK4	1,274E+00	6,327E-05	GLP2R	1,535E+00	4,711E-05	CLEC4OP	2,498E+00	6,700E-12
MAP3K7CL	1,088E+00	2,240E-10	NWD2	1,276E+00	1,662E-04	KRT16P3	1,535E+00	5,088E-08	ANXA13	2,506E+00	1,267E-17
RN7SL862P	1,088E+00	6,329E-03	CKM	1,278E+00	8,944E-03	SEC14L2	1,535E+00	1,379E-18	KRT16P6	2,510E+00	7,637E-13
CCL5	1,089E+00	7,843E-06	EQTN	1,278E+00	2,399E-06	NIPAL4	1,538E+00	8,892E-11	GYG2P1	2,517E+00	1,671E-06
HTR1D	1,089E+00	1,155E-04	OR2W3	1,278E+00	2,742E-03	IFNE	1,540E+00	1,631E-08	LINC01043	2,530E+00	1,178E-03
SEC61G	1,090E+00	2,180E-06	IL33	1,278E+00	8,907E-06	SERPINE1	1,543E+00	1,042E-12	CXCL11	2,548E+00	1,394E-12
FOXL1	1,090E+00	1,841E-06	MYH3	1,283E+00	2,066E-04	SLC6A14	1,546E+00	5,512E-06	DSCAM	2,565E+00	2,507E-14
KRT6C	1,096E+00	2,473E-03	MYOM3	1,285E+00	6,629E-06	LINC02873	1,549E+00	2,312E-09	LINC01537	2,581E+00	6,810E-10
RUNDC3A	1,096E+00	4,255E-06	LAMC2	1,285E+00	6,612E-08	ART3	1,553E+00	1,120E-12	KRTAP11-1	2,602E+00	3,478E-04
RPL21P40	1,097E+00	7,603E-03	SLC38A5	1,286E+00	1,131E-08	XDH	1,562E+00	2,183E-13	FGF5	2,608E+00	3,829E-11
EPSTH1	1,097E+00	1,771E-08	TNFRSF12A	1,287E+00	8,115E-22	MGAM	1,563E+00	3,281E-03	ROSI	2,639E+00	2,778E-12
LINC01940	1,098E+00	1,040E-03	FIGN	1,290E+00	1,168E-07	MYO7B	1,563E+00	1,853E-15	PAGE3	2,649E+00	8,822E-03
LIPG	1,101E+00	2,133E-05	RNU2-63P	1,291E+00	8,033E-04	FLRT3	1,569E+00	1,445E-07	MTCO3P12	2,683E+00	3,363E-03
GOLGA7B	1,103E+00	7,772E-07	DCBLD1	1,291E+00	2,198E-20	GLDC	1,570E+00	7,438E-06	PSG6	2,708E+00	3,850E-05
SERPINA10	1,103E+00	1,659E-04	FLRT2	1,293E+00	8,097E-09	GNLY	1,574E+00	3,435E-09	ARSFP1	2,723E+00	7,181E-03
ORM2	1,104E+00	2,123E-03	ARSI	1,294E+00	6,821E-09	FAM83A	1,582E+00	3,080E-11	PPP1R1A	2,756E+00	5,899E-06
LOC105370616	1,105E+00	3,558E-03	LINC00705	1,297E+00	4,895E-05	SHISAL1	1,589E+00	1,088E-10	IL20	2,783E+00	3,029E-18
IFIT3	1,106E+00	1,288E-06	BMP2	1,298E+00	1,645E-10	TFDP3	1,598E+00	5,442E-04	PCDH8	2,785E+00	7,543E-03
RNA5SP160	1,107E+00	9,382E-03	SERPINB4	1,298E+00	7,938E-05	JPH2	1,602E+00	2,932E-08	CGB8	2,811E+00	2,750E-11
GBP5	1,108E+00	8,493E-05	CXCL3	1,299E+00	1,159E-06	MIR31HG	1,603E+00	7,006E-10	LINC01714	2,850E+00	7,731E-04
IGSF23	1,108E+00	8,301E-04	SPTLC3	1,301E+00	1,537E-08	INHBA	1,609E+00	5,835E-12	SPANXB1	2,871E+00	1,441E-04
ACTN1	1,112E+00	1,873E-15	MYMK	1,306E+00	7,047E-03	SH3TC2-DT	1,612E+00	1,002E-07	LINC02167	2,935E+00	9,354E-05
DRD5	1,113E+00	8,196E-04	SPOCD1	1,307E+00	3,784E-11	ADAMTS15	1,612E+00	4,109E-12	TRHDE	2,941E+00	1,203E-12
ANXA3	1,113E+00	1,195E-09	LINC01615	1,309E+00	9,879E-08	FGL1	1,614E+00	5,993E-06	LINC01998	2,974E+00	5,441E-17
ACTA1	1,114E+00	6,766E-03	IL12RB2	1,309E+00	3,296E-07	CTSE	1,617E+00	2,456E-05	KRTAP2-3	3,024E+00	5,876E-09

GBP1	1,115E+00	1,368E-07	MAP3K20-AS1	1,313E+00	5,713E-12	DISP2	1,619E+00	3,608E-10	LOC101928228	3,083E+00	5,384E-17
LOC101929532	1,116E+00	2,026E-06	KCNF1	1,314E+00	1,152E-07	FAM178B	1,639E+00	2,069E-05	CES5A	3,104E+00	2,532E-07
CASP5	1,117E+00	2,590E-06	MARCHF4	1,316E+00	1,868E-05	ELAVL2	1,644E+00	3,131E-07	PAEP	3,109E+00	1,269E-22
ASB5	1,118E+00	1,627E-04	KHSRPP1	1,316E+00	1,177E-06	SCAT1	1,652E+00	2,309E-14	RPEP6	3,193E+00	3,620E-05
THBS1	1,120E+00	6,033E-07	CSPG4P13	1,317E+00	2,474E-06	NOG	1,662E+00	3,879E-06	DHRS2	3,226E+00	2,558E-10
PRDM8	1,120E+00	3,872E-11	ODAPH	1,317E+00	1,529E-06	USP17L7	1,664E+00	8,964E-03	TH	3,271E+00	5,484E-19
LINC00313	1,121E+00	4,010E-05	IFIT1B	1,318E+00	9,649E-06	MT1E	1,667E+00	4,292E-11	OR2L13	3,301E+00	2,361E-09
HLA-J	1,123E+00	1,132E-05	TSPAN2	1,318E+00	2,916E-08	KRT19P2	1,676E+00	6,142E-07	TRIML2	3,328E+00	3,184E-23
LGALS1	1,123E+00	1,038E-09	LINC02910	1,318E+00	1,242E-04	KIF12	1,685E+00	9,527E-10	AADA3L3	3,357E+00	5,153E-05
RIMS3	1,124E+00	6,702E-07	LINC00165	1,320E+00	4,687E-03	AGBL1	1,688E+00	3,913E-03	PSG5	3,407E+00	2,635E-12
STEAP4	1,128E+00	7,347E-05	HJV	1,321E+00	5,643E-04	SLC6A11	1,690E+00	2,027E-09	OR2L2	3,433E+00	3,403E-11
MIR6772	1,129E+00	1,021E-04	PLEK2	1,321E+00	1,594E-14	MYH6	1,692E+00	5,103E-03	LINC01179	3,442E+00	1,109E-19
CYP26B1	1,131E+00	1,132E-07	RASGRF1	1,321E+00	2,706E-08	CYP27B1	1,696E+00	1,425E-15	LINC02383	3,571E+00	3,685E-04
DCT	1,131E+00	2,281E-03	MYF6	1,323E+00	2,644E-03	TMEM92	1,716E+00	7,785E-19	CCK	3,617E+00	9,211E-13
ARSF	1,132E+00	4,714E-03	DPYD-AS2	1,324E+00	3,282E-03	USP17L1	1,717E+00	2,329E-03	SCGB2A1	3,630E+00	1,029E-03
SAA4	1,133E+00	7,388E-04	MMP3	1,327E+00	1,812E-05	CWH43	1,720E+00	6,478E-06	F11	3,758E+00	6,045E-03
KRT6B	1,133E+00	1,889E-04	IL1RL1	1,327E+00	6,183E-06	SYT16	1,722E+00	2,659E-08	CD300LD-AS1	3,760E+00	1,079E-11
FRMPD4	1,134E+00	4,894E-04	NCF4-AS1	1,328E+00	7,006E-05	LINC02872	1,722E+00	1,118E-09	CGB5	3,831E+00	3,198E-08
CEBPB	1,136E+00	1,360E-21	ISG15	1,330E+00	1,809E-06	DIO3OS	1,723E+00	1,552E-10	KRT81	4,227E+00	4,200E-06
RNY3P16	1,137E+00	1,607E-05	NTRK1	1,330E+00	3,008E-04	CSMD3	1,723E+00	5,147E-06	FABP4	4,748E+00	1,220E-19
LINC01990	1,138E+00	7,696E-05	FHL2	1,333E+00	1,599E-25	IFNWP19	1,723E+00	1,425E-10	IFNK	5,219E+00	1,520E-17

Supplementary Table 3. Co-expressed Genes by Inverse SOX2/SOX9 Expression in TCGA-HNSCC Cohort

Correlated Gene	Spearman's Correlation SOX2	Spearman's Correlation SOX9	Correlated Gene	Spearman's Correlation SOX2	Spearman's Correlation SOX9
UGT1A7	0,672168643	-0,202442872	LINC01119	-0,252720723	0,211804643
ALDH3A1	0,670920841	-0,282802301	ITGB6	-0,252774555	0,249268487
UGT1A8	0,559019725	-0,218505464	APCDD1L	-0,254850061	0,246704945
AKR1C3	0,508214372	-0,20757116	FOXC2	-0,256316334	0,264770937
NPL	0,48615644	-0,238394264	CTSV	-0,256497048	0,238358554
KHK	0,450784104	-0,218463779	ABL2	-0,258186237	0,326130382
GSTM2	0,448848192	-0,240476263	ICAM1	-0,260477085	0,287577696
SLC25A48	0,436236069	-0,215300643	DCBLD2	-0,261329796	0,306197734
KEAP1	0,428427507	-0,204687311	ADRA1B	-0,262021885	0,319458884
ABCB6	0,424134737	-0,30615012	ELFN2	-0,263208684	0,243649722
HOGA1	0,418008486	-0,235814018	PTPRE	-0,264347828	0,210007012
EVPLL	0,414120911	-0,26753799	ENC1	-0,264367306	0,335379261
SERPINI1	0,401966428	-0,248346521	PPP4R4	-0,264571898	0,269197962
GCHFR	0,401024984	-0,240710001	SLC35F2	-0,265129119	0,219038385
OTOP3	0,392911286	-0,239677186	MCL1	-0,269443531	0,227367462
SEMA3G	0,389681121	-0,207308205	GFPT2	-0,276619068	0,278544159
METTL7A	0,383728379	-0,211212494	TINAGL1	-0,276828999	0,344370596
ARMC7	0,381863887	-0,271074285	LTBP1	-0,282466844	0,231988114
SMAD5-AS1	0,371651921	-0,255862926	TMCC3	-0,28461377	0,277151471
ADAD2	0,370934688	-0,223306161	COL13A1	-0,287433774	0,228344616
PRDX2	0,365645073	-0,241926303	SRPX	-0,2899075	0,230836738
CYP4F12	0,363667501	-0,215614443	FRMD5	-0,292728733	0,218613077
AHRR	0,359295194	-0,216495403	FOXD1	-0,293141957	0,222809573
MIR9-3HG	0,351908643	-0,244138156	MED15	-0,293824774	0,218438891
EXPH5	0,349985716	-0,208927056	MYADM	-0,298066685	0,210422546
ATPAF1	0,343850096	-0,248197188	LITAF	-0,307841474	0,302639723
PANK1	0,336018794	-0,21406929	MICAL2	-0,308202902	0,260476003
NKD2	0,330143965	-0,214899276	TENM3	-0,313222856	0,231001221
SLC48A1	0,32906401	-0,266786926	DNAJB5	-0,31427143	0,233730327
LCMT2	0,326787229	-0,261539727	PCDHGC5	-0,315364996	0,211089192
ACSS2	0,325718095	-0,26797834	MANCR	-0,316318061	0,211442132
REEP1	0,313048635	-0,257784771	C1ORF74	-0,31735439	0,254371764
ADI1	0,309383495	-0,342225834	PDE7A	-0,317529693	0,234361203
CBFA2T3	0,300567464	-0,236096923	C10ORF55	-0,31762492	0,263642286
CBR1	0,300473319	-0,213639689	ST3GAL5	-0,319606281	0,23397164
PCBD2	0,298995143	-0,209219229	TIMP3	-0,320916729	0,237933281
MLYCD	0,293147367	-0,26250606	PLAU	-0,325032031	0,276207862
GABRQ	0,291363532	-0,214386837	SERPINB7	-0,33058006	0,228498277
DCXR	0,290702859	-0,270730171	CCDC71L	-0,332492165	0,221861636
SHMT1	0,28984582	-0,232784554	SEMA7A	-0,333543986	0,279884906

POMGNT2	0,287249814	-0,200468125	GNB1	-0,333696565	0,203643065
RPP25	0,28642091	-0,214973943	TRAF3IP2	-0,333795038	0,309814178
SFXN2	0,282614012	-0,206939202	PMEPA1	-0,33634343	0,244247451
FBXL22	0,282319676	-0,254306837	FOSL1	-0,337122557	0,214035745
SULT1A1	0,281650925	-0,200055837	TREM1	-0,337424468	0,208960602
SLC25A10	0,279528888	-0,235250706	NME7	-0,344884603	0,22195578
GSTA3	0,277254233	-0,224424297	TTPAL	-0,353500442	0,21787186
PROS1	0,265582527	-0,230094404	CGB8	-0,354312725	0,213925218
KAT14	0,263089323	-0,238544679	SH3KBP1	-0,354525209	0,32235703
MCUR1	0,262756642	-0,218350757	RTTN	-0,355802111	0,215443583
MYCL	0,26258289	-0,232382006	PRDM8	-0,357313832	0,270248628
TSGA10IP	0,261158532	-0,263929874	MSN	-0,358283412	0,223475077
LDHD	0,256726457	-0,247454854	TENM3-AS1	-0,360546124	0,30489919
SPIN2A	0,248888587	-0,20475557	SERPINE1	-0,361473501	0,375701214
ZNF503-AS2	0,24110281	-0,200040688	DUSP6	-0,361817615	0,264876985
PDZD4	0,237965744	-0,221744767	DUSP4	-0,365217636	0,295447954
ZNF540	0,236796613	-0,225359116	MRTFA	-0,366096317	0,275284814
ADAL	0,233998693	-0,205250013	BACE2	-0,368057119	0,235119769
SELENOP	0,231711091	-0,204594248	PLAUR	-0,36872164	0,247093493
KRT4	0,226189636	-0,217778315	CITED2	-0,369365402	0,209303634
CYP1A2	0,221491903	-0,251949431	DYRK3	-0,369488763	0,269235764
MPC1	0,22112363	-0,222888568	MAP3K20	-0,370114229	0,259200184
USH1G	0,220489567	-0,200113244	STK17A	-0,370199716	0,20004718
HADH	0,216987768	-0,211814152	PLK3	-0,37511254	0,296585262
BRCC3	0,216245433	-0,206617813	CHST11	-0,377114462	0,312103944
VAR52	0,215755233	-0,243515937	ADTRP	-0,378122998	0,208973588
GPR55	0,214783006	-0,243484622	ERRFI1	-0,378416253	0,320010994
CCDC152	0,213959996	-0,246138996	RPSAP52	-0,383618418	0,298827738
STX19	0,208639269	-0,209853949	NOX5	-0,387888514	0,2886696
PDIA2	0,205545596	-0,208677474	ARHGAP29	-0,3882182	0,287702218
NECAB3	0,201245087	-0,234845993	INHBA	-0,39146229	0,296172975
AVPR2	0,200891513	-0,262888914	EVA1A	-0,392218692	0,294862527
ACTG1	-0,200351256	0,263227834	PPP1R18	-0,396411907	0,210978756
PLEKHG2	-0,200494096	0,231519556	CANT1	-0,400063088	0,260493928
RHBDF1	-0,201278633	0,201897605	SMIM3	-0,406374985	0,246359749
KLF7	-0,2023034	0,228880266	IER5	-0,40728613	0,219333804
LINC00452	-0,203031215	0,232753103	ANXA5	-0,412270374	0,225774581
EZR	-0,203661461	0,206677329	L1CAM	-0,412718372	0,223612506
NFIL3	-0,206202279	0,351672741	ITGA5	-0,416788225	0,293067291
CHI3L2	-0,208849144	0,200122929	ACTN1	-0,425553396	0,326932233
CSNK1E	-0,214851663	0,262313443	CAMK2N1	-0,427583453	0,278602594
WWC2	-0,214922001	0,201686592	DCBLD1	-0,430887165	0,337741529
MARK3	-0,218930173	0,206113545	SH3TC2	-0,431979682	0,234008495
UBASH3B	-0,220984037	0,325303643	SEC14L2	-0,433178013	0,391663565

GMEB1	-0,222990287	0,22984227	KLF6	-0,440251831	0,235589409
LTBR	-0,224643766	0,204119198	ANTXR2	-0,440625162	0,232224016
LRRRC8C	-0,225342816	0,289502788	INPP4B	-0,444617103	0,238038246
ADAMTS6	-0,22655918	0,227853395	MICALCL	-0,457940163	0,212071697
KDSR	-0,228802353	0,202958083	KIF12	-0,458160498	0,220438708
LASP1	-0,230940622	0,229001463	TNS4	-0,460269967	0,204412452
VIM	-0,231811728	0,225368786	CCBE1	-0,461885697	0,216249279
C3ORF52	-0,234749684	0,299155297	CEBPB	-0,464469675	0,24036264
PPP1R14C	-0,235664075	0,276851724	LGALS1	-0,464565983	0,241795367
GPRC5A	-0,237015643	0,221535918	CDH3	-0,469836989	0,232836496
CHST15	-0,237321883	0,239712286	TNFRSF12A	-0,497349888	0,290194262
LAMC2	-0,244683545	0,259394965	TRIML2	-0,498719069	0,205441012
CDCP1	-0,245683424	0,235381642	TMEM51	-0,516323779	0,274026308
SLC20A1	-0,246153064	0,339954464	FHL2	-0,547325432	0,226000744
THBS1	-0,246403033	0,216005203	BCAR3	-0,549906288	0,27681926
TES	-0,248531563	0,263387988	TMEM92	-0,586926693	0,224619959
RP1	-0,248857616	0,251604001			

Supplementary Table 4. IPA Gene Set Signature – Top 50 Canonical Pathways

<i>Ingenuity Canonical Pathways</i>	<i>-log(p-value)</i>	<i>Molecules</i>	<i>Key Biological Function</i>
<i>TGF-β Signaling</i>	2,500	INHBA, PMEPA1, SERPINE1	Cell migration, cell differentiation, apoptosis, cell growth, etc.
<i>Integrin Signaling</i>	1,610	ACTN1, BCAR3, ITGB6	Promote intracellular signals in response to the stimulus by the extracellular matrix resulting in modification in cellular shape, mobility, and progression through the cell cycle.
<i>NRF2-mediated Oxidative Stress Response</i>	1,570	FOSL1, GSTA3, GSTM2	It is activated under oxidative stress conditions and subsequently activates several antioxidative genes and proteins.
<i>ERK/MAPK Signaling</i>	1,550	DUSP4, DUSP6, ITGB6	It is a chain of proteins in the cell that communicates a signal from a receptor on the surface of the cell to the DNA in the nucleus of the cell.
<i>Paxillin Signaling</i>	1,380	ACTN1, ITGB6	Focal adhesions and it functions to adhere cells to the extracellular matrix.
<i>TWEAK Signaling</i>	0,983	TNFRSF12A	It is related to TNF and presented apoptotic ability
<i>IL-17A Signaling in Fibroblasts</i>	0,959	CEBPB	Play a central role in the control of infections and chronic inflammatory conditions
<i>ILK Signaling</i>	0,921	ACTN1, ITGB6	Control of cytoskeletal organization and cell motility.
<i>HIF1-alpha Signaling</i>	0,889	ADRA1B, SERPINE1	Master transcriptional regulator of cellular and developmental response to hypoxia
<i>RHOGDI Signaling</i>	0,863	CDH3, ITGB6	Regulates cell migration, adhesion, proliferation and differentiation.
<i>Transcriptional Regulatory Network in Embryonic Stem Cells</i>	0,818	L1CAM	Process that modulate cellular identity and modulates biological activity.
<i>cAMP-mediated signaling</i>	0,804	DUSP4, DUSP6	Signaling cascade used in cell communication.
<i>Wound Healing Signaling Pathway</i>	0,785	CEBPB, LAMC2	Process of tissue regeneration and repair.
<i>Actin Cytoskeleton Signaling</i>	0,777	ACTN1, ITGB6	Mediate cell motility and cell shape changes during the cell cycle and in response to extracellular stimulus
<i>Eicosanoid Signaling</i>	0,742	AKR1C3	Cell signaling, immune response, cell growth, inflammation, etc.
<i>Remodeling of Epithelial Adherents Junctions</i>	0,738	ACTN1	Principle mediators of cell-cell adhesion
<i>Signaling by Rho Family GTPases</i>	0,712	CDH3, ITGB6	Rho proteins play a role in organelle development, cytoskeletal dynamics, cell movement and other common cellular functions.
<i>TREM1 Signaling</i>	0,703	TREM1	Modulate the immune response stimulated by pathogen and cytokine receptors like IL-1R.
<i>Senescence Pathway</i>	0,646	CEBPB, SERPINE1	Process of stable cell cycle arrest that response to various intrinsic and extrinsic stimuli, associated with cellular and molecular changes.
<i>p53 Signaling</i>	0,587	THBS1	Controlling cell division and cell death.
<i>Role of Tissue Factor in Cancer</i>	0,527	PDIA2	Expressed by tumor cells and contribute to tumor growth, metastasis and tumor angiogenesis.
<i>PAK Signaling</i>	0,521	ITGB6	Activation of MAPK signaling in breast cancer
<i>IL-6 Signaling</i>	0,488	CEBPB	Cytokine produced by numerous different cell types and plays a critical role in regulating the acute inflammation and cancer progression.
<i>HGF Signaling</i>	0,480	ITGB6	Produced by stromal and mesenchymal cells, and it stimulates epithelial cell proliferation, motility, morphogenesis and angiogenesis.
<i>IL-12 Signaling and Production in Macrophages</i>	0,472	CEBPB	Promotes Th1 responses and induces IFNγ production by T and NK cells
<i>Protein Kinase A Signaling</i>	0,470	DUSP4, DUSP6	Play a role in many cellular processes, such as: proliferation, cell cycle, metabolism, etc.
<i>NAD Signaling Pathway</i>	0,445	CEBPB	involved in cell metabolism

<i>PTEN Signaling</i>	0,433	ITGB6	Promote the activation of signals from growth factor receptors on the cell surface to transcription factors in the nucleus.
<i>HMGB1 Signaling</i>	0,413	SERPINE1	Secreted in the tumor microenvironment leading to cell proliferation, differentiation, mobilization and senescence.
<i>Human Embryonic Stem Cell Pluripotency</i>	0,405	INHBA	
<i>Gαq Signaling</i>	0,401	ADRA1B	Signal transduction in cells leading to RhoA activation, phosphatidylinositol 3-kinase (PI3K) and implications in the regulation of the Akt pathway.
<i>WNT/β-catenin Signaling</i>	0,390	CDH3	Biological processes such as proliferation, apoptosis, differentiation, adhesion, metabolism, gene expression, etc.
<i>IL-17 Signaling</i>	0,374	CEBPB	Pro-inflammatory cytokine involved in control infections.
<i>Leukocyte Extravasation Signaling</i>	0,357	ACTN1	Promote the movement of leukocytes out of the vessels and towards the site of tissue damage or infection.
<i>Regulation of The Epithelial Mesenchymal Transition by Growth Factors Pathway</i>	0,351	FOXC2	
<i>Regulation of the Epithelial-Mesenchymal Transition Pathway</i>	0,350	FOXC2	Biological process in which a non-motile epithelial cell changes to a mesenchymal phenotype with invasive capacities
<i>PI3K/AKT Signaling</i>	0,343	ITGB6	Intracellular signaling which promotes cell proliferation and apoptosis, angiogenesis and glucose metabolism.
<i>FAK Signaling</i>	0,332	ADRA1B, BCAR3, ITGB6	Plays critical roles in integrin-mediated signal transductions.
<i>HER-2 Signaling in Breast Cancer</i>	0,305	ITGB6	Promote cells in the breast to grow and divide at an uncontrolled rate.
<i>AMPK Signaling</i>	0,279	ADRA1B	Play a role as regulator of cellular energy homeostasis.
<i>Sirtuin Signaling Pathway</i>	0,234	DUSP6	Involved in metabolic regulation.
<i>SAPK/JNK Signaling</i>	0,219	DUSP4	Activated by a variety of environmental stresses, inflammatory cytokines, growth factors which may regulate the activity of multiple transcription factors.
<i>Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis</i>	0,206	CEBPB	Secretion of cytokines involved in the regulation of angiogenesis and ECM reorganization.
<i>Glucocorticoid Receptor Signaling</i>	0,000	SERPINE1	Regulation of gene expression
<i>Molecular Mechanisms of Cancer</i>	0,000	ITGB6	Causes by specific DNA damage
<i>CDC42 Signaling</i>	0,000	ITGB6	Involved with Ras-mediated cellular transformation, tumorigenesis, and metastasis.
<i>Phospholipase C Signaling</i>	0,000	ITGB6	Activation of Protein kinase C
<i>TEC Kinase Signaling</i>	0,000	ITGB6	Signal transducers that mediate cell-to-cell communication.
<i>T Cell Receptor Signaling</i>	0,000	DUSP6	Regulation of cytokine production, cell survival, proliferation, and differentiation.

Supplementary Table 5. Summary of Clinical and Histopathological Features of the HIPO Cohort

Sample ID	Subsite	Gender	Age [years]	Tobacco	Alcohol	HPV	PFS [event]	PFS [months]	DSS [event]	DSS [months]	Pathological grading	Therapy	SOX2 [gene]	SOX9 [gene]	SOX2 [IHC/IF]	SOX2 [IHC/TC]	SOX9 [IHC/IF]	SOX9 [IHC/TC]
HNC01	oropharynx	male	60,69	no	yes	yes	0	31,5	0	31,5	3	SRCT	8,55164419	8,85140507	15	15	20	9
HNC02	oropharynx	female	57,48	no	no	yes	0	38,3	0	38,3	3	SRT	9,83068882	6,57900604	20	20	4	4
HNC03	oropharynx	male	61,44	yes	yes	yes	0	35	0	35	3	SRCT	9,06074746	7,37264673	15	15	9	4
HNC04	oropharynx	male	62,04	no	yes	yes	1	6,97	1	11	2	SRCT	10,6130605	7,29015528	20	20	8	4
HNC06	oropharynx	male	64,66	yes	yes	yes	0	54,53	0	54,53	3	SRCT	10,4159696	9,08489319	15	12	8	12
HNC07	oropharynx	male	48,21	yes	no	yes	0	25,97	0	25,97	1	S	8,68236619	7,23399158	15	15	12	15
HNC09	nasal cavity	male	45,84	yes	yes	no	0	13,53	0	13,53	2	S	8,59342551	9,44826206	NA	NA	NA	NA
HNC10	oropharynx	male	54,79	yes	yes	no	0	45,43	0	45,43	2	SRCT	10,2042985	7,1310828	1	4	1	1
HNC11	oropharynx	male	64,52	no	no	yes	1	5,7	1	5,7	3	SRCT	9,57213941	9,07769064	5	9	1	1
HNC12	hypopharynx	male	65,24	yes	yes	no	0	55,33	0	55,33	2	SRCT	11,3386616	7,6814489	20	20	6	4
HNC13	oropharynx	male	66,25	yes	no	yes	0	28,93	0	28,93	2	S	10,5137004	7,60660516	20	20	12	9
HNC14	nasal cavity	male	66,07	no	yes	no	0	38,27	0	38,27	2	S	10,406232	8,77519452	20	15	15	20
HNC15	oropharynx	male	52,88	yes	yes	yes	0	27,13	0	27,13	3	SRT	11,3651244	7,90905288	15	15	12	12
HNC16	oropharynx	male	56,51	yes	yes	no	1	18,27	1	50,07	2	SRCT	10,0501759	6,74092064	16	20	20	16
HNC17	nasal cavity	male	75,30	no	no	no	0	43,5	0	43,5	2	S	4,70584506	7,14543801	12	4	15	15
HNC18	larynx	male	61,99	yes	no	no	1	14,13	1	17,27	2	SRCT	11,6369513	9,47667265	16	16	9	4
HNC19	nasal cavity	male	67,64	yes	no	no	0	53,43	0	53,43	2	SRT	5,34872639	9,19091679	1	4	1	4
HNC20	oral cavity	male	59,88	yes	yes	no	0	42,73	0	42,73	2	SRCT	7,74376776	7,25655674	12	12	6	6
HNC21	larynx	male	53,90	yes	yes	no	1	11,57	1	12,57	2	SRCT	6,70836243	8,32100805	4	4	9	4
HNC22	oropharynx	male	54,59	yes	no	no	1	17,17	1	17,17	3	SRCT	7,57803919	8,08742797	6	9	1	12
HNC23	larynx	male	66,21	yes	yes	no	0	59,87	0	59,87	2	SRT	11,0831755	8,43384952	1	15	1	6
HNC24	oral cavity	female	76,27	yes	no	no	1	16,33	1	16,77	3	S	11,3189535	9,79056882	20	20	9	9
HNC26	nasal cavity	male	45,82	no	no	no	0	60,57	0	60,57	2	S	7,71887722	9,10431944	9	6	1	1
HNC27	oropharynx	male	63,62	yes	yes	no	0	45,03	0	45,03	3	SRCT	7,65216986	7,74921074	12	12	20	20
HNC28	oropharynx	male	71,55	yes	yes	yes	0	40,8	0	40,8	2	S	9,74435299	6,21645572	20	20	12	9
HNC29	oral cavity	male	56,02	no	no	no	1	10,77	1	28,17	3	SRCT	5,18576669	8,70205658	1	1	4	4
HNC31	oropharynx	male	67,75	yes	yes	no	0	55,83	0	55,83	3	SRCT	8,94631254	8,86284354	15	15	8	1
HNC33	oropharynx	female	46,27	yes	no	no	0	51,87	0	51,87	2	SRT	8,41806075	7,30919284	6	15	1	9
HNC34	oral cavity	female	73,39	no	no	no	0	42,37	0	42,37	2	SRT	9,07824413	6,78086719	4	6	4	8
HNC35	oropharynx	male	40,61	no	no	yes	0	55,57	0	55,57	2	SRCT	9,83555173	8,12922988	NA	NA	NA	NA
HNC36	oropharynx	male	50,33	yes	yes	yes	1	14,2	1	34,33	3	SRCT	9,64015234	5,83607897	10	8	1	1
HNC37	larynx	male	64,47	yes	no	no	1	15,4	0	30,33	2	SRCT	11,1173613	8,62236656	20	9	4	4
HNC38	oropharynx	female	54,57	yes	yes	no	1	46,43	0	48,93	2	SRCT	9,09763623	6,77537001	20	15	4	1
HNC39	oropharynx	male	56,60	yes	no	yes	0	31,3	0	31,3	3	SRT	10,519821	4,09033755	1	12	10	6
HNC40	oral cavity	male	60,44	yes	yes	no	0	57,73	0	57,73	2	SRT	7,69027848	7,54315753	15	15	1	12
HNC42	oropharynx	male	48,12	no	no	yes	0	32,53	0	32,53	3	S	10,3939246	5,54576025	10	20	9	12
HNC43	larynx	male	69,22	yes	yes	no	0	40,33	0	40,33	2	S	7,42736521	7,18069571	4	4	12	9
HNC44	oropharynx	male	64,27	yes	yes	yes	0	35,3	0	35,3	2	SRCT	10,1596331	6,92975478	20	15	9	12
HNC45	oral cavity	female	75,26	no	yes	no	0	2,63	0	2,63	3	S	3,33537075	7,84972908	1	6	9	6
HNC46	oral cavity	male	77,99	no	no	no	0	44,67	0	44,67	2	S	7,12027445	8,35214519	NA	NA	NA	NA
HNC47	larynx	female	59,26	yes	yes	no	0	8,23	0	8,23	2	SRT	3,90413815	7,49791613	1	1	9	6



HNC48	nasal cavity	male	55,70	no	no	no	0	25,23	0	25,23	2	S	5,73589813	9,30048127	15	9	1	6
HNC49	oral cavity	male	57,21	yes	yes	no	1	9,7	1	17,7	2	SRCT	6,32403632	7,6551884	1	8	4	6
HNC50	oropharynx	male	46,86	no	no	yes	0	33,67	0	33,67	2	SRT	11,2288238	6,10096533	12	15	15	12
HNC52	oral cavity	female	67,68	no	no	no	0	44,33	0	44,33	3	S	7,63061241	8,94953995	4	6	1	9
HNC53	nasal cavity	male	55,55	yes	yes	no	1	18,3	1	24,9	3	SRT	7,13668043	6,19934622	15	15	4	10
HNC54	nasal cavity	female	59,49	no	yes	no	0	27,53	0	27,53	3	S	11,3419134	6,37312227	20	15	4	6
HNC55	larynx	female	82,52	yes	no	no	0	48,43	0	48,43	2	S	11,7161035	7,46246646	4	2	1	1
HNC56	hypopharynx	male	68,44	yes	no	no	0	26,07	0	26,07	2	SRT	10,2725458	6,78340535	20	20	12	9
HNC57	oropharynx	male	56,94	no	yes	yes	0	56,2	0	56,2	3	SRCT	10,1316016	8,35214519	20	15	15	6
HNC58	oral cavity	male	73,68	yes	yes	no	0	1,8	0	1,8	3	S	7,0618009	6,95692139	20	9	12	4
HNC59	hypopharynx	male	59,59	yes	yes	no	0	5,53	0	5,53	2	S	10,2182751	7,04235515	20	9	12	15
HNC60	oral cavity	female	68,06	no	no	no	0	38,47	0	38,47	2	SRT	7,40027608	5,39674341	1	6	1	4
HNC62	larynx	male	63,52	yes	no	yes	0	28,33	0	28,33	2	SRT	10,7726224	6,13988019	NA	NA	NA	NA
HNC63	oropharynx	female	56,65	yes	no	no	1	10,33	0	65,17	2	SRCT	11,2627458	8,31555322	20	20	12	12
HNC64	larynx	male	60,88	yes	no	no	1	4,93	1	7,5	2	S	6,82559073	8,04467156	4	6	12	12
HNC65	larynx	male	61,75	yes	no	no	1	19,9	0	19,9	3	SRT	7,40215608	10,7803478	6	6	20	15
HNC66	oropharynx	female	72,35	no	no	yes	0	17,8	0	17,8	3	SRCT	11,0062458	7,43034317	20	15	20	15
HNC67	oropharynx	male	51,55	yes	yes	yes	0	43,83	0	43,83	3	SRCT	8,60801302	7,23641174	15	15	15	15
HNC68	larynx	male	49,45	yes	yes	no	1	19,87	0	19,87	2	SRCT	10,8679004	7,53407961	1	1	1	1
HNC69	oral cavity	male	74,59	yes	no	no	0	38,93	0	38,93	3	SRT	8,65489044	8,40554513	6	6	1	1
HNC70	oral cavity	male	63,29	yes	yes	no	1	4,57	1	15,07	2	S	7,12653782	9,29039949	16	16	16	12
HNC71	oral cavity	male	39,68	no	no	no	0	26,37	0	26,37	2	SRCT	8,55518143	7,93542086	15	6	15	15
HNC72	oropharynx	male	62,53	yes	yes	no	1	7,07	1	14,6	2	SRCT	5,09329199	9,08546928	12	12	16	20
HNC73	oropharynx	male	57,12	yes	yes	no	1	5,63	1	9,73	3	SRT	9,01437516	9,53685314	4	12	12	12
HNC74	oral cavity	male	56,31	yes	no	no	1	14,73	0	40,93	2	S	8,3721437	7,97217651	16	15	20	20
HNC75	larynx	male	50,78	yes	yes	no	0	29,17	0	29,17	2	S	8,80944661	5,80054189	16	4	6	6
HNC76	larynx	male	55,68	yes	no	no	0	20,37	0	20,37	3	SRT	9,59184116	6,9443709	1	4	1	4
HNC77	oral cavity	female	63,45	yes	no	no	0	2,1	0	2,1	3	SRT	7,95132425	6,97642318	1	12	4	9
HNC78	oral cavity	male	70,08	yes	yes	no	0	37,43	0	37,43	2	SRT	7,55381549	7,53701663	12	6	1	6
HNC79	nasal cavity	male	58,01	no	no	no	0	27,77	0	27,77	2	S	9,12258834	8,18907474	20	20	12	20
HNC80	oral cavity	female	52,00	yes	no	no	0	53,93	0	53,93	2	S	7,70379557	7,83301138	6	4	8	8
HNC82	nasal cavity	male	75,95	yes	no	no	1	9,3	1	11,5	3	SRCT	11,1841251	7,6108414	1	15	1	6
HNC83	oropharynx	male	56,84	yes	yes	yes	0	25,2	0	25,2	3	SRT	9,93240664	5,86704804	12	10	4	6
HNC84	nasal cavity	female	68,07	no	no	yes	0	35,5	0	35,5	2	SRT	10,7201689	8,47923558	20	20	6	9
HNC85	oropharynx	female	66,58	yes	no	no	0	15	0	15	2	SRT	10,8967114	7,85266868	20	20	9	6
HNC86	oropharynx	male	61,39	no	yes	yes	0	24,33	0	24,33	3	S	10,3540598	7,37947988	20	20	9	20

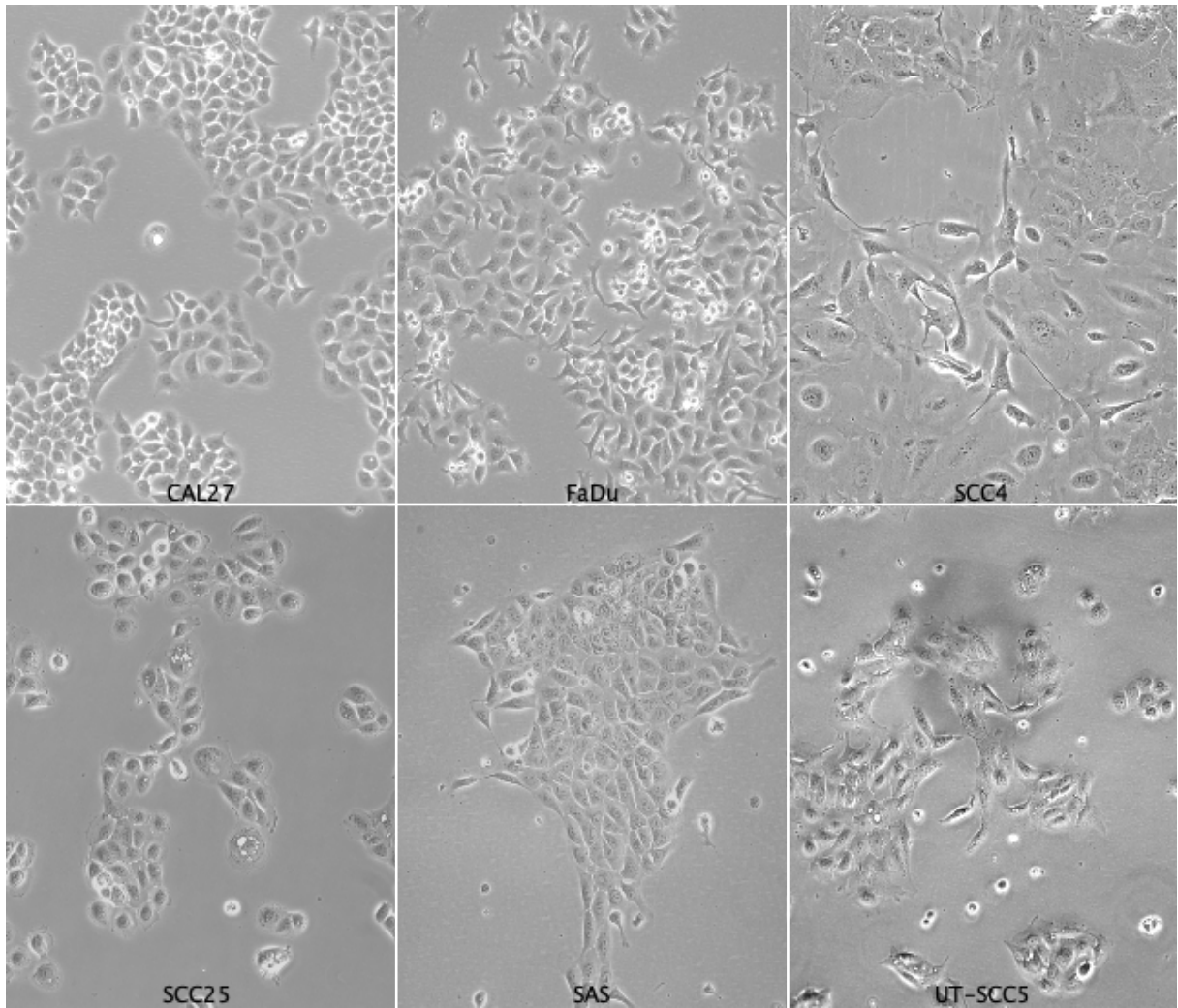
HPV-related= DNA and RNA positive

S=surgery; RT=adjuvant radiotherapy; RCT=adjuvant radio chemotherapy

IHC = immunohistochemistry

IF= invasion front of the tumor sample, TC= tumor core of the sample

NA= cases excluded of the analysis



Supplementary Figure 1. Morphological characteristic of HNSCC cells. Phase contrast images of cell morphology for Cal27, FaDu, SCC4, SCC25, SAS and UT-SCC5 cell lines. Images were taken in Inverted Zeiss microscope Axiovert 100 using 20x objective.

## **Curriculum Vitae**

Name: Silvia Barbosa  
Date of Birth: 01.05.1985  
Place of Birth: Porto Alegre – Rio Grande do Sul - Brazil  
Marital Status: Married

## **Education**

2018- present      Medical Faculty of Heidelberg,  
University of Heidelberg, Germany  
DKFZ International PhD Program in Heidelberg  
(Doctoral Candidate)

2014- 2016      Federal University of Rio Grande do Sul, Brazil  
Master of Science – Cellular and Molecular Biology

2009-2013      Federal University of Health Sciences of Porto Alegre, Brazil  
Bachelor of science – Biomedicine  
Major in Cellular and Molecular Biology / Minor in Human Genetics

## **Professional Career/Experience**

2018- present      Doctoral research in the Section of Molecular Mechanisms of  
Head and Neck Tumors and Division of Radiooncology /  
Radiobiology at Deutsches Krebsforschungszentrum (DKFZ)

2014- 2016      Master research in the Cellular Migration Laboratory (LAMOC) at  
Federal University of Rio Grande do Sul - Brazil

2009-2022      Research assistant in the Department of Morphological Sciences  
at Federal University of Health Sciences of Porto Alegre, Brazil

## **Acknowledgement**

I would like to start by thanking the DAAD who supported me financially during the entire period of my PhD, without this support this dream would not have come true. I am infinitely grateful to the DKFZ graduate school and the amazing staff that always helped and supported me. My enormous gratitude and thanks to Dr. Wolfgang Hagmann for his last-minute participation and for helping me so much in the last steps of this thesis.

My enormous thanks to my Doktor Vater, Prof. Dr. Jochen Hess, for the opportunity, the challenges, the patience and for all his contribution. To Dr. Ina Kurth, for all the support, criticism, help and especially for opening the doors of her lab and receiving me with great respect and friendship.

To Dr. Adriana Jou, my deepest thanks. To my laboratory colleagues at Kopfklinik and DKFZ. A special thanks to all the technicians who helped me, supported me and taught me a lot during these years. Antje, Nataly, Ines, Sonja, Rosemarie and Mahnaz, you live in my heart and you will have my eternal gratitude.

To my PhD colleagues, thank you very much for all the talks, complaints, and many partnerships in the lab. A special thanks to Dr. Maria J. Besso, for all the teachings, all the talks and all the tips you gave me. To Dr. Wahyu W. Hadiwikarta, thank you for teaching me bioinformatics and for believing that I could do it.

To Dr. Alexander Rölle and Dr. Julia Pollmann, my incredible thanks, with all my heart and respect, for taking time for fruitful discussions, for developing the cytokine assay, and for teaching me FACs. To Dr. Mareike Roscher, who literally did the NK assay with me and with enormous patience, thank you, my dear.

I also thank my family for all their support and for always loving me. My husband, Sven Markus, who has been an amazing partner and never stopped believing that this thesis would happen. To my friends at UFRGS, who gave me all the support and believed in me. To Laura and her family, who have always been there for everything and listened to many stories throughout these years.

I would also like to thank all the cancer patients that donated their samples so that the scientific research could be carried out, their families, and all the doctors that helped in the research and in the treatment of these patients.

My deepest gratitude!

## Eidesstattliche Versicherung

1. Bei der eingereichten Dissertation zu dem Thema

*The Role of SOX2 and SOX9 as Regulatory Genes of the Head and Neck Squamous Cell Carcinoma Microenvironment*

handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslandes als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärung bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eid statt, dass ich nach bestem Wissen die reine Wahrheit erkläre und nichts verschwiegen habe.

Heidelberg, August 2022



Silvia Barbosa