The role of the homeobox gene HOXD10 in Lung and Head & Neck carcinoma

Fahad Mohammad I Hakami

BSc, MSc

A thesis submitted to the University of Sheffield in part fulfilment for the degree of Doctor of Philosophy

The School of Clinical Dentistry

University of Sheffield

December 2013

This thesis is dedicated to the memory of my father, Mohammad Ibrhaim Aiafi Hakami.

ACKNOWLEDGEMENTS

First and foremost, I thank The Almighty Allah for blessing, protecting and guiding me throughout my personal and professional life.

This thesis is the end of my journey in obtaining my PhD, and it would not end happily without the advice, motivation, patience and, most importantly, the friendship of my principal PhD supervisor, Dr Keith Hunter. Keith has always been available to guide me through all the challenges I faced during my work until I developed independent thinking. I am deeply grateful for his endless support and his believing in me. It has been an honour and a privilege working with Dr Hunter.

I would also like to extend my appreciation to my co-supervisors Professor Penella Woll for introducing me to Keith and for her contributions during discussing the progress of my work, and Dr Daniel Lambert for all his insightful discussions regarding my work and for his generosity in answering my endless questions.

I gratefully acknowledge Dr Lynn Bingle for teaching me all the IHC optimisation tricks, Mrs Prachi Stafford for all the technical help, Dr Paul Heath for his help in the microarray analysis and Mrs Brenka McCabe for keeping my drawer full with all items I needed for my work. A special thank you to Dr Martin Nicklin for all the invaluable molecular biology knowledge I gained before I started my PhD studies.

Also, it is my pleasure to thank everyone in the School of Clinical Dentistry, particularly, in the Unit of Oral and Maxillofacial Pathology. Big thank you to my fellow colleagues, Mr Lav Darda, Mrs Nada AlHindi, Mrs Ibtisam Zargoun, Miss Emma Hinsley, Miss Tasnuva Kabir, Miss Genevieve Melling, Mrs Hanan Niaz, Miss Dalal Al-Otaibi, Miss Kate Naylor, Miss Hayley Lunn and Miss Abigaile Rice.

I am very grateful to Dr Mohammad Abdelaal, Mr Mohammad Al-Shehri and my wise mentor, Dr Baraa AlHaj-Huessin. Without their help and opportunities they offered me at the beginning of my career, I would not have been able to complete my postgraduate studies. Also, a special thank goes to the movies expert, Fawaz Albloui, who made my stay at Sheffield more enjoyable.

Last but not least, words are short to express my deep sense of gratitude towards my wife, Hana Mobarki. I am grateful for her understanding when I worked late for long days, standing by me through the good and bad times, and for taking good care of our daughter, Hala. I doubt that I will ever be able to fully convey my appreciation to her.

The work described in this thesis has been supported by the National Guard Health Affairs, Jeddah, through the Royal Embassy of Saudi Arabia Cultural Bureau in London.

ABSTRACT

Alterations in gene expression in head and neck squamous cell carcinoma (HNSCC) and non-small cell lung cancer (NSCLC) are variable and correlate with cancer site and stage. Previous microarray analysis comparing HNSCC to normal oral keratinocytes (NOKs) revealed changes in HOX gene expression, particularly HOXD10. Similar data were identified in NSCLC in other study when a comprehensive analysis system based on the quantitative real-time polymerase chain reaction (qPCR) was performed. Moreover, HOXD10 expression is altered in many cancers including breast, ovarian, prostate and gastric cancers. This study demonstrates a low HOXD10 mRNA and protein expression in normal cells, variable levels in precancerous cells, high expression in most primary tumours, but low in cells derived from lymph node metastases. Moreover, it was demonstrated that HOXD10 expression is reduced in HNSCC metastases relative to their paired primary tumours. Over-expression of HOXD10 decreased cell invasion but increased proliferation, adhesion and migration, with siRNA and shRNA knock-down causing reciprocal effects, indicating possible role of HOXD10 in promoting the development of the primary tumour. Microarray analysis using the Agilent Sureprint G3 Array in cells with stable over-expression or transient knockdown of HOXD10 demonstrated increases in expression of a number of genes known to promote cell proliferation and modulate epithelial-to-mesenchymal transition (EMT), in keeping with its suggested role as a suppressor of metastasis. This was supported further by the observation that HOXD10 acts as a negative regulator of a number of genes known to promote cell invasion. Gene ontology (GO) analysis has mapped many of the genes affected by HOXD10 manipulation to different pathways responsible for different biological processes, particularly cell proliferation and migration. The microarray analysis also identified a number of HOXD10-associating molecules. Validation by qPCR and filtering by *in-silico* promoter analysis and applying some selection criteria identified a number of HOXD10 putative targets, including Angiomotin p80 isoform (AMOT-p80) and miR146a. These were confirmed as direct targets of HOXD10 by dual luciferase reporter (DLR) assay in the wild-type and mutated HOXD10 binding site(s) on the promoter. Suppression of miR-146a by HOXD10 was found to induce the expression of genes previously reported to induce tumour growth and are targets of miR-146a. Manipulation of AMOT-p80 expression in HNSCC cells resulted in similar phenotypic changes to those seen with manipulation of HOXD10 expression. Indeed, AMOT-p80 is able to rescue the migratory and proliferative phenotype of knockdown of HOXD10. Thus, HOXD10 expression varies by stage of disease in HNSCC and produces differential effects, with high expression giving premalignant and cancer cells a probable proliferative and migratory advantage over normal cells while low expression may later support metastasis.

TABLE OF CONTENTS

Abstract ..... **ii**

Table of contents ..... **iii**

List of Figures ..... **vii**

List of Tables ..... **x**

Abbreviations ..... **xi**

Publications ..... **xvi**

CHAPTER 1: INTRODUCTION **1**

1.1 Cancer 2

1.2 Head and neck cancer 2

1.2.1 Epidemiology and incidence *6*

1.2.2 Diagnosis and staging *6*

1.2.3 Aetiology and risk factors *7*

1.2.3.1 Smoking and alcohol 7

1.2.3.2 Human papillomavirus (HBV) 9

1.2.3.3 Other HNSCC aetiological factors 10

1.2.1 Genetic alterations and molecular markers *12*

1.2.1.1 Tumour protein p53 (TP53) 13

1.2.1.2 Cyclin-Dependent Kinase Inhibitor 2A (CDKN2A) 14

1.2.1.3 Epidermal growth factor receptor (EGFR) 15

1.3 Lung cancer 17

1.3.1 Epidemiology *20*

1.3.2 Diagnosis and staging *20*

1.3.3 Aetiology and risk factors *23*

1.3.3.1 Smoking 23

1.3.3.2 Radon and Asbestos 24

1.3.3.3 Genetic factors 24

1.3.3.4 Other aetiological factors: 25

1.3.4 Genetic alterations and molecular markers *25*

1.3.4.1 Epidermal growth factor receptor (EGFR) 29

1.3.4.2 Kirsten rat sarcoma viral oncogene homolog (KRAS) 30

1.3.4.3 Tumour protein p53 (TP53) 30

1.4 The need for new HNSCC and NSCLC biomarkers 31

1.5 *HOX* genes 32

1.5.1 Discovery: *32*

1.5.2 HOX of homeobox *34*

1.5.3 Specificity of HOX genes *34*

1.5.4 Regulation of HOX genes *35*

1.5.5 HOX genes in human development *39*

1.5.6 Congenital disorders of HOX genes *40*

1.5.7 HOX genes and cancer *41*

1.5.8 HOX genes as potential therapeutic targets in cancer ... 45
1.6 HOXD10 ... 46
1.6.1 Pathologies associated with abnormal HOXD10 expression ... 46
1.6.2 Cancer-related HOXD10 interactions ... 49
1.7 Project hypothesis and aims ... 50

CHAPTER 2: MATERIALS AND METHODS ... 51

2.1 Materials ... 52
2.1.1 Equipment ... 52
2.1.2 Reagents and chemicals ... 52
2.1.3 Buffers and solutions ... 52
2.1.4 Antibodies ... 52
2.1.5 Peptides ... 53
2.1.6 Antibiotics ... 53
2.1.7 Media for eukaryotic cell culture ... 53
2.1.8 Cells ... 53
2.1.9 Primers and probes ... 55
2.1.10 Expression vectors and siRNA/shRNA ... 55
2.2 Methods ... 56
2.2.1 Cell culture ... 56
2.2.1.1 Routine cell culture procedure ... 56
2.2.1.2 Cell passaging and harvesting ... 56
2.2.1.3 Cell counting ... 57
2.2.1.4 Cryogenic preservation and recovery ... 57
2.2.2 Quantitative real-time polymerase chain reaction (qPCR) ... 57
2.2.2.1 Total RNA extraction ... 57
2.2.2.2 Complementary DNA (cDNA) preparation ... 58
2.2.2.3 Designing qPCR primers ... 59
2.2.2.4 Running qPCR reactions ... 60
2.2.2.5 Assessment of the quality of SYBR® Green-based qPCR reaction ... 61
2.2.2.6 qPCR data analysis ... 61
2.2.3 Detection of protein expression in cell lysates ... 62
2.2.3.1 Protein extraction ... 62
2.2.3.2 Protein quantification ... 63
2.2.3.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting (WB) 63
2.2.4 Detection of protein expression in tissue sections ... 65
2.2.4.1 Scoring of immunohistochemistry (IHC) staining ... 66
2.2.5 DNA cloning into plasmid vectors ... 66
2.2.5.1 Cloning of a coding sequence ... 66
2.2.5.1.1 Designing the PCR primers ... 66
2.2.5.1.2 PCR reaction ... 66
2.2.5.1.3 Assessment of PCR products using agarose gel electrophoresis ... 67
2.2.5.1.4 Insertion of a gene's coding sequence into a mammalian expression vector ... 67
2.2.5.2 Cloning of a gene's promoter ... 68

2.2.5.2.1 Detecting a promoter region ... 68
2.2.5.2.2 Designing the PCR/cloning primers ... 68
2.2.5.2.3 Mutagenic PCR ... 69
2.2.5.2.4 Promoter DNA to vector ligation ... 70
2.2.6 Amplification and purification of a cloned vector ... 70
2.2.6.1 Bacterial transformation ... 70
2.2.6.2 Purification of plasmid DNA from miniprep cultures ... 71
2.2.6.3 DNA sequencing ... 71
2.2.6.4 Maxiprep amplification and extraction ... 71
2.2.7 Gene over-expression and silencing ... 72
2.2.7.1 In-vitro transient and stable transfection of a plasmid DNA ... 72
2.2.7.2 In-vitro transient transfection of siRNA ... 73
2.2.8 Assays of cell function ... 73
2.2.8.1 Transwell cell migration assay ... 73
2.2.8.2 Transwell cell invasion assay ... 75
2.2.8.3 Cell adhesion assay ... 76
2.2.8.4 Cell proliferation assay ... 77
2.2.9 Dual luciferase reporter assay ... 77
2.2.10 Microarray ... 79
2.2.10.1 RNA preparation and quality control ... 79
2.2.10.2 Sample labelling and hybridisation ... 79
2.2.10.3 Data acquisition and analysis ... 81
2.2.10.4 Gene Ontology (GO) and pathway analysis ... 82
2.2.1 Statistical Analysis ... 82
2.2.2 Ethics ... 82

CHAPTER 3: DIFFERNTIAL AND ALTERED HOXD10 EXPRESSION IN HNSCC AND NSCLC CELLS ... 83

3.1 Introduction ... 84
3.2 HOXD10 mRNA and protein level in a panel of HNSCC and NSCLC cell lines ... 85
3.3 HOXD10 protein level in HNSCC tissue sections ... 89
3.4 Basal expression level of HOXD10 versus its previously reported target molecules ... 93
3.5 *In-vitro* HOXD10 over-expression and knocking down ... 96
3.6 Differential *IGFBP3* and *miR-7* expression levels following HOXD10 manipulation ... 98
3.7 HOXD10 over-expression induces cell proliferation and vice versa ... 101
3.8 HOXD10 over-expression promotes cell migration and vice versa ... 101
3.9 HOXD10 over-expression enhances cell adhesion and vice versa ... 104
3.10 HOXD10 over-expression impairs cell invasion and vice versa ... 106
3.11 Discussion ... 109

CHAPTER 4: MICROARRAY ANALYSIS ... 112

4.1 Introduction ... 113
4.2 Pre-analysis Quality Control ... 113
4.3 Microarray data analysis ... 115
4.4 Using *T*-test to identify HOXD10-associating genes ... 117
4.5 Filtering the list of HOXD10-associating genes identified by *T*-test ... 119

4.6 qPCR analysis to validate the candidate genes that correlate with HOXD10 expression level 119
4.7 Using ANOVA test to identify HOXD10-associating genes 124
4.8 Gene Ontology and pathway analysis 126
4.9 Validation of HOXD10-associating genes identified by ANOVA test 131
4.10 Discussion 139

CHAPTER 5: VALIDATION OF HOXD10 PUTATIVE TARGETS 142

5.1 Introduction 143
5.2 Assessment of HOXD10 binding to the promoters of its correlating genes 143
5.3 AMOT-p80 protein in HNSCC tissues and cell lines 149
5.4 AMOT-p80 protein expression in cells transfected with cloned HOXD10 or AMOT-p80 149
5.5 The function of AMOT-p80 in HNSCC cells 153
5.6 The effect of miR-146a on its downstream targets in HNSCC and NSCLC 157
5.7 Discussion 161

CHAPTER 6: GENERAL DISCUSSION AND FUTURE WORK 163

6.1 General discussion 164
6.2 Future work 169

CHAPTER 7: REFERENCES 170

APPENDICES 211

LIST OF FIGURES

Figure 1.1: A cross-section diagram of the head and neck. 3

Figure 1.2: The progression of epithelial precursor lesions in HNSCC. 5

Figure 1.3: Proposed mechanism of HPV oncogenic role in HNSCC. 11

Figure 1.4: Cross section of human respiratory system. 18

Figure 1.5: SCLC and subtypes of NSCLC. 19

Figure 1.6: A model of molecular changes that occur during lung carcinogenesis. 28

Figure 1.7: Homeotic transformation in the head of *Drosophila Melanogaster*. 33

Figure 1.8: Clusters of human HOX genes. 37

Figure 1.9: 3D structure of a homeodomain interaction with DNA. 38

Figure 1.10: Congenital disorders associated with a mutated HOXD10 gene. 48

Figure 2.1: A diagram of a two-step mutagenic PCR. 69

Figure 2.2: A diagrammatic representation of Transwell migration assay. 74

Figure 2.3: A diagrammatic representation of Transwell invasion assay. 75

Figure 2.4: A diagram of cell adhesion assay. 76

Figure 2.5: A diagrammatic sketch of the principle of luciferase reporter assay. 78

Figure 2.6: Steps of the Agilent one-colour microarray sample labelling and hybridisation. 80

Figure 3.1: A screenshot for a result of RNA sample analysis using the NanoDrop spectrophotometer. 86

Figure 3.2: HOXD10 mRNA level in a panel of HNSCC and NSCLC cell lines. 87

Figure 3.3: HOXD10 protein level in a panel of cell lines. 88

Figure 3.4: Photomicrographs of IHC staining for HOXD10 in prostate tissue sections. 90

Figure 3.5: HOXD10 protein level in a cohort of HNSCC tissue sections. 91

Figure 3.6: A screen for HOXD10 protein level in primary tumours vs. metastasis. 92

Figure 3.7: HOXD10 basal expression level vs. its previously reported downstream targets, *IGFBP3* and *miR-7*, in normal oral keratinocytes and a panel of HNSCC cell lines. 94

Figure 3.8: HOXD10 basal expression level vs. previously reported regulator, miR-10b, in normal oral keratinocytes and a panel of HNSCC cell lines. 95

Figure 3.9: Transient manipulation of HOXD10 expression. 97

Figure 3.10: The effect of HOXD10 level manipulation on *miR-7* expression. 99

Figure 3.11: The effect of HOXD10 level manipulation on *IGFBP3* expression. 100

Figure 3.12: The effect of HOXD10 transient over-expression or silencing on cell proliferation..... 102

Figure 3.13: The effect of HOXD10 transient over-expression or silencing on cell migration..... 103

Figure 3.14: The impact of HOXD10 transient over-expression or silencing on cell adhesion..... 105

Figure 3.15: Stable manipulation of HOXD10 expression. 107

Figure 3.16: The effect of HOXD10 level manipulation on cell invasion. 108

Figure 4.1: Examples of monitoring microarray pre-analytical QC. 114

Figure 4.2: A scatter and Box & whisker plot for samples' PCA scores and normalised values, respectively. 116

Figure 4.3: Venn diagram of the array analysis using *T*-test. 118

Figure 4.4: A dissociation curve and an agarose gel electrophoresis following a SYBR® Green qPCR run. 121

Figure 4.5: HOXD10 basal and manipulated expression level versus USP6 level. 122

Figure 4.6: HOXD10 basal and manipulated expression level versus CYP2J2 level. 123

Figure 4.7: A snapshot for the gene expression analysis output using the ANOVA statistical method. 125

Figure 4.8: Pathway analysis using GePS analysis tool. 128

Figure 4.9: Two cell proliferation pathways identified by GO analysis. 129

Figure 4.10: OCLN and JAG2 genes associated with EMT are affected by HOXD10 manipulation. 130

Figure 4.11: A heat-map of the 48 genes associated with HOXD10 according to ANOVA statistical analysis. 133

Figure 4.12: A scheme of filtering HOXD10-associated genes identified by ANOVA test. 134

Figure 4.13: HOXD10 basal and manipulated expression level versus GBP2 level. 136

Figure 4.14: HOXD10 basal and manipulated expression level versus AMOT-p80 level. 137

Figure 4.15: HOXD10 basal and manipulated expression level versus miR-146a level. 138

Figure 5.1: A schematic diagram of HOXD10 binding sites in AMOT-p80 and miR-146a promoters. 145

Figure 5.2: Snapshots of the sequencing results demonstrated on Finch TV for HOXD10- binding sites in AMOT-p80 and miR-146a promoters. 146

Figure 5.3: DLR analysis of HOXD10 binding to the wild-type and mutated promoters of *AMOT-p80*, *miR-146a* and *CYP2J2* genes. 147

Figure 5.4: DLR analysis of HOXD10 binding to WT and mutated promoters of AMOT-p80. 148

Figure 5.5: Detection of AMOT-p80 protein in normal and tumour tissue sections. 150

Figure 5.6: AMOT-p80 protein and mRNA levels in a panel of cell lines. 151

Figure 5.7: qPCR and WB analysis of AMOT-p80 expression driven by either HOXD10 or AMOT-p80 transfection. 152

Figure 5.8: Effect of AMOT-p80 over-expression on cell proliferation, migration and adhesion. 154

Figure 5.9: Rescuing AMOT-p80 expression following HOXD10 knocks down. 155

Figure 5.10: The impact of rescuing AMOT-p80 expression on the migration and proliferation of HOXD10-silenced cells. 156

Figure 5.11: The effect of miR-146a over-expression on HOXD10 level. 158

Figure 5.12: miR-146a and NOTCH1 levels in HOXD10-manipulated cells and in a panel of cell lines. 159

Figure 5.13: Effect of HOXD10 manipulation on miR-146a and its downstream targets, CXCR4 and L1CAM. 160

Figure 6.1: The identified HOXD10-associating molecules and their effect on cellular functions. 168

LIST OF TABLES

Table 1.1: TNM staging (7th edition) for cancers of the lips and oral cavity..... 8

Table 1.2: TNM staging system (7th edition) for NSCLC..... 22

Table 1.3: Stage grouping and sub-grouping of NSCLC (7th edition)..... 22

Table 2.1: Cell lines used in this study..... 54

Table 2.2: TaqMan® pre-developed probes/primers..... 55

Table 3.1: The Abs tested to detect HOXD10 protein in cell lysate and tissue sections. 89

Table 4.1: The 19 associating genes with HOXD10 identified by *T*-test. 120

Table 4.2: Top 20 GO terms ranked by their *p*-values..... 127

Table 4.3: The 48 genes associated with HOXD10 according to ANOVA test. 132

Table 4.4: The final list of the eight HOXD10-correlated genes identified by ANOVA test. 135

Table 5.1: The cloning and mutagenic primers used to amplify, clone and mutate AMOT-p80, CYP2J2 and miR-146a promoters. 144

ABBREVIATIONS

3D	Three Dimensional
AAH	Atypical Adenomatous Hyperplasia
Ab	Antibody
ABC	Avidin and Biotinylated Horseradish Peroxidase Macromolecular Complex
ABCA3	Atp-Binding Cassette, Sub-Family A (Abc1), Member 3
ACTA1	Actin, Alpha 1, Skeletal Muscle
AIS	Adenocarcinoma In Situ
AJCC	American Joint Committee for Cancer
ALDH3A1	Aldehyde Dehydrogenase 3 Family, Member A1
ALK	Anaplastic Lymphoma Receptor Tyrosine Kinase
ALL	Acute Lymphoid Leukaemia
AML	Acute Myeloid Leukaemia
AMOT	Angiomotin
ANT-C	Antennapedia Complex
A-P	Anterior-Posterior
APS	Ammonium Persulfate
AR	Antigen Retrieval
AREG	Amphiregulin
B2M	Beta-2 Microglobulin
BAI3	Brain-Specific Angiogenesis Inhibitor 3
BCA	Bicinchoninic Acid Protein Assay
bp	Base Pair
BRAF	V-Raf Murine Sarcoma Viral Oncogene Homolog B
BSA	Bovine Serum Albumin
BX-C	Bithorax Complex
CDK4	Cyclin-Dependent Kinase 4
CDK6	Cyclin-Dependent Kinase 6
CDKN2A	Cyclin-Dependent Kinase Inhibitor 2A
cDNA	Complimentary DNA
CDX	Caudal Type Homeobox 2
CEL	Carboxyl Ester Lipase
ChIP	Chromatin Immunoprecipitation
CML	Chronic Myeloid Leukaemia
CMT	Charcot-Marie-Tooth
COL17A1	Collagen, Type XVII, Alpha 1
cRNA	Complimentary RNA

CRTAC1	Cartilage Acidic Protein 1
CT	Computed Tomography
C_T	Threshold Cycle
CVT	Congenital Vertical Talus
CXCL9	Chemokine (C-X-C Motif) Ligand 9
CXR	Chest Radiography
CYP3A4	Cytochrome P450, Family 3, Subfamily A, Polypeptide 4
dH_2O	Deionized Water
DCC	Deleted in Colorectal Carcinoma
DIPNECH	Diffuse Idiopathic Pulmonary Neuroendocrine Cell Hyperplasia
DLEC1	Deleted in Lung and Esophageal Cancer 1
DLR	Dual-Luciferase Reporter
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
DNMT	DNA Methyltransferase
DPX	Distrene-Plasticiser-Xylene
EBV	Epstein-Barr Virus
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic Acid
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
EET	Regioisomeric Epoxyeicosatrienoic Acids
EMT	Epithelial-Mesenchymal-Transition
EPGN	Epigen
EREG	Epiregulin
ERα	Estrogen Receptor Alpha
EtBr	Ethidium Bromide
EtOH	Ethanol
FBXW7	F-Box and Wd Repeat Domain Containing 7, E3 Ubiquitin Protein Ligase
FCS	Foetal Calf Serum
FDA	Food And Drug Administration
FDG-PET	Fluorodeoxyglucose Positron-Emission Tomography
FDR	False Discovery Rate
FE	Feature Extraction
FFPE	Formalin-Fixed Paraffin-Embedded
FGF	Fibroblast Growth Factor
FGFR3	Fibroblast Growth Factor Receptor 3
FHIT	Fragile Histidine Triad

GFP	Green Fluorescent Protein
GO	Gene Ontology
GPRC5C	G Protein-Coupled Receptor, Family C, Group 5, Member C
GPX2	Glutathione Peroxidase 2
GRM8	Glutamate Receptor, Metabotropic 8
GSTT1	Glutathione S-Transferase Theta 1
HBEGF	Heparin-Binding EGF-Like Growth Factor
HDM2	Human Double Minute 2
HCLS1	Hematopoietic Cell-Specific Lyn Substrate 1
hGH	Human Growth Hormone
HMGA2	High Mobility Group AT-Hook 2
HNC	Head and Neck Cancer
HNSCC	Head and Neck Squamous Cell Carcinoma
HOM-C	Homeotic Gene Complex
HPV	Human Papillomavirus
HRAS	Harvey Rat Sarcoma Viral Oncogene Homolog
HRP	Horseradish Peroxidase
IARC	International Agency For Research On Cancer
IGF2R	Insulin-Like Growth Factor 2 Receptor
IGFBP3	Insulin-Like Growth Factor Binding Protein 3
IHC	Immunohistochemistry
IRF6	Interferon Regulatory Factor 6
JAG2	Jagged 2
KGM	Keratinocyte Growth Medium
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
KRT5	Keratin 5
L1CAM	L1 Cell Adhesion Molecule
LATS1	Large Tumour Suppressor Kinase 1
LC	Lung Cancer
LCM	Laser Capture Microdissection
LEC	Lung Epithelial Cell
LOH	Loss of Heterozygosity
LTF	Lactotransferrin
MAPK	Mitogen-Activated Protein Kinase
miR	MicroRNA
MLH1	Mutl Homolog 1
MLL	Mixed Lineage Leukaemia
MMP14	Matrix Metalloproteinase-14

MRI	Magnetic Resonance Imaging
mRNA	Messenger RNA
MTS-1	Multiple Tumour Suppressor 1
MUC1	Mucin 1
NAPSA	Napsin A Aspartic Peptidase
NGS	Next Generation Sequencing
NPC	Nasopharyngeal Carcinoma
NSCLC	Non-Small Cell Lung Carcinoma
NTS	Neurotensin
OCLN	Occludin
OK	Oral Keratinocyte
OPL	Oral Pre-Malignant Lesion
ORF	Open Reading Frame
PARK2	Parkin Rbr E3 Ubiquitin Protein Ligase
PBS	Phosphate Buffer Saline
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
PDGFRA	Platelet-Derived Growth Factor Receptor, Alpha Polypeptide
PDGFRL	Platelet-Derived Growth Factor Receptor-Like
PET	Positron Emission Tomography
PI3-K	Phosphatidylinositol 3-Kinase
PIGN	Phosphatidylinositol Glycan Anchor Biosynthesis, Class N
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase, Catalytic Subunit Alpha
PRDX6	Peroxiredoxin 6
pRb	Retinoblastoma Protein
PTEN	Phosphatase and Tensin Homolog
qPCR	Real-Time Quantitative Polymerase Chain Reaction
RARβ	Retinoic Acid Receptor Beta
Rb	Retinoblastoma
RGS17	Regulator of G-Protein Signaling 17
RhoC	Ras Homolog Gene Family Member C
RIN	Rna Integrity Number
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
RPMI	Roswell Park Memorial Institute
rRNA	Ribosomal RNA
SASH1	Sam and Sh3 Domain Containing 1
SCC	Squamous Cell Carcinoma

SCLC	Small Cell Lung Carcinoma
SD/CIS	Squamous Dysplasia/Carcinoma In Situ
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
shRNA	Short Hairpin RNA
siRNA	Short Interfering RNA
SNP	Single Nucleotide Polymorphisms
SOC	Super Optimal Broth with Catabolite Repression
SOX2	Sry (Sex Determining Region Y)-Box 2
TAE	Tris Acetate-EDTA Buffer
TBST	Tris Buffered Saline with Tween-20
TCF21	Transcription Factor 21
TE	Trypsin/Ethylenediaminetetraacetic Acid
TEMED	Tetramethylethylenediamine
TET1	Ten-Eleven Translocation 1
TF	Transcription Factor
TGFα	Transforming Growth Factor Alpha
TKI	Tyrosine Kinase Inhibitor
Tm	Melting Temperature
TMA	Tissue Microarray
TNM	Tumour/Nodes/Metastasis
TOPO	Topoisomerase
TP53	Tumour Protein P53
TP63	Tumour Protein P63
TSG	Tumour Suppressor Gene
TSS	Transcription Starting Site
TUSC3	Tumor Suppressor Candidate 3
TWIST1	Twist Basic Helix-Loop-Helix Transcription Factor 1
UCSC	University of California Santa Cruz
uPAR	Urokinase Plasminogen Activator Receptor
UV	Ultraviolet
VEGFA	Vascular Endothelial Growth Factor A
w/v	Weight/Volume
WT	Wild Type
WHO	World Health Organization
ΔCT	Delta Ct

PUBLICATIONS

Hakami, F., Darda, L., Stafford, P., Woll, P., Lambert D., Hunter, K. (2013). The role of HOXD10 in the development and progression of head and neck squamous cell carcinoma. *British Journal of Cancer*.

Hunter, K., Darda, L., Hakami, F., Morgan, R., Murdoch, C. and Lambert, D. (2013). The role of HOXB9 and miR-196a in head and neck squamous cell carcinoma. *International Journal of Cancer*.

CHAPTER 1: INTRODUCTION

1.1 Cancer

Cancer can be defined as the out-of-control growth of malignant cells occurring in any part of the body, which impairs the normal function of that part and can spread to other body parts. According to a recent global statistical study, cancer is one of the main causes of death worldwide (Ferlay et al., 2010). Many factors are believed to cause or facilitate cancer initiation and progression. These include chemicals, alcohol, sunlight exposure, radiation and viruses. Although the mechanism of cancer onset driven by any of these factors remains largely unclear, it is known that cancer develops when cells are affected by genetic abnormalities including the activation of oncogenes, which are genes able to induce cancer, and inactivation of tumour suppressor genes, which are genes that inhibit uncontrolled cell proliferation (Lodish, 1999, Mehrotra and Yadav, 2006). Such genetic alterations and their subsequent effects are what drive cells into acquiring the main hallmarks of cancer; continuous and uncontrolled proliferation, growth signal autonomy, insensitivity to anti-growth signals, avoidance of apoptosis, invasion and metastasising capabilities and induction of angiogenesis (Hanahan and Weinberg, 2011).

1.2 Head and neck cancer

Head and neck cancer (HNC) describes all malignancies that occur in the tissues of the mouth, nose, throat, sinuses, thyroid glands, salivary glands and lymph nodes in the neck (Figure 1.1). Therefore, it includes the following: oral and oropharyngeal cancer, nasal and paranasal sinuses cancer, nasopharyngeal cancer, laryngeal cancer, thyroid cancer and salivary gland cancer (Marur and Forastiere, 2008). Cancers of the remaining parts of the head and neck, such as the brain, eyes and ears are not usually included in HNC cancer classification. The mouth is the most common site for HNC cancer, which can develop in the lips, tongue, buccal mucosa, floor and roof of the mouth, cheek and gum. Due to the anatomical site, HNC usually affects important functions including taste, speech, swallowing and smell. Therefore, HNC treatment, which typically includes surgery, radiation and chemotherapy, might cause an obvious subsequent impairment to the quality of the patient's life. The treatment plan depends on different factors such as the age of the patient, the stage of the tumour and its exact location. Thus, the management of HNC is very complex and requires a close cooperation between the oncologists, pathologists, surgeons and radiologists (Gregoire et al., 2010). Like most cancers, the efficiency of HNC treatment depends mainly on the disease stage, with the earlier the stage, the better the chance for prognosis. However, two out of three patients with HNC present with an advanced stage (Jemal et al., 2007). This might be due to the fast progression of the tumour while its symptoms are still mild and bearable. In general, patients at the early stages of HNC present with non-specific symptoms and minimal physical findings.

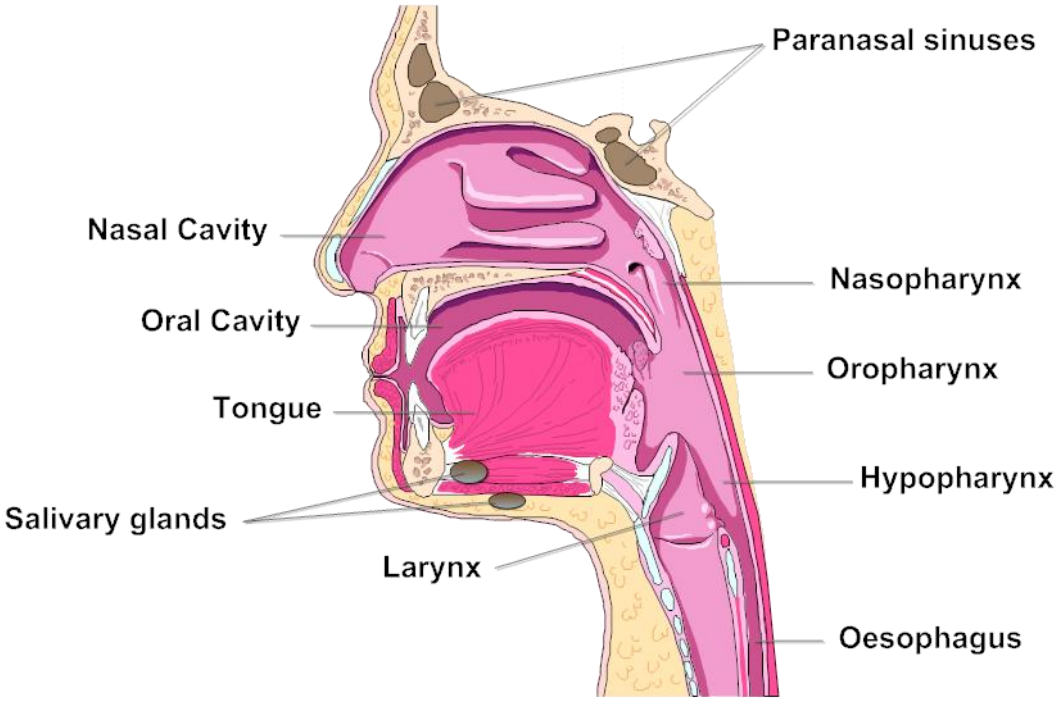


Figure 1.1: A cross-section diagram of the head and neck.

HNC describes all malignancies that occur in the tissues of the mouth, nose, throat, sinuses, thyroid glands, salivary glands and neck lymph nodes. *This diagram was generated by SmartDraw software (available on http://www.smartdraw.com/)*.

Clinical features of this cancer depend mainly on the tumour site. Exophytic masses which bleed, nonhealing ulcers and extensive white or red patches are common in patients suffering from a cancer in the oral cavity. Cancer that occurs in the oropharynx may be accompanied by sore throat, odynophagia and dysphagia, while otalgia and hoarseness are common in patients with a cancer in the hypopharynx (reviewed by Goy et al., 2009). The hoarseness is persistent in those with a cancer in the larynx in addition to shortening of breath, or dyspnoea. Symptoms of the nasal cavity and paranasal sinuses cancers include sinusitis, nasal airway obstruction and epistaxis (Marur and Forastiere, 2008, Harrison et al., 2009).

Due to the diversity of anatomical structures it contains, the head and neck region includes a large number of histologies. Accordingly, a wide variety of tumour types have been classified in the head and neck. For instance, more than fifty different tumours were distinguished in the nasal cavity and paranasal sinuses only according to the classification of the World Health Organization (WHO) (Barnes, 2005). However, squamous cell carcinoma (SCC), which originates from the mucosal squamous epithelium, is by far the dominant malignant tumour type accounting for more than ninety percent of all HNC cases (Cooper et al., 2009). Therefore, head and neck squamous cell carcinoma (HNSCC) will be the focus of the next discussion and this project in general. HNSCC arises as a result of a multi-step pathological process initiated by a premalignant lesion. Most common premalignant lesions are leukoplakia, which appear clinically as a white plaque on a mucous membrane, erythroplakia, which is a red plaque on a mucous membrane, or mixed white/red speckled leukoplakia lesion. Erythroplakia is more likely to transform into a malignancy than leukoplakia, and is mostly associated with histopathologic alterations, such as dysplasia (Harrison et al., 2009). Although, non-dysplastic lesions may still develop and form a tumour (Warnakulasuriya et al., 2008), dysplasia, which describes an altered epithelium and loss of cell maturation, is typically the onset of HNSCC progression and development. It can be graded into the following:

- Mild dysplasia: variations in the size and shape of the keratinocytes but limited to the lower third of the epithelium.
- Moderate dysplasia: disturbance extended into the middle third of the epithelium.
- Severe dysplasia: two-thirds of the epithelium showing architectural disturbance.

Severe dysplasia can progress into an in-situ carcinoma presented by almost full thickness of the epithelium and marked cellular atypia. This leads eventually to an invasive carcinoma, which is characterised by an irregular interface between epithelium and connective tissue with pointed projections and/or epithelial islands in the connective tissue separated from the surface epithelium (Figure 1.2) (Warnakulasuriya et al., 2008, Vermorken and Specenier, 2010).

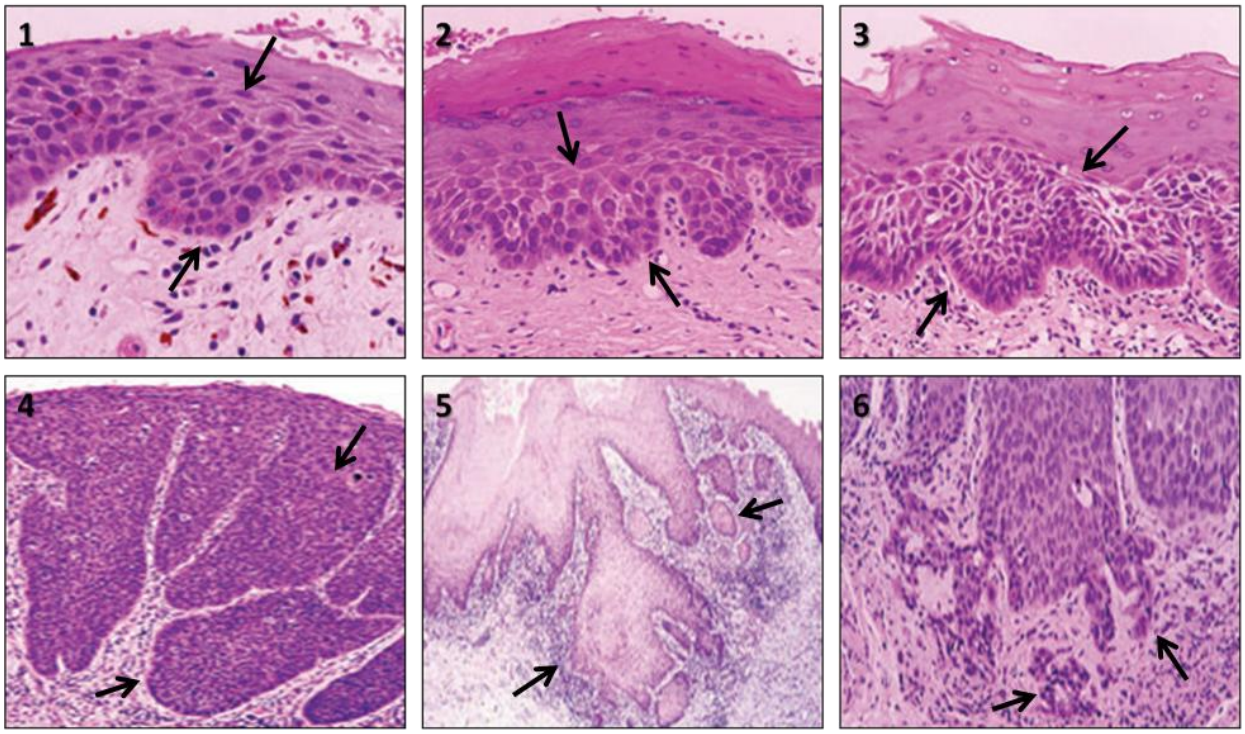


Figure 1.2: The progression of epithelial precursor lesions in HNSCC.

(1) Mild dysplasia: variations in the size and shape of the keratinocytes but limited to the lower third of the epithelium (x100 magnifications). (2) Moderate dysplasia: disturbance extended into the middle third of the epithelium (x100 magnification). (3) Severe dysplasia: two-thirds of the epithelium showing architectural disturbance (x100 magnification). (4) In-situ carcinoma: almost full thickness of the epithelium and marked cellular atypia (x20 magnification). (5) & (6) Invasive carcinoma: epithelial islands in the connective tissue separated from the surface epithelium and/or irregular interface between epithelium and connective tissue with pointed projections (x20 and x40 magnification, respectively).

The above haematoxylin and eosin stained tissue sections photomicrographs courtesy of Dr K Hunter.

1.2.1 Epidemiology and incidence

HNSCC ranks as the sixth most common type of cancer worldwide affecting approximately half a million people every year with a fifty percent chance of surviving for five years (Conway et al., 2009, Ferlay et al., 2010). In the UK, summing up all subsites affected by HNSCC makes it the fifth most common cancer (Cancer Research UK, 2010). Despite the global awareness of HNSCC possible risk factors, the incidence of HNSCC around the world is still increasing, especially in young patients, who were recently found to have greater disease-free survival (Lassig et al., 2013). HNSCC is more common in males than females. The sixth decade of life is a median age of HNSCC diagnosis (Mehanna et al., 2010, Hartwig et al., 2012).

Oral cancer is the most common type of HNSCC affecting nearly 300,000 people every year worldwide (Lippman et al., 2005). The highest rates of oral cancer are in Melanesia, South-Central Asia, Central and Eastern Europe (Jemal et al., 2011). Larynx and pharynx are the second and third most common sites for HNSCC accounting for 160,000 and 65,000 annual new cancer cases, respectively, around the world (Mehanna et al., 2010). The high morbidity and mortality rates caused by HNSCC are mostly due to local tumour invasion and metastasis (Harrison et al., 2009). Moreover, despite advancements in surgical procedures, chemotherapy methods and other molecular-based therapies, recurrence remains a major challenge and HNSCC-associated mortality rates have remained largely unchanged for decades (Fung and Grandis, 2010, Abuzeid et al., 2011, Gonzalez-Moles et al., 2013).

1.2.2 Diagnosis and staging

An accurate histological examination and tumour classification are essential for planning the most efficient treatment regime, predicting any metastatic potential or secondary tumour formation and evaluation of the treatment efficacy. A variety of basic and advanced methods are used in the clinic to diagnose HNSCC. These includes physical examination to detect mucosal abnormalities, biopsies for histological examination, endoscopy, radiographic imaging, computed tomography (CT) and three dimensional (3D) CT scan with reconstruction, positron emission tomography (PET), magnetic resonance imaging (MRI) and MR spectroscopy (Sengupta, 2012). Once diagnosis has been established, the tumour can be classified according to the Tumour-Node-Metastasis (TNM) staging system, which is the universal system used to describe all malignant tumours including all types of HNSCC, and is periodically updated by the American Joint Committee for Cancer (AJCC).

The TNM system was established by Pierre Denoix during the 1940s (Denoix, 1944), then reviewed and updated several times over the years until its latest version in 2010 (Sobin et al., 2011). It combines and describes three parameters: 1) the anatomic extent of the primary

tumour (divided into four stages: T1 to T4), 2) the involvement level of regional lymph node metastasis (divided into four stages: N0 to N3) and 3) the presence of distant metastasis (divided into two stage: M0 and M1) (Sobin et al., 2011). According to this classification, there are 32 possible stages combinations. However, for the sake of simplicity, they are cross combined and clustered into four main stages: I, II, III and IV. Each main grouped stage might be further subcategorised into a, b and c subgroups according to some factors (Takes et al., 2010). For example, stage IV of lip and oral cancer has three subdivisions: IVa, IVb and IVc, identified by the ability to remove the tumour surgically (Table 1.1). The TNM system can be combined with supplementary information, such as histological grading G1 to G4, where G1 means a well-differentiated tumour, while G4 stands for undifferentiated tumour, and certainty factor C1 to C5, which reflects staging validity; C1 means that the standard diagnostic methods were used for the classification, while C5 means that the staging relied on autopsy (van der Schroeff and Baatenburg de Jong, 2009). It is worth mentioning that the TNM system can be a clinical staging system (cTNM), which is a classification that relies on a clinical examination, or a pathological staging system (pTNM), which is a classification driven from a histopathological examination (van der Schroeff and Baatenburg de Jong, 2009).

1.2.3 Aetiology and risk factors

1.2.3.1 Smoking and alcohol

A recent epidemiological study has shown a strong relationship between tobacco smoking and alcohol consumption and the risk of developing a HNSCC. Moreover, these major risk factors were found to associate with around seventy percent of all analysed HNSCC cases (Hashibe et al., 2009). In addition, the risk is higher for those who combine both factors in their normal lifestyle than others who smoke only or drink alcohol only (Ansary-Moghaddam et al., 2009, Hashibe et al., 2009). Furthermore, according to a previous study, heavy smokers have a 21-fold increase in their risk of developing HNSCC, while it is a fivefold increase for those who are heavy drinkers, and a 48-fold increase for combination of high tobacco and alcohol consumption (Rodriguez et al., 2004).

One of the possible reasons why tobacco is a highly suspected agent in causing HNSCC is the pro-carcinogens it contains, such as nitrosamines and benz-(a)-pyrene. These chemicals result in changes to guanine nucleotides and drive mutations in different genes, including the tumour suppressor gene TP53, which is frequently mutated in HNSCC (Lacko et al., 2009). In the case of alcohol, its effects are not yet fully described. However, it might have an impact on DNA integrity when it is metabolised to acetaldehyde, which is a DNA damaging agent (Poschl and Seitz, 2004).

Table 1.1: TNM staging (7th edition) for cancers of the lips and oral cavity.

Taken from (Edge and Compton, 2010).

Primary tumor (T)	
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Tis	Carcinoma in situ
T1	Tumor ≤ 2 cm in greatest dimension
T2	Tumor > 2 cm but ≤ 4 cm in greatest dimension
T3	Tumor > 4 cm in greatest dimension
T4	(lip) Tumor invades through cortical bone, inferior alveolar nerve, floor of mouth, or skin of face, ie, chin or nose[a]
T4a	Moderately advanced local disease[a] (lip) Tumor invades through the cortical bone, mouth, or skin of the face (ie, chin or nose) (oral cavity) Tumor invades adjacent structures (eg, through cortical bone [mandible or maxilla] into the deep [extrinsic] muscle of the tongue [genioglossus, hyoglossus, palatoglossus, and styloglossus], maxillary sinus, or skin of the face)
T4b	Very advanced local disease Tumor involves masticator space, pterygoid plates, or skull base and/or encases internal carotid artery

Regional lymph nodes (N)	
NX	Regional nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in a single ipsilateral lymph node, ≤ 3 cm in greatest dimension
N2	Metastasis in a single ipsilateral lymph node, > 3 cm ≤ 6 cm in greatest dimension; or in multiple ipsilateral lymph nodes, none > 6 cm in greatest dimension; or in bilateral or contralateral lymph nodes, none > 6 cm in greatest dimension
N2a	Metastasis in a single ipsilateral lymph node, > 3 cm but ≤ 6 cm in greatest dimension
N2b	Metastasis in multiple ipsilateral lymph nodes, none > 6 cm in greatest dimension
N2c	Metastasis in bilateral or contralateral lymph nodes, none > 6 cm in greatest dimension
N3	Metastasis in a lymph node, > 6 cm in greatest dimension

[a] Superficial erosion alone of bone/tooth socket by gingival primary is not sufficient to classify a tumor as T4.

Distant metastases (M)	
M0	No distant metastasis
M1	Distant metastasis

Stage grouping			
Stage 0	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Stage IVA	T4a	N0	M0
	T4a	N1	M0
	T1	N2	M0
	T2	N2	M0
	T3	N2	M0
	T4a	N2	M0
Stage IVB	Any T	N3	M0
	T4b	Any N	M0
Stage IVC	Any T	Any N	M1

1.2.3.2 Human papillomavirus (HBV)

Another important risk factor sometimes considered an oncogenic factor in HNSCC is the HPV (Gillison and Shah, 2001). This multi-type small non-enveloped DNA virus causes a wide range of type-dependent infections ranging from benign skin and mucosal lesions to more aggressive malignant lesions (Giovannelli et al., 2002). HPV DNA was detected for the first time in HNSCC tissues in 1985 (De Villiers et al., 1985), and since then, it has been detected in a variable proportion of HNSCC cases, ranging from less than ten percent in some studies to almost 100% in others (Langendijk and Psyrri, 2010). Among the 120 identified types of HPV, HPV16 was the most frequently detected genotype in HPV-positive HNSCC accounting for 87% of the HPV-associated pharyngeal carcinomas and approximately seventy percent of the HPV-associated oral and laryngeal carcinomas (reviewed by Kreimer et al., 2005).

The proposed mechanism by which HPV plays a carcinogenesis role starts when its genome becomes integrated into the host DNA (Figure 1.3). The HPV genome replicates as a circular double-stranded DNA, or what is called an 'episome' (LaPorta and Taichman, 1982). The HPV genome is composed of 'early' coding genes, *E1*, *E2*, *E4*, *E5*, *E6* and *E7*, 'late' coding genes, *L1* and *L2*, and a non-coding long control region (*LCR*) (McLaughlin-Drubin and Munger, 2009). Each gene encodes a protein with a specific function, for example, *E1* encodes a helicase enzyme that acts on initiating the viral replication, while E2 protein controls the expression of *E6* and *E7* genes (Hughes and Romanos, 1993). Disruption of the *E2* sequence allows HPV genome/host DNA integration, which leads to uncontrolled over-expression of E6 and E7, which are oncoproteins (Bechtold et al., 2003). E6 and E7 proteins are able to degrade and antagonise the TP53 tumour suppressor protein and retinoblastoma protein (pRb), respectively, leading to apoptosis inhibition and an increase in DNA synthesis and cell proliferation (Munger et al., 1989, Scheffner et al., 1990, Scheffner et al., 1993, Boyer et al., 1996). Eventually, this will increase the risk of developing carcinoma in the infected cells.

Although recombination between HPV and chromosomal DNA has been reported in the progression of oral carcinoma (Balderas-Loaeza et al., 2007), linking HPV to the carcinogenesis of HNSCC has been a point of controversy because of the variable and inconsistent detection of HPV DNA in HNSCC patients. In various studies, detection of HPV in different head and neck sites of HNSCC cases ranged from complete absence to complete presence in HNSCC patients (Terai et al., 1999, Giovannelli et al., 2002, Campisi et al., 2007, Chaudhary et al., 2009). However, this variability in the prevalence of HPV in head and neck tumours may be due to using different detection techniques (Boy et al., 2006), or it might be attributed to the differences in the epidemiological background studies (Machado et al., 2010).

In general, there has been an upward trend in the incidence of HPV-positive versus HPV-negative HNSCC cases in many parts of the world (Psyrri et al., 2009). The change in sexual behaviours that leads to the transmission of the virus to the oral cavity and then to the

other internal sites of the head and neck might be the main factor behind this increase (Chaturvedi et al., 2008). This suggestion is further supported by the fact that almost 99% of cervical squamous cell carcinoma incidents are associated with HPV infection (Walboomers et al., 1999, Munoz et al., 2003) and that the cervical mucosa share many histological similarities with the mucosa of the head and neck, particularly in the tonsil (Georgieva et al., 2009).

1.2.3.3 Other HNSCC aetiological factors

Betel nut chewing might not be a major HNSCC risk factor globally, but it can be a unique aetiological factor in some parts of the world such as India and other Southeast Asian countries, where it is associated with HNSCC (Zain et al., 1999, Muttagi et al., 2012). Infection with Epstein-Barr virus (EBV) is common and has been reported to associate with a variety of diseases such as infectious mononucleosis, hairy leukoplakia, inflammatory pseudotumours, Burkitt's lymphoma and, most importantly, nasopharyngeal carcinoma (NPC) (Perez-Ordonez, 2007). Exposure to sunlight, which contains ultraviolet rays, is another possible cause of HNSCC affecting the lips, especially in the case of outdoor workers (Håkansson et al., 2001). Dietary behaviour is also a suspected aetiological factor. Some studies have shown that frequent consumption of preserved meat might increase the incidence of HNSCC (Farrow et al., 1998), while regular fruit and vegetables consumption might have the opposite effect (Boeing et al., 2006, Freedman et al., 2008). Finally, some diseases that involve genetic abnormalities able to affect DNA repair mechanism or other important cellular pathways may contribute to the risk of developing HNSCC. Examples of this are Lynch-II syndrome, Bloom syndrome, ataxia telangiectasia, Li-Fraumeni syndrome and Fanconi's anaemia (reviewed by Baez, 2008). Patients of the latter disease are highly susceptible to develop HNSCC if they survive bone marrow transplantation (Rosenberg et al., 2008).

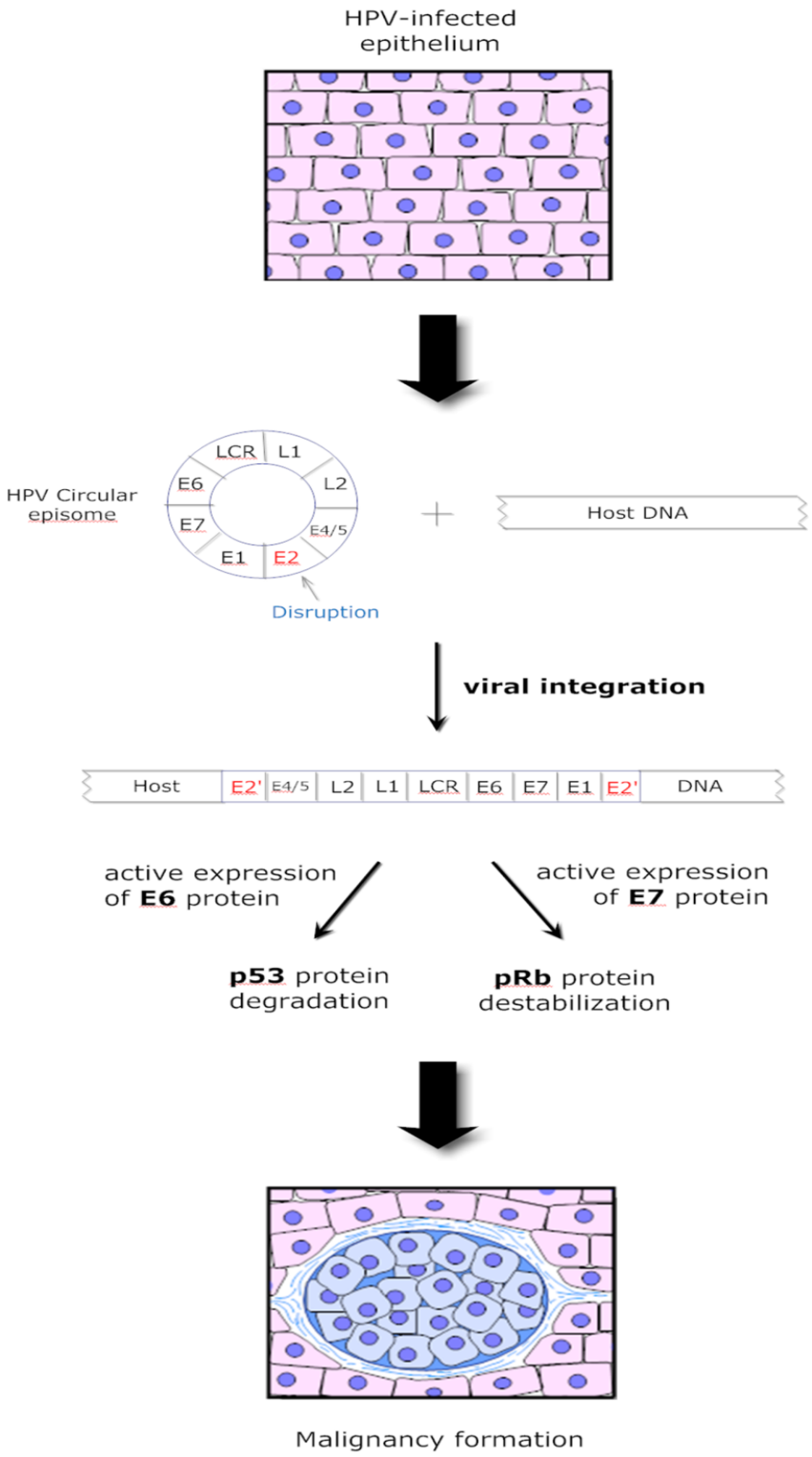


Figure 1.3: Proposed mechanism of HPV oncogenic role in HNSCC.

HPV genome composed of different genes including *E1*, *E2*, *E4*, *E5*, *E6* and *E7*. Disruption of *E2* sequence in the HPV circular episome allows HPV genome/host DNA integration, leading to uncontrolled over-expression of E6 and E7 proteins. These proteins degrade or destabilise TP53 and pRb, respectively, which allows unchecked cell growth and possible carcinoma formation. *This diagram was generated by SmartDraw software (available on http://www.smartdraw.com/).*

1.2.1 Genetic alterations and molecular markers

The accumulation of genetic imbalances including over-expression, down-regulation, deletion and mutation of some oncogenes or tumour suppressor genes (TSGs) is believed to drive the progression of HNSCC through different pathways (reviewed by Stadler et al., 2008). Both gene deletion and mutation occur at a chromosomal level. Studies using cytogenetic analysis have demonstrated that a variety of chromosomal abnormalities are frequently detected in HNSCC. These include gain of chromosome 1q, 3q, 7p, 8q, 5p, 9q, 11q, q13, 14q, 16q and 20q, and loss of chromosome 3p, 5q, 8p, 9p, 15p, 18q, 21q and 22q (Van Dyke et al., 1994, Bockmühl et al., 2000, Bérgamo et al., 2005, Bauer et al., 2008). The effect of chromosomal changes is reflected in the expression of several genes located within their particular DNA sequences. For example, tumour suppressor genes *p16* (*INK4A*) & *p14* (*ARF*) and *TP53* are affected by the loss of heterozygosity (LOH) at their chromosomal loci 9p21 and 17p13, respectively (Gonzalez et al., 1995, Ohta et al., 2009). Although the exact mechanism behind such structural changes in these chromosomes and other chromosomes in different cancers are not fully described, a very recent study might have provided new understanding of how abnormal chromosomal changes originate. The findings of this study revealed that the collisions between the machineries of DNA duplication and RNA transcription are the reason behind many of the chromosomal abnormalities detected in different types of tumours (Barlow et al., 2013). However, why such collisions are avoidable in normal cells is still under investigation.

The abnormal over-expression and down-regulation of non-mutated or mutated genes are caused by either pre-transcriptional or post-transcriptional regulation mechanisms. Epigenetic changes, such as DNA methylation and chromatin remodelling, are examples of pre-transcriptional control of gene expression. Such mechanisms were found to be implicated in the development and progression of HNSCC. An example of this is the aberrant promoter hypermethylation of the tumour suppressor genes *RAR-β2* (Youssef et al., 2004), *CDKN2A* (Viswanathan et al., 2003) and *DCC* (Carvalho et al., 2006). The non-coding small microRNAs (miRs), are the main post-transcriptional modulators. Through controlling the translation of the mRNA of their target genes, several miRs were proved to be involved in HNSCC pathogenesis and can be either TSGs or oncogenes. miR-21, miR-31 and miR-504 are examples of HNSCC oncogenic miRs that target several TSGs and have an impact on a number of cellular pathways (Tu et al., 2013). Moreover, some miRs, such as miR-323, miR-345, miR-371, miR-181a, miR-342 and miR-326, were shown to be HNSCC specific (Lindenbergh-van der Plas et al., 2013). On the other hand, a microRNA, such as miR-16, can be a TSG by targeting the oncogene *BCL2* (Cimmino et al., 2005).

A model for the sequential pathological genetic progression of HNSCC carcinogenesis was proposed many years ago (Califano et al., 1996). However, it needs updating since many genetic alterations have been identified since then, especially with the help of highly advanced

techniques, such as microarrays and next generation sequencing (NGS). NGS not only validated and assessed the previously identified HNSCC genetic abnormalities, but also revealed abnormalities in many genes not previously implicated in this type of cancer, such as *NOTCH1*, *FBXW7*, *IRF6* and *TP63* genes (Agrawal et al., 2011, Stransky et al., 2011). Moreover, such advanced techniques allowed correlation between the detected genetic abnormalities and the presence or absence of some known HNSCC aetiological factors, such as HPV virus (Nichols et al., 2012, Lechner et al., 2013).

Through the identification of the recurrent gene expression signatures, HNSCC was classified into four different molecular subtypes: basal, mesenchymal, atypical and classical (Walter et al., 2013). Interestingly, there was a major similarity between these subtypes and the molecular subtypes of lung cancer. This similarity might be related to their shared exposure to smoking-associated mutagens. *EGFR*, *TP63* and *COL17A1* are examples of genes associated with basal subtypes, while the differential expression of other genes including *ACTA1*, *TWIST1* and *PDGFRA* is characteristic of the mesenchymal subtype. Genes of atypical subtype include *CDKN2A*, *ALDH3A1* and *FGFR3*, while *PIK3CA*, *SOX2* and *GPX2* are some of the genes associated with the classical subtype (Walter et al., 2013). In general, it might be that the most frequently detected genetic alteration or aberrant expression in HNSCC occurs in tumour suppressor genes *TP53*, *PTEN*, *NOTCH1* and *CDKN2A*, and oncogenes *Cyclin D1*, *HRAS*, *EGFR* and *PIK3CA* (Hunt et al., 2013) (some of these will be discussed briefly in the next sections).

1.2.1.1 Tumour protein p53 (TP53)

The short arm of chromosome 17 contains the tumour suppressor gene *TP53*, which encodes a protein that has the function of sustaining the genomic stability in addition to other important roles in cellular proliferation, differentiation and apoptosis. TP53 is crucial to guarantee a normal cell progeny by protecting DNA from damage or directing the cell to apoptosis if this damage is irreversible (McBride et al., 1986, Weinberg, 1991). Mutated *TP53* is the most commonly detected genetic alteration among tumour suppressor genes across all cancers (Olivier et al., 2002). Around fifty percent of HNSCC patients have a mutated *TP53*, which was found to be associated with the progression of the malignancy from its pre-malignant phase to its invasive condition (Boyle et al., 1993, Forastiere et al., 2001). The fact that many pre-malignant lesions in the head and neck region harbour *TP53* mutations suggests that this genetic alteration can be an early event in the pathology of HNSCC (Boyle et al., 1993). Moreover, *TP53* mutations increase the risk of disease recurrence following surgical therapy (reviewed by Brennan and Sidransky, 1996), and have been correlated with a weak response to treatment and, therefore, poorer survival (Temam et al., 2000).

Mutations of *TP53* are highly associated with some aetiological factors such as high exposure to alcohol and tobacco (Zhang et al., 2000). In addition, E6 oncoprotein product of HPV virus is believed to work on suppressing TP53 functions (Scheffner et al., 1990). The nature of *TP53* mutations are divergent and have variable effects on the functions of the encoded TP53 protein depending on the position of the mutation and its impact on the protein structure and its binding ability. For example, alterations in the genetic sequence that lead to complete blocking of its DNA-binding were found to accelerate tumour progression and affect treatment response (Poeta et al., 2007). Such alterations were considered "disruptive" compared with "non-disruptive" mutations, which have much less impact on the mutated TP53 function and generally have better clinical outcomes (Poeta et al., 2007).

1.2.1.2 Cyclin-Dependent Kinase Inhibitor 2A (CDKN2A)

CDKN2A locus in chromosome 9p21 encodes tumour suppressor proteins *p14* (*ARF*) and, more importantly, *p16* (*INK4A*) (Stott et al., 1998). The p16 protein, also known as multiple tumour suppressor 1 (*MTS-1*), is a cyclin-dependent kinase inhibitor of cyclin-protein kinase complexes, such as CDK4 and CDK6 (Takeuchi et al., 1995), and a suppressor of Cyclin D1 (Kamb et al., 1994). Therefore, the loss of an active p16 might results in Cyclin D1 amplification (reviewed by Arora et al., 2012), which leads to reduction in cancer cells requirements for growth factors and enhances their resistance to some chemotherapeutics (Nakashima et al., 2005). Restraining CDK4/6 activity by p16 leads to inhibition of the phosphorylation of retinoblastoma protein (pRb) and, eventually, impairment of G1-to-S phase transition of the cell cycle (Serrano et al., 1996, Li et al., 2011). Therefore, inactivation of p16 promotes uncontrolled growth. Moreover, p16 inactivation by hypermethylation was found to induce cell migration and invasion of HNSCC cells (Su et al., 2010), while restoring its expression represses angiogenesis in some other cancers (Harada et al., 1999). Inducing the expression of a wild-type *p16* gene through engineered adenoviruses is also able to arrest the growth of HNSCC cells (Rocco et al., 1998).

Genetic alteration and abnormal expression of *p16* is a very early event in HNSCC, which occurs by mutation, complete deletion of its locus, or aberrant hypermethylation of its promoter (Reed et al., 1996, Ai et al., 2003, Demokan et al., 2012). Down-regulation of p16 expression is detected in 55% to 89% of all HNSCC cases (Higuchi et al., 2007). Tumours of the larynx have shown weaker p16 expression than other tumour types of the head and neck cancer (Yuen et al., 2002). Tobacco and alcohol are considered very important aetiologic factors that cause p16 inactivation in HNSCC (Ai et al., 2003). However, p16 over-expression is a surrogate clue of HPV infection because HPV oncorprotein products can bind and interfere with pRb function leading to the loss of p16/pRb negative feedback control, which results in p16 up-regulation (Andl et al., 1998). Low levels of p16 protein correlated with higher risk of tumour recurrence, metastasis and poor survival in HNSCC patients (Bova et al., 1999, Danahey et al.,

1999, Kumar et al., 2008b), while high levels of p16 might be an indicator of better prognosis in HPV-positive HNSCC patients (Lassen et al., 2009).

The p14 protein shares some similarities with p16 in controlling the cell cycle but through different mechanisms. Through sequestering HDM2 protein, which is an E3 ubiquitin ligase, p14 abrogates its role in attenuating TP53 activity and subsequently allows TP53-dependent apoptosis and growth arrest (Pomerantz et al., 1998). Therefore, inactivity of p14 leads to indirect loss of TP53 function. Moreover, p14 is capable of repressing cell proliferation in a mechanism independent of TP53 by delaying the S-phase of the cell cycle (Korgaonkar et al., 2002, Yarbrough et al., 2002). Genetic aberration of *p14* is frequently detected in HNSCC (Shintani et al., 2001, Weber et al., 2002), however, free-disease survival is improved when active p14 is detected (Kwong et al., 2005).

1.2.1.3 Epidermal growth factor receptor (EGFR)

EGFR gene is located on chromosome 7p12 (Kondo and Shimizu, 1983), and encodes a type I receptor tyrosine kinase that is involved in diverse signalling pathways responsible for cell proliferation, differentiation, migration in addition to angiogenesis and invasion (Wang et al., 2004, Kalyankrishna and Grandis, 2006). EGFR structure is composed of three parts: 1) an extracellular ligand-binding domain, 2) a transmembrane segment and 3) an intracellular kinase domain with five autophosphorylation sites (Sharafinski et al., 2010). This receptor can be activated by many ligands, including EGF, transforming growth factor-alpha (TGFα), heparin-binding EGF-like growth factor (HBEGF), amphiregulin (AREG), epiregulin (EREG) and epigen (EPGN) (Schneider and Wolf, 2009). Downstream EGFR signalling triggers several pathways including mitogen-activated protein kinase (MAPK) pathway resulting in producing anti-apoptotic proteins such as BCl-X_2 and suppressing the apoptosis-inducing BAD protein (Yarden, 2001, Kari et al., 2003). Phosphatidylinositol 3-kinase (PI3-K)/protein kinase B (Akt) pathway is also activated by EGFR and leads to activating cell-survival signals, inhibition of TP53 TSG and regulating cell growth and motility (Brazil et al., 2004, Peng et al., 2006, Ratushny et al., 2009).

Up-regulation of EGFR is one of the early markers of HNSCC (Grandis and Tweardy, 1993), occurring in around ninety percent of all head and neck malignancies compared to the corresponding normal tissues (Rubin Grandis et al., 1998). However, laryngeal tumours showed lower level of EGFR expression when compared to tumours that occur in the oral cavity and oropharynx (Takes et al., 1997). The increase in EGFR expression correlates with the progression of the disease from dysplasia to invasive cancer (Shin et al., 1994). Moreover, over-expression of EGFR was found to be associated with the increase in tumour growth, metastasis and poor prognosis accompanied by resistance to chemotherapy and radiotherapy (Rubin Grandis et al., 1998, Ang et al., 2002, Normanno et al., 2006). Based on this, and the fact that

EGFR is a surface receptor that can be antagonised and blocked using specific antibodies, it became a goal to develop many EGFR-targeted therapies as a treatment of HNSCC and other cancers. In 2006, the Food and Drug Administration (FDA) approved the first anti-EGFR antibody, cetuximab, to be used in combination with radiotherapy as a treatment regimen for HNSCC patients (Robert et al., 2001, Bonner et al., 2006, Mehra et al., 2008). Cetuximab has proven to significantly improve the survival of patients with locally advanced HNSCC. Tyrosine kinase inhibitors (TKI) constitute an alternative approach to inhibiting EGFR in the treatment of head and neck cancer. These agents work by blocking the signalling cascade initiated by EGFR/ligand interaction rather than interfering with EGFR/ligand binding (Egloff and Grandis, 2008). Gefinitib and erlotinib are examples of EGFR TKIs that showed an average efficiency (Christy and Bojan, 2013).

1.3 Lung cancer

The term lung cancer (LC) refers to malignancies that originate in the epithelium, lymphatics, mesothelium, or soft tissue of the trachea, bronchi, bronchioles and the lungs (Figure 1.4). Each part contains different predominant cell types, resulting in a variety of LC subtypes. Around 99% of all lung cancers are carcinomas that originate in the epithelium lining of the airways, while the other histologies comprise well under one percent (Travis, 2004, Pass et al., 2005). LC is mostly a result of the progression of a pre-invasive precursor lesion, which can be either squamous dysplasia/carcinoma in situ (SD/CIS), atypical adenomatous hyperplasia (AAH), adenocarcinoma in situ (AIS), or diffuse idiopathic pulmonary neuroendocrine cell hyperplasia (DIPNECH) (Travis, 2011). Although each lesion involves different morphological changes with a series of correlating events on the molecular level, the mechanism of lesion progression towards LC development is not completely understood.

According to the histopathological classification, LC is divided into two main types: non-small cell lung carcinoma (NSCLC), which constitutes 75% to 80% of all LC cases, and the more aggressive small cell lung carcinoma (SCLC) type, which has the worst prognosis and accounts for the remaining 15% to 20%. SCLC generally originates in the central parts of the lungs, and is characterised by neuroendocrine pathophysiologic features (Gould et al., 2004). The SCLC cells have distinct morphology, described as “oat cells” because of their large nuclei and thin cytoplasmic area in addition to their lack of nucleoli (Figure 1.5, A) (Hunt et al., 2009). NSCLC includes squamous cell carcinoma (30%), adenocarcinoma (45%) and large cell lung carcinoma (5% to 10%) (Jemal et al., 2004, Jackman and Johnson, 2005, MacKinnon et al., 2010). The proportions of NSCLC subtypes might differ according to sex (Travis, 2004). NSCLC major types can be further sub-classified into subtypes such as lepidic, acinar, papillary and solid predominant patterns of adenocarcinoma, or the basaloid large cell carcinoma subtype. Lung squamous cell carcinoma, a malignant epithelial tumour with keratinisation arises, from the epithelium layer of the bronchi (Figure 1.5, B). Adenocarcinoma is a malignancy which occurs mostly in the respiratory epithelium of peripheral airways and alveoli. This tumour has a variety of histological subtypes characterised by glandular differentiation and mucin production, in addition to exhibiting different growth patterns (Figure 1.5, C). Large cell lung carcinoma is the undifferentiated non-small cell carcinoma that lacks the cytological and morphological features of small cell carcinoma. Nests of large cells characterise this type of tumour, and they usually have abundant cytoplasm and vesicular nuclei with prominent nucleoli (Figure 1.5, D) (Travis, 2004, Travis, 2011). NSCLC will be the focus of this project due to its predominance and sharing some similarities with HNSCC.

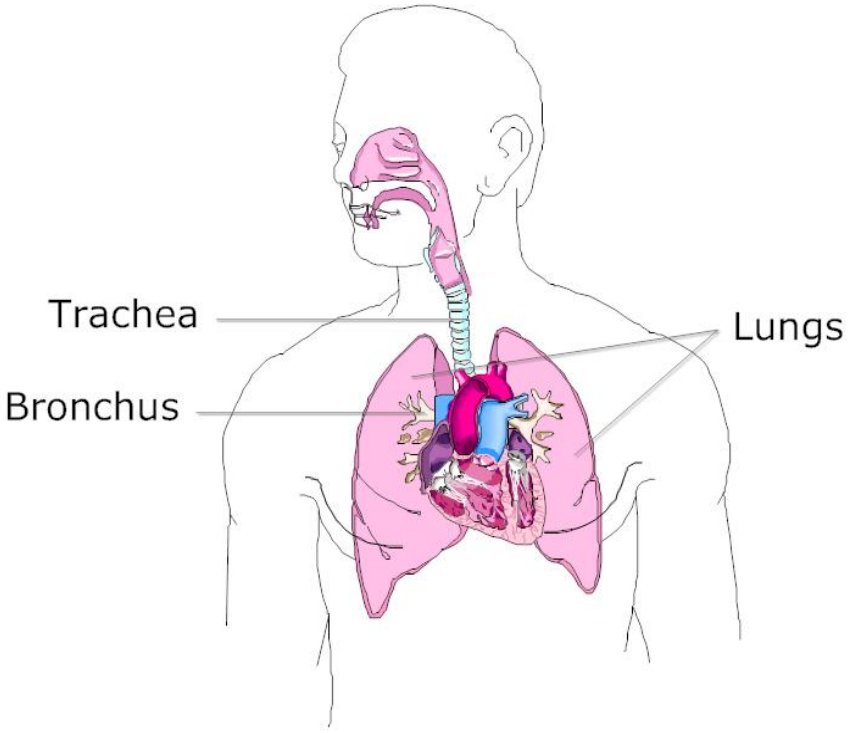


Figure 1.4: Cross section of human respiratory system.

This diagram was generated by SmartDraw software (available on http://www.smartdraw.com/).

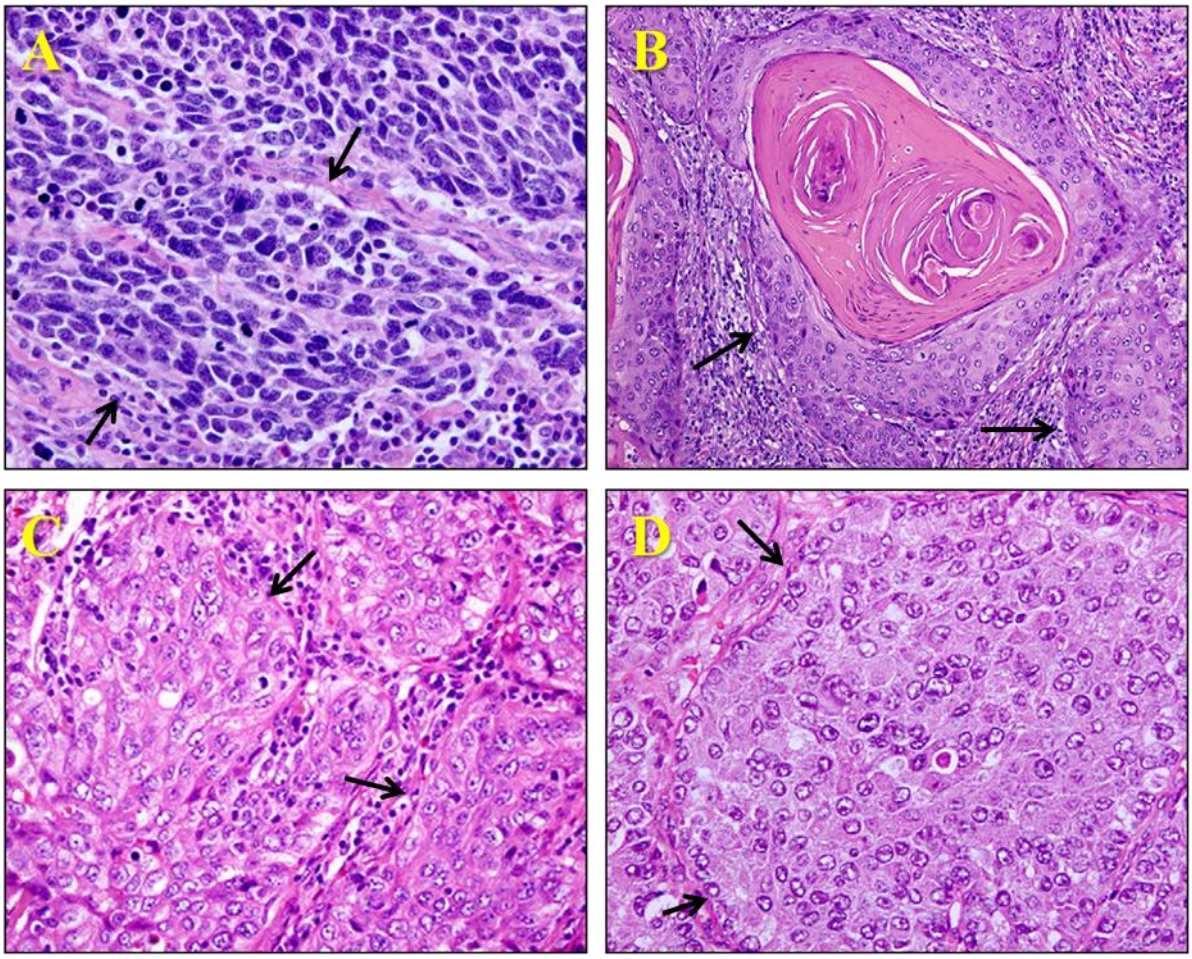


Figure 1.5: SCLC and subtypes of NSCLC.

(A) Small cell carcinoma. Cells of this type are described as “oat cells” because of their large nuclei, thin cytoplasmic area and absence of nucleoli (x100 magnification). (B) Squamous cell carcinoma. These malignant epithelial cells have abundant eosinophilic keratinised cytoplasm (x100 magnification). (C) Solid Adenocarcinoma. These mucin producing cells consist of sheets of tumour cells with abundant cytoplasm and vesicular nuclei with multi conspicuous nucleoli (x100 magnification). (D) Large cell carcinoma. Nests of large cells characterise this type of tumour. These cells have abundant cytoplasm and vesicular nuclei with prominent nucleoli (x100 magnification).

The above haematoxylin and eosin stained tissue sections photomicrographs were taken from (Travis, 2011).

1.3.1 Epidemiology

Lung cancer is one of the most common cancers in the world and is the leading cause of cancer-related mortality among both sexes. In 2008 and 2010, LC killed 1.4 million and 1.5 million people, respectively, with the highest incidence rates detected in North America and Europe (Jemal et al., 2011, Lozano et al., 2012). The lifetime chance of men developing LC in the USA is one in thirteen, while it is one in sixteen for women (Jemal et al., 2010). In the United Kingdom, LC is the second most common type of cancer, affecting around 38,000 new cases and killing approximately 33,500 people annually (Hunt, 2008, Hunt et al., 2009). Moreover, the five year survival in the UK is as low as 7.8% in men and 9.1% in women (Field et al., 2013). However, progress in understanding this disease has allowed the development of some molecular-based targeted therapies that have succeeded in increasing the overall survival by 8 to 12 months according to some clinical studies (Reck et al., 2013). LC mortality and morbidity rates are variable across different parts of the world (Kamangar et al., 2006).

1.3.2 Diagnosis and staging

In addition to systemic signs of cancer, which include weight loss, fever and general fatigue, most common LC manifestations are dyspnea, persistent cough with or without blood, rib pain, hoarseness and upper airway inflow partial obstruction. Also, in the case of very advanced stage of the disease accompanied with metastasis, bone pain and headache are common (Hammerschmidt and Wirtz, 2009, Hunt et al., 2009). In most of the cases, these symptoms appear late while the cancer has already spread beyond its original site and unfortunately there are no specific early symptoms that might help in LC early detection. Another possible reason of the late diagnosis is the inability to visually detect a premalignant lesion in addition to the lack of an effective screening test. Incidental chest X-ray might be the best chance for an early LC detection. Moreover, by reducing LC-associated mortality by twenty percent, periodic low-dose CT screening for high risk individuals might be a useful approach for LC screening if the cost can be managed (Aberle et al., 2011). LC-associated auto-antibodies can be detected three to five years before symptomatic disease (Chapman et al., 2008). Based on this, auto-antibodies for the antigens TP53, NY-ESO-1, SOX2, GBU4-5, CAGE and Annexin 1, which associate with LC, were used to design the EarlyCDT screening test for LC early detection in high-risk populations with a sensitivity of around forty percent (Lam et al., 2011).

When LC is suspected, analysis of sputum cytology and bronchoscopy might be the first investigation attempt. For a definite diagnosis of LC, histological examination and a variety of radiological investigation techniques are applied. Examples of this are chest radiography (CXR), chest CT scan and MRI to detect chest wall invasion and pleural effusion, which present in around one-third of all NSCLC patients (Read et al., 2006, Kernstine and Reckamp, 2010). For

more accurate LC diagnosis and staging, more advanced techniques have been developed including fluorodeoxyglucose positron-emission tomography (FDG-PET), which is also used for prognosis prediction and assessment of the treatment efficiency (Reck et al., 2013).

The universal TNM staging system is applied for identifying the stage of NSCLC. Staging the disease is highly important because it can determine the degree of tumour spread and identify the key information for planning the most effective treatment procedure in addition to predicting the outcome. Similar to many different cancers, TNM staging criteria of NSCLC were developed over many years. The first NSCLC staging system was proposed in 1973, followed by many updates until introduction of its latest version in 2009 (Salgado et al., 2013). The current description of each TNM grade is illustrated in (Table 1.2). Grouping of NSCLC stages is based on the combination of T, N and M characteristics, and in the latest version, the groups are sub-grouped based on the survival rate (Table 1.3).

The management of NSCLC is complicated and requires a multidisciplinary approach. The universal treatment regimen is generally a combination of surgery, radiotherapy and adjuvant chemotherapy, depending on the nature and stage of the cancer (Chang, 2011). Identifying the NSCLC histological subgroup is critical for obtaining the best outcome of the new implemented therapies since their efficacy is subgroup-dependent (Rossi et al., 2010, Warth et al., 2012). For example, cisplatin-pemetrexed (brand name: Alimta®) is a new chemotherapy drug found to be more efficient in treating adenocarcinoma compared with squamous cell carcinoma (Syrigos et al., 2010). Because of the asymptomatic nature of NSCLC in its early stages, only 25% to 30% of the patients are diagnosed with stage I disease, which might be suitable for surgical intervention (Rami-Porta et al., 2009). However, around half of those patients will relapse within five years after receiving surgical resection (Hyun et al., 2013). The larger proportion of NSCLC patients present with advanced stage disease (Morgensztern et al., 2010), and therefore less benefit from the treatment regimen is expected.

Table 1.2: TNM staging system (7th edition) for NSCLC. *Taken from (Tsim et al., 2010).*

Primary tumour (T)

- T1 – Tumour ≤3 cm in diameter, surrounded by lung or visceral pleura, without invasion more proximal than lobar bronchus.
- T1a – Tumour ≤2 cm in diameter.
- T1b – Tumour >2 cm but ≤3 cm in diameter.
- T2 – Tumour >3 cm but ≤7 cm in diameter, or tumour with any of the following features:
  - Involvement of the main bronchus ≥2 cm distal to the carina.
  - Invasion of visceral pleura.
  - Associated atelectasis or obstructive pneumonitis that does not involve the entire lung.
- T2a – Tumour ≤5 cm in diameter.
- T2b – Tumour >5 cm but ≤7 cm in diameter.
- T3 – Tumour >7 cm in diameter, or tumour with any of the following features:
  - Direct invasion of the chest wall, diaphragm, phrenic nerve.
  - Direct invasion of mediastinal pleura or parietal pericardium.
  - Associated atelectasis or obstructive pneumonitis of the entire lung.
  - Tumour within the main bronchus <2 cm to the carina, without involvement of the carina.
  - Satellite tumour nodules in the same lobe.
- T4 – Tumour of any size that has any of the following features:
  - Invasion of the mediastinum.
  - Invasion the heart or great vessels.
  - Invasion of the trachea, oesophagus or recurrent laryngeal nerve.
  - Invasion of a vertebral body or carina.
  - Separate tumour nodules in a different ipsilateral lobe.

Regional lymph nodes (N)

- N0 – No regional lymph node metastasis.
- N1 – Involvement of ipsilateral hilar or peri-bronchial nodes.
- N2 – Involvement of ipsilateral mediastinal or subcarinal nodes.
- N3 – Involvement of contralateral mediastinal or hilar nodes, OR involvement of ipsilateral/contralateral scalene or supraclavicular nodes.

Distant Metastasis (M)

- M0 – No distant metastasis.
- M1 – Distant metastasis present.
- M1a – Separate tumour nodule(s) in a contralateral lobe or tumour with pleural nodules or malignant pleural/ pericardial effusion.
- M1b – Distant metastases.

Table 1.3: Stage grouping and sub-grouping of NSCLC (7th edition).

Taken from (Tsim et al., 2010).

Stage	cTNM Subset	Five-year Survival
0	Carcinoma in situ	
IA	T1a/T1b, N0M0	50–80%
IB	T2aN0M0	47%
IIA	T1a/T1b, N1M0 T2aN1M0 T2bN0M0	36%
IIB	T2bN1M0 T3N0M0	26%
IIIA	T1/T2, N2M0 T3, N1/N2, M0 T4, N0/N1, M0	19%
IIIB	T4N2M0 Any T, N3, M0	7%
IV	Any T, Any N, M1a/M1b	2%

1.3.3 Aetiology and risk factors

1.3.3.1 Smoking

Tobacco smoking is the major population-attributable risk factor that can lead to developing a malignancy in the lungs with a strong dose-response relationship (reviewed by Bach, 2009). According to a 50-year follow-up study, smokers who consume 25 cigarettes per day have relatively a 25 fold increase in their risk to develop LC compared to never-smokers (Doll et al., 2005). Moreover, smoking is the cause of around ninety percent of all lung cancer cases (Khuder, 2001). Despite the efforts made to educate the public about the harm of smoking, around one-third of the world's adult population smoke tobacco (Roemer et al., 2005), and in 25 European countries, 22% of boys and 18% of girls smoke cigarettes (Baska et al., 2009). The inhaled tobacco smoke contains over 4000 chemical compounds, at least sixty of which are known to cause cancer (Roth et al., 2008). Continuous exposure to these agents damages the cells, and over-time, this damage causes a transformation of these cells from normal to malignant with the aid of some environmental and/or genetic factors. The most important role of these factors is facilitating the metabolism of the carcinogenic compounds and enabling them to react with and affect the cellular components (Schwartz et al., 2007). Moreover, *ABCA3*, *CRTAC1*, *CYP3A4*, *GPRC5C*, *LTF*, *PIGN* and *SEMA3C* genes were identified as having a possible role in smoking-induced LC initiation, carcinogenesis and progression, through a number of cell signalling pathways (Wan et al., 2012). Cannabis (marijuana) smoke is known to cause major microscopic injury to the tissues of the lungs and airways (Tashkin, 2013), and it might be another causative agent of lung cancer, especially among adolescents, according to a recent 40-year cohort study (Callaghan et al., 2013). The harm of tobacco and cannabis components is not limited to smokers. Second-hand smokers, or passive smokers, are also affected since urine metabolites of tobacco-specific carcinogens are detectable in non-smokers who have been exposed to cigarette smoke. Therefore, it is not surprising that tobacco smoke was found to cause lung cancer among those who involuntarily inhale it (Samet et al., 2009). In the United Kingdom, passive smoking is responsible for 600 lung cancer deaths every year (Hunt et al., 2009).

LC in never-smokers differ from typical LC at the pathology and molecular levels (Sun et al., 2007, Yano et al., 2008). Several studies suggest that this LC type is actually a result of inflammatory processes occurring during some lung infections, such as tuberculosis or bacterial pneumonia, in addition to the presence of a helping genetic factor (Ballaz and Mulshine, 2003, Engels, 2008, Brenner et al., 2013). Moreover, it was noted that LC in never-smokers is more common in women than men (Wakelee et al., 2007). This might be related to a hormonal effect since some studies have shown an association between the incidents of lung cancer in never-smoker women and the increase in the expression of estrogen receptors alpha and beta (Wu et al., 2005, Mazieres et al., 2013).

1.3.3.2 Radon and Asbestos

Radon is the second most important aetiological factor of LC after smoking, and its harmful effect increases in smokers more than non-smokers (Lantz et al., 2013). Radon is an inert radioactive gas produced when uranium decays, and it produces high energy mass α-particles which can cause DNA damage in the cells of the bronchial epithelium leading to LC development (ATSDR, 2010, Robertson et al., 2013). Uranium traces are naturally found in many rocks and soils. Epidemiological studies have linked exposure to radon gases with lung cancer diagnosed in miners working under ground (Lubin et al., 1995, Heidenreich et al., 2012). In Europe, radon in indoor spaces is estimated to be the cause of about nine percent of all LC deaths (Darby et al., 2005).

Asbestos fibres are naturally occurring silicate minerals characterised by their durability and heat-resistance, which made them a useful component of many construction materials (Mossman et al., 1990, Cugell and Kamp, 2004). Asbestos is one of the classic proven human carcinogens (Nicholson, 1986). An estimated 20,000 asbestos-related lung cancers occur annually worldwide (Tossavainen, 2000). It was estimated that asbestos was behind three percent of all LC-related deaths between 1980 and 2000 in the UK (Darnton et al., 2006). The risk of lung cancer increases upon the exposure to asbestos or to its principal commercial forms (Alberg et al., 2007). Thus, all forms of this mineral have been banned in more than 50 countries, including all European Union member countries (LaDou et al., 2010). Nevertheless, other countries are still importing and exporting asbestos products, mainly chrysotile. To date, the function of asbestos in cancer initiation is not well described.

1.3.3.3 Genetic factors

Host genes play an important role in susceptibility to LC and might even affect the prognosis. Smoking and other LC aetiological factors can cause the disease in only a portion of people exposed to them, which suggest that development of LC might be partially inherent. For example, the risk of developing LC increases with family history of this disease (Matakidou et al., 2005). Moreover, LC incidents among the majority of East Asian women could not be linked to any known LC risk factor, raising the possibility of a genetic factor (Christiani, 2009). It was estimated that people with a positive family history of LC have a twofold higher LC risk compared with others who have no family history (Lissowska et al., 2010).

A number of studies have identified many candidate genes that conferred an increased risk of LC when their normal expression level is affected by an inherited factor. Such factors can be promoter methylation or single nucleotide polymorphisms (SNP). For instance, the locus 6q23-25 was identified by a large-scale linkage analysis as a susceptible region for inherited LC (Bailey-Wilson et al., 2004). This locus contains many genes such as *SASH1*, *LATS1*, *RGS17*,

IGF2R, *PARK2* and *TCF21*, some of which were found to be frequently inactivated by methylation in LC patients (Tessema et al., 2008). Fine mapping of the locus 6q23-25 identified *RGS17* as a candidate gene that associates with familial LC (You et al., 2009). A more recent study has identified a locus at chromosome 12p13.33 to have some inherited variations able to affect LC development risk (Shi et al., 2012). The effect of SNP in LC development can be represented by the T309G polymorphism in the coding sequence of the *MDM2* gene, a negative regulator of the tumour suppressor gene *TP53* (Chen et al., 1996). This SNP increases the risk of LC according to different meta-analyses studies (Wilkening et al., 2007, Gui et al., 2009).

1.3.3.4 Other aetiological factors:

Numerous occupational and environmental carcinogens besides those previously mentioned were demonstrated to have an impact on the risk of developing LC. Such factors might or might not associate with smoking and their role in LC development increases when they occur in a cumulative or synergistic manner. One example is air pollution with many air-borne solid and liquid harmful particles, which might drive the development of LC after their deposition in the lungs (Vineis et al., 2004). Moreover, around 27 principal carcinogens were identified by the International Agency for Research on Cancer (IARC) as potential LC risk factors (Field and Withers, 2012). Examples of these possible cancer-causing agents are coal-tar pitch, arsenic, beryllium, cadmium, nickel and silica (Field and Withers, 2012).

Some disorders such as diabetes mellitus can be an independent risk factor of LC development (Lee et al., 2012). Viral factors were also suggested to play some role in the development of LC. The genotypes 16 and 18 of HPV virus were detectable in more than half of the tested LC cases, especially in non-smoker women (Cheng et al., 2001), and might associate with adenocarcinoma (Badillo-Almaraz et al., 2013). However, some other studies demonstrated no relationship between HPV infection and LC development (Simen-Kapeu et al., 2010, van Boerdonk et al., 2013), but may be attributed to pulmonary metastases (van Boerdonk et al., 2013). Linking alcohol consumption to the risk of LC development has been a controversial subject. However, different studies showed that considering alcohol as an independent factor in LC aetiology can be ruled out (Wakai et al., 2007, Bagnardi et al., 2011).

1.3.4 Genetic alterations and molecular markers

Lung cancer occurs through a complex and multistage process with extensive genomic instability, which ranges from chromosomal imbalance and loss of heterozygosity (LOH), DNA copy number aberration, to genes fusion and mutations. Examples of these include allelic imbalance of a chromosome 8p, an event found to associate with lymph node metastasis

(Woenckhaus et al., 2003). Also, microsatellite instability of chromosome 3p, which was detected in around 78% of pre-invasive bronchial lesions (Lantuéjoul et al., 2009), and contains important genes for NSCLC diagnosis, such as MLH1 and RARβ (Antczak et al., 2013). Other abnormalities detected at the chromosomal level include autosomal aneuploidy of chromosomes 9, 17, 18, 19, 20 and 22 (Demirhan et al., 2010), and homozygous deletion of chromosomes 2q24, 3p14, 5q11, 9p21, 9p23, 11q14 and 21q21 (Sato et al., 2005).

Gene mutations and fusions are commonly detected in LC, and with the help of newly developed sequencing techniques, the number of such abnormalities is growing with an emphasis on distinguishing "driver" mutations from "passenger" ones. Examples of genes commonly detected in LC which have one or more mutations are *TP53*, *GRM8*, *BAI3*, *ERBB4*, *RUNX1T1*, *KEAP1*, *FBXW7* and *KRAS* (Kan et al., 2010). Amplification of several members of proto-oncogenes such as, *EGFR*, *MYC*, *PIK3CA*, *NKX2-1* and *ALK*, is one more abnormality believed to participate in the pathogenesis of LC (Kok et al., 1997, Varella-Garcia, 2010).

Describing the expression signature of some genes at the time of diagnosis was found to help in identifying the disease subtype. For instance, the over-expression of *GSTT1*, *CEL* and *PRDX6* genes characterise squamous cell carcinoma subtype, while adenocarcinomas associate with abnormal expression of *SFTPA2*, *SFTPB*, *MUC1* and *NAPSA* genes (Bhattacharjee et al., 2001, Garber et al., 2001, Nacht et al., 2001, McDoniels-Silvers et al., 2002). MicroRNAs are also reported to be players in LC progression since several of them were detected as affecting the expression of a variety of genes responsible for regulating cancer cells survival, migration, invasion and metastasis. Examples include, miR-133b, miR-126, miR-449a and miR-137, which target EGFR, VEGFA, c-MET and Cdc42 mRNAs, respectively (Zhang et al., 2013a).

The timing and nature of genetic and epigenetic changes in LC are believed to correlate with its pathological progress (Figure 1.6). This is supported by identifying different types of genetic alterations and abnormalities of particular genes at different stages of the disease. Brzeziańska et al. proposed a model for sequential molecular changes occurring during lung carcinogenesis (Brzezianska et al., 2013). In this model, LOH of chromosomes 3p and 9p are proposed to be an early event accompanied by promoter methylation of *p16*, *CDK4*, *CDK6* and other genes. Cell cycle deregulation caused by aberrant telomerase activity and impaired DNA repairing is the main abnormalities in hyperplasia. Furthermore, aneuploidy, LOH of *TP53* and hypermethylation of some genes including *RARβ* and *FHIT* might drive dysplasia formation. Resistance to apoptosis and uncontrolled cell growth characterise the in situ carcinoma stage. The genetic abnormalities at this stage include mutations of some well studied genes, such as *TP53*, *EGFR*, *KRAS* and *Her2*, in addition to disruption of some miRNA expression. Invasive carcinoma and metastasising of cancer cells to other organs is a result of aberrant expression of lamins and integrins in addition to down-regulation of miR-212 and *DLEC1* tumour suppressors.

As mentioned earlier, pathogenesis of LC includes extensive genomic instability, but the most important genes associated with LC are *ALK*, *EGFR*, *KRAS*, *MET*, *MEK*-1, *TP53*, *ROS1*, *RET*, *Her2*, *p16* (*INK4A*), *PTEN* and *BRAF* (Johnson et al., 2012, Cooper et al., 2013) (some of these will be discussed briefly in the next sections). The large number of abnormally expressed genes has encouraged researchers to investigate the benefit of neutralising their effect through targeting therapy strategy. In addition to EGFR, the most promising targets include Her2, BRAF, MET and EML4-ALK fusion (Farhat and Houhou, 2013).

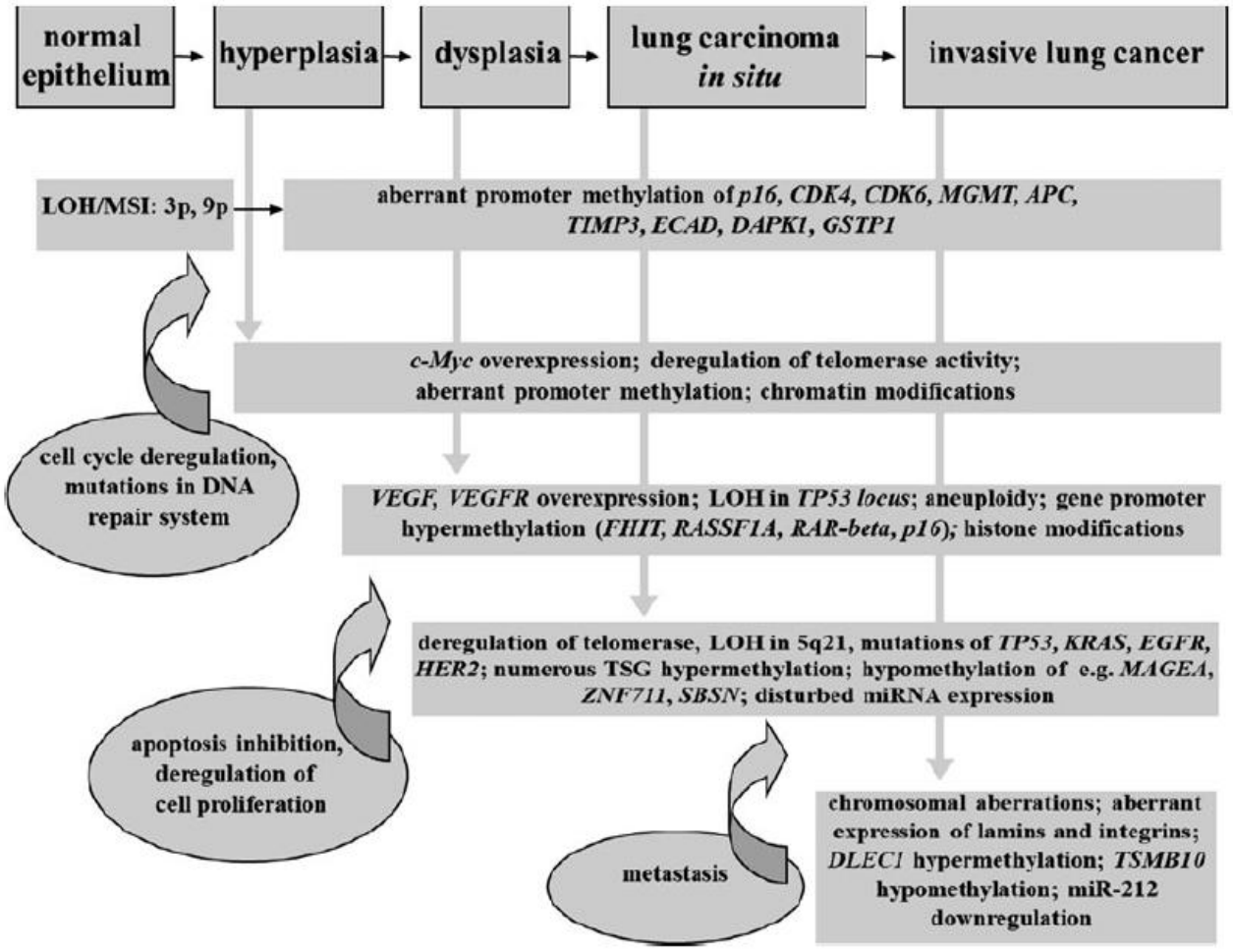


Figure 1.6: A model of molecular changes that occur during lung carcinogenesis.

The sketch was taken from (Brzezianska et al., 2013).

1.3.4.1 Epidermal growth factor receptor (EGFR)

The aberrant expression of EGFR protein is a prevalent molecular event in NSCLC and other cancers. Over-expression of EGFR protein is believed to occur in 50% to 80% of NSCLCs, while in 65% of cases, its DNA copy number increases (Hirsch et al., 2003, Raben et al., 2004, Cappuzzo et al., 2005). It is widely accepted that aberrant expression of EGFR is an early oncogenic event in NSCLC and plays an important role in its pathogenesis as excessive signalling of EGFR impairs the normal balance between cell proliferation and apoptosis (reviewed by Hynes and Lane, 2005), which can be a tumour-contributing factor. Moreover, EGFR over-expression has been associated with worse prognosis and poor overall survival (Rokita et al., 2013). Mutations in the *EGFR* gene, which were discovered for the first time in 2004 (Lynch et al., 2004), are more common in women, Asians, and in 40% to 60% of never-smokers NSCLCs (Rosell et al., 2009, D'Angelo et al., 2011). While EGFR over-expression is more frequent in squamous cell carcinoma, mutated *EGFR* is more common in adenocarcinoma cases (Skrzypski et al., 2013). Many *EGFR* mutations lead to ligand-independent persistent activation of tyrosine kinase (reviewed by Red Brewer et al., 2013). Over 250 different sites for EGFR mutations have been described so far and the vast majority of *EGFR* mutations occur in exons 18 to 21 of tyrosine kinase domain. However, small in-frame deletions in exon 19 and missense mutation a in exon 21 are by far the most commonly detected mutations accounting for ninety percent of all *EGFR* mutations (De Luca and Normanno, 2010).

Repression of EGFR signalling activates intrinsic the mitochondrial apoptotic pathway which causes cell death (Costa et al., 2007). Therefore, EGFR has been a target for developing many molecular-based therapies, such as the anti-EGFR monoclonal antibody nimotuzumab and cetuximab, and some TKIs, including gefitinib, afatinib and erlotinib (Farhat and Houhou, 2013). Moreover, it was found that the presence of some *EGFR* mutations, rather than over-expression, confer a sensitivity to TKI agents and can predict a favourable clinical outcome (Kumar et al., 2008a). Although 70% to 80% of NSCLCs with EGFR mutations respond to first-line TKIs (Mitsudomi et al., 2005, Mitsudomi and Yatabe, 2007), a major proportion develop drug resistance through various mechanisms acquiring different types of *EGFR* mutations insensitive to these agents. However, there are different mechanisms through which tumour cells become insensitive to EGFR TKIs. Examples of these include activation of alternative tyrosine kinase pathways (Zhang et al., 2012), transformation of NSCLC tumour cells into mesenchymal cells (Nurwidya et al., 2012) and acquisition of small cell features (Sequist et al., 2011). *EGFR* mutation status is currently the first step in clinical trials to evaluate the efficacy of EGFR second- and third-line TKIs, such as Dacomitinib (Brzezniak et al., 2013). Routine methods used for this purpose include the diagnostic *EGFR* Mutation Kit (Zhou et al., 2011), high-resolution melting analysis (Fukui et al., 2008) and direct sequencing of the PCR-amplified exon sequences (Angulo et al., 2012).

1.3.4.2 Kirsten rat sarcoma viral oncogene homolog (KRAS)

The small GTP-binding protein encoded by *KRAS* participates in signalling networks important for different cellular processes including differentiation, apoptosis and cell growth (reviewed by Downward, 2003). *KRAS* mutations in lung cancer were detected for the first time in 1984 (Santos et al., 1984). Similar to the mutations of *EGFR*, mutations of *KRAS* are considered driver mutations, which are significantly involved in NSCLC progression (Luo and Lam, 2013). *KRAS* mutations predominantly occur in adenocarcinoma with an incidence of 15% in never-smoker and 25% in smoker Caucasians (Riely et al., 2008). These percentages vary based on geographical location. *KRAS* mutations correlate with poor survival (Mascaux et al., 2005), and resistance to EGFR TKIs such as gefitinib and erlotinib (Pao et al., 2005, Massarelli et al., 2007).

In NSCLC, most mutations of *KRAS* coding sequence occur at codons 12, and less often at codons 13 and 61 of exon 2. These mutations result in constitutive activation of KRAS downstream pathways (reviewed by Harris and McCormick, 2010). Some of these pathways are shared with EGFR signal transduction, and in the presence of an EGFR downstream active KRAS signalling, these pathways become independent of EGFR signalling (Shigematsu et al., 2005, Harris and McCormick, 2010). Hence, persistently active KRAS signalling can bypass the effect of EGFR inhibition driven by its targeting TKIs, which prompted the need to develop agents able to block mutated-KRAS activity as well. The development of clinically effective anti-KRAS agents has been a difficult goal. An alternative approach is tackling the oncogenic effect of its contingent downstream signalling. For example, atorvastatin was designed to restrain KRAS-dependent AKT and ERK activity through interruption of KRAS/PI3K and KRAS/RAF complexes leading to induction of cancer cell apoptosis (Chen et al., 2013b). Selumetinib, which inhibits KRAS-downstream MAPK kinases, MEK 1 and 2, has also been developed as an anti-KRAS agent (Paolo et al., 2013).

1.3.4.3 Tumour protein p53 (TP53)

TP53 inactivation is one of the most common genetic abnormalities in many cancers including NSCLC. Eighty percent of *TP53* mutations are inactivating missense mutations within the DNA-binding domain (Olivier et al., 2002). These mutations occur early during LC development and are maintained during tumour progression (Mogi and Kuwano, 2011). According to a recent meta-analysis study, mutations of *TP53* occur in around thirty percent of lung adenocarcinomas and in fifty percent of lung squamous cell carcinomas in western countries (Dearden et al., 2013). Alterations in *TP53* coding sequence were found to associate with smoking, more aggressive tumours and poor survival (Mitsudomi et al., 2000, Toyooka et al., 2003). Moreover,

recurrence of NSCLC was found to associate significantly with the presence of *TP53* mutations (Ludovini et al., 2008).

In contrast with the effect of *KRAS* and *EGFR* mutations, which allow them to gain an abnormal function, the majority of *TP53* mutations result in its loss of normal function and also acquisition of some abnormal functions. Mutated TP53 modulates the expression of some genes to enhance tumour progression and metastasis (Scian et al., 2004, Vaughan et al., 2012). Targeted inactivation of protein with abnormal function, such as mutated EGFR, tends to be more conducive to drug development than re-activating protein, such as mutated TP53, or even substituting it with a wild-type one e.g. by gene replacement therapy, in which *TP53* wild-type coding sequence is encapsulated in adenovirus designed to infect cancer cells. Since several studies demonstrated a central role of *TP53* mutations in LC chemoresistance (Kandioler et al., 2008, Huang et al., 2011), combining chemotherapy with *TP53* gene replacement therapy was indicated as a rational strategy. The results of such am approach were promising in pre-clinical trials and more so in the associated clinical studies where the effect of chemotherapy and even radiotherapy was improved in the presence of the introduced wild-type *TP53* (Guan et al., 2009, Liu et al., 2009). Moreover, a low-molecular weight PRIMA-1 compound can be of therapeutic value since it selectively inhibits the proliferation of mutated-*TP53* harbouring NSCLC cells (Bykov et al., 2002, Magrini et al., 2008). One further approach to using TP53 normal anti-tumour function is through its small molecule activator, RITA. This drug activates the cellular wild-type TP53 in cancer cells leading to cell apoptosis (Zhao et al., 2010).

1.4 The need for new HNSCC and NSCLC biomarkers

The early detection of any cancer is a key factor for increasing the treatment efficiency and improving the patients' survival. As previously discussed, the majority of HNSCC and NSCLC patients are diagnosed at advanced stages. One of the reasons is the lack of reliable biomarkers. Current biomarkers of HNSCC and NSCLC are insufficient for full understanding of the underlying biology and pathophysiology of these diseases. Moreover, targeting some of these biomarkers, such as EGFR, by molecular-based therapies to tackle cancer cells might show some promising results at the beginning, but developing resistance by cancer cells is frequent. In addition, the effect of these therapies is not universal since many HNSCC and NSCLC cases do not show any response. Advanced techniques, such as NGS and microarrays, continuously identify new genetic alterations in HNSCC and NSCLC. In-depth investigation might help in understanding their roles in HNSCC and NSCLC development and progression, identifying potential new biomarkers that are useful for cancer screening and monitoring, and eventually translating these findings to new therapeutics if possible.

1.5 *HOX* genes

1.5.1 Discovery:

In 1894, William Bateson, who was the first scientist to coin the term 'genetics' for the study of inheritance (Keynes and Cox, 2008), discussed the un-expected structural variations and body part replacement in animals or insects descending from the same species, and named this observation "homeosis" (Bateson, 1894). Not many years later, the first "homeotic" mutation was identified in Drosophila (Bridges and Morgan, 1923). Discoveries of genetic mutations linked to homeosis succeeded one after the other until two homeotic gene complexes (*HOM-C*) controlling Drosophila segmentation were described separately – *Antennapedia complex* (*ANT-C*) and *Bithorax complex* (*BX-C*) (Lewis, 1978, Kaufman et al., 1980). *HOM-C* genes were detected to control the segments development along the anterior-posterior (A-P) body axis of Drosophila, while mutations affecting these genes resulted in the formation of body parts in wrong locations (Lewis, 1978). An example of this is the homeotic transformation resulting from some mutations in *ANT-C* genes that caused the formation of legs rather than antennae in the head of Drosophila (Figure 1.7) (Schneuwly et al., 1987). Interestingly, co-linearity between genes locations within the complex and their corresponding segment was demonstrated (Lewis, 1978, Hoey et al., 1986).

Homeotic proteins encoded by the *HOM-C* genes were identified to be transcription factors that switch on or off the process of transcription of a large number of target genes (Levine and Hoey, 1988, Liang and Biggin, 1998). In addition, the function of most of these homeotic genes was found to be mediated by a highly conserved 180 base pair DNA sequence, named homeobox, which encodes a sixty amino acid long DNA-binding protein motif, termed the homeodomain (McGinnis et al., 1984b, Scott and Weiner, 1984, Gehring, 1987). However, quite a number of homeotic genes were found to contain an atypical homeobox motif consisting of 183 bp (Duboule, 1995a, Banerjee-Basu and Baxevanis, 2001).

Through homolgy searches, many homeobox-containing genes, named homeobox genes, were detected in different invertebrates and vertebrates, including human (Levine et al., 1984, McGinnis et al., 1984a, Akam, 1989). More and more studies in this field have demonstrated that not all homeobox genes are homeotic genes, therefore, the term "homeotic" is not interchangeable with the term "homeobox". This is because not all homeotic genes contain the homeobox motif in their coding sequence and not all homeobox genes have the homeotic transformation feature if get mutated (Meyers, 2004, Slack, 2006). The main functions of non-homeotic homeobox genes involves determining cell fate and its differentiation pathway, and regulating the required cell behaviour during morphogenesis, such as migration and invasion (Hombría and Lovegrove, 2003).

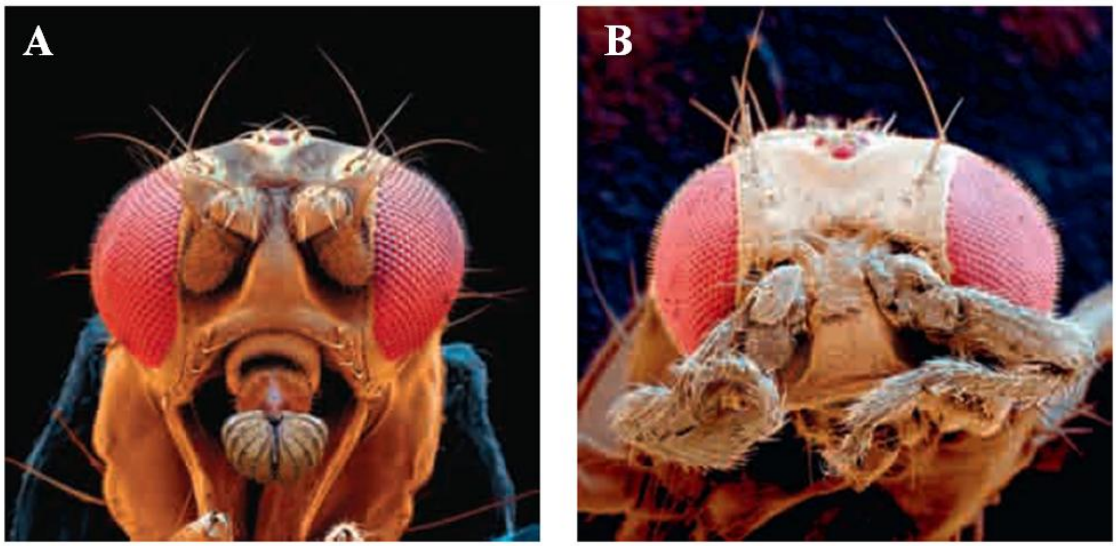


Figure 1.7: Homeotic transformation in the head of *Drosophila Melanogaster*.

(A) A normal fly with antennae and wild-type ANT-C genes. (B) A fly with a mutation in one of its ANT-C homeotic genes that normally gives rise to antennae. The abnormal expression of this particular homeotic gene caused the formation of legs rather than antennae in the head of Drosophila fly. *The photographs were taken from (Russell et al., 2010)*.

1.5.2 HOX of homeobox

A long list of homeobox genes has been identified in metazoans and grouped into different classes and families. The human genome is estimated to possess around 235 functional homeobox genes (Holland et al., 2007). It is important to mention that there is an innocent, but confusion-driving, mistake commonly seen in the literature and different scientific resources, which is using the word "Hox" as an abbreviation for the term "homeobox". Homeobox was shortened to Hox during the 1980s when scientists were referring to the newly discovered *HOM-C* homologues in vertebrates. The following studies of homeobox genes during the last two decades made such abbreviation no longer valid since it was revealed that the firstly discovered clustered Hox genes are just one family of homeobox genes superclass, which contains 11 classes subdivided into 102 families of clustered (Class I) and non-clustered (Class II) dispersed genes (Holland et al., 2007). Moreover, *Hox* (in lower case letters) is designated to describe *HOM-C* homologues genes in animals, while *HOX* (in capitals) is reserved to describe those of human, a rule applied to all other human genes (Wain et al., 2002).

A nomenclature system applied in 1992 to identify the genes of *HOX* and *Hox* clusters. The clusters were named A, B, C and D, and based on sequence similarity; their genes were sorted into 13 paralogous groups with an equivalent position in each cluster. The genes were numbered in each cluster from 1 at the 3′ end to 13 at the 5′ end. When a paralogous group is absent from a cluster, the corresponding gene number is omitted (Scott, 1992, Scott, 1993). In humans, there are 39 *HOX* genes clustered on chromosomes 7p15.3, 17q21.3, 12q13.3 and 2q31 (Figure 1.8) (Apiou et al., 1996).

1.5.3 Specificity of *HOX* genes

Most homeobox genes, including *HOX* genes, have a highly conserved homeobox motif in their coding sequences which modulate their functions. A sensible question that arises is how the different members of HOX proteins can fulfil their job in controlling the expression of their target genes in a very specific manner when their function-mediator, the homeodomain, is highly similar among all of them. Multiple molecular events were found to enhance homeodomain binding and specificity. Firstly, the homeodomain was found to exist as a self-folding helix-turn-helix motif that exerts its effect on the regulatory regions of its target genes through an interaction between its third α-helix, also called "recognition helix", and the bases of its target at the DNA major groove. The N-terminal motif of this folded structure forms a flexible arm and makes contact with the minor groove to contribute in a homeodomain-to-DNA specific binding with high affinity (Figure 1.9) (Affolter et al., 1991, Qian et al., 1993, Gehring et al., 1994). Moreover, the recognition helixes of one to eight paralog HOX proteins can identify the canonical TAAT core sites in their target DNA sequence, while those of the remaining posterior

groups preferentially bind to TTAT core sequence. The specificity of homeodomains binding to either TAAT or TTAT sites depends on their preference for the bases occurring immediately 3' or 5' to these cores (Deutsch, 2010). Furthermore, HOX proteins contain functional motifs that contribute to homeodomain specificity through mediating the interaction of HOX proteins with other transcription factors, such as PBX, MEIS and EXD (Mann and Chan, 1996, Shanmugam et al., 1999). For example, the hexapeptide and YPWM flanking motifs of HOX proteins are binding sites for PBX proteins, which allow the formation of a HOX·PBX complex that has more specificity and binding affinity than a solo HOX protein (Shanmugam et al., 1997, LaRonde-LeBlanc and Wolberger, 2003, Moens and Selleri, 2006). With the aid of these co-factors, some HOX proteins have distinct functions while many others appear to exhibit overlapping or redundant functions (McIntyre et al., 2007).

1.5.4 Regulation of *HOX* genes

The expression of *HOX* genes and most homeobox genes in general are tightly organised along the A-P axis through different phenomena noticed earlier during *HOM-C* genes studies (Lewis, 1978, Hoey et al., 1986, Dolle et al., 1991), and later called Spatial and Temporal co-linearities. Spatial co-linearity is a unique controlling system in which *HOX* genes residing towards the 3′ end of each cluster expressed more at the anterior tissues of the embryo, while genes closer to the 5′ end expressed more at the posterior ones. Temporal co-linearity is the term used to describe the correlation between the *HOX* genes order and the timing of their expression during embryogenesis; in other words, an early expression of 3′ genes during the embryo life occurring versus later expression of 5′ genes (Duboule, 1994). Moreover, another kind of co-linearity, named quantitative co-linearity, was identified and described as a co-expression of several *HOX* genes at a given position along the A-P body axis with expression of the *HOX* genes positioned more towards the 5′ end stronger than the 3′ genes (Dolle et al., 1991, Shah and Sukumar, 2010). In addition, genes of *HOX* clusters were noticed to undergo a "posterior prevalence", which is a functional dominance of posterior *HOX* proteins over the anterior ones (Duboule and Morata, 1994).

It may be questioned what triggers the early expression of *HOX* genes in a freshly fertilised egg to allow them to perform their roles in development. The answer might come from the well-studied Drosophila embryogenesis model. Coordinated changes and a cascade of events have been identified to be responsible for regulating oocyte maturation, pre- and post-implantation stages, reaching the point of A-P pattering, segmentation and complete embryo formation. It was found that maternal mRNAs of some genes called "maternal-effect genes" are deposited into the unfertilised egg from the surrounding cells during oogenesis. Immediately after fertilisation, these dormant mRNAs are translated into active transcription factor proteins

that induce the sequential expression of different classes of "segmentation genes", which in turn establish segmental periodicity and regulate the expression of HOM-C genes, which define the identity of individual segments (Papageorgiou, 2007, Russell et al., 2010). A similar cascade of events in higher chordates is still not fully understood and of course vertebrates might have much more complicated mechanism of *HOM-C* homologues genes activation. However, a number of studies demonstrated the presence of maternal factors in early mammalian development (Farley and Ryder, 2008, Li et al., 2010a, Lindeman and Pelegri, 2010), in addition to some clues of very early *Hox* genes expression during implantation derived by signalling molecules, including members of the WNT and fibroblast growth factor (FGF) growth factor families, produced from the surrounding tissues. Such events were found to occur during and even before gastrulation (Adjaye and Monk, 2000, Forlani et al., 2003, Mallo et al., 2010, Paul et al., 2011). Collectively, this evidence indicates a scenario of *HOX*/*Hox* gene activation similar, to some extent, to what has been demonstrated in Drosophila embryogenesis.

Other factors associated with the regulation of *HOX* genes expression are microRNAs, chromosome modifications, epigenetic changes and hormones, such as estrogen and progesterone (Pearson et al., 2005, Daftary and Taylor, 2006, Duester, 2008, Vasanthi and Mishra, 2008, Brock et al., 2009, Soshnikova and Duboule, 2009). Furthermore, a recent study has proposed the possibility that *Hox* genes have an auto-regulatory loop (Sheth et al., 2013). Vitamins are other regulators of *HOX* genes expression during development. Some studies showed that retinoic acid (vitaminutes A) is a vital regulator of *HOX*/*Hox* gene expression during embryogenesis in a concentration-dependent manner (Boncinelli et al., 1991, Hill et al., 1995, Tabin, 1995, Marshall et al., 1996, Koop et al., 2010). Vitamin D is also important for normal differentiation of hematopoietic and maturation of endometrial stromal cells during development and has some roles in *HOX* genes regulation (Du et al., 2005).

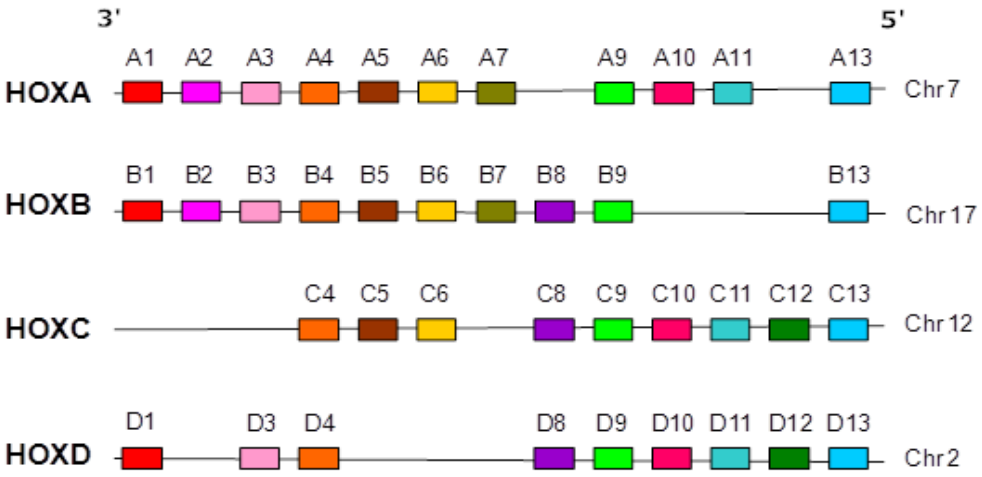


Figure 1.8: Clusters of human HOX genes.

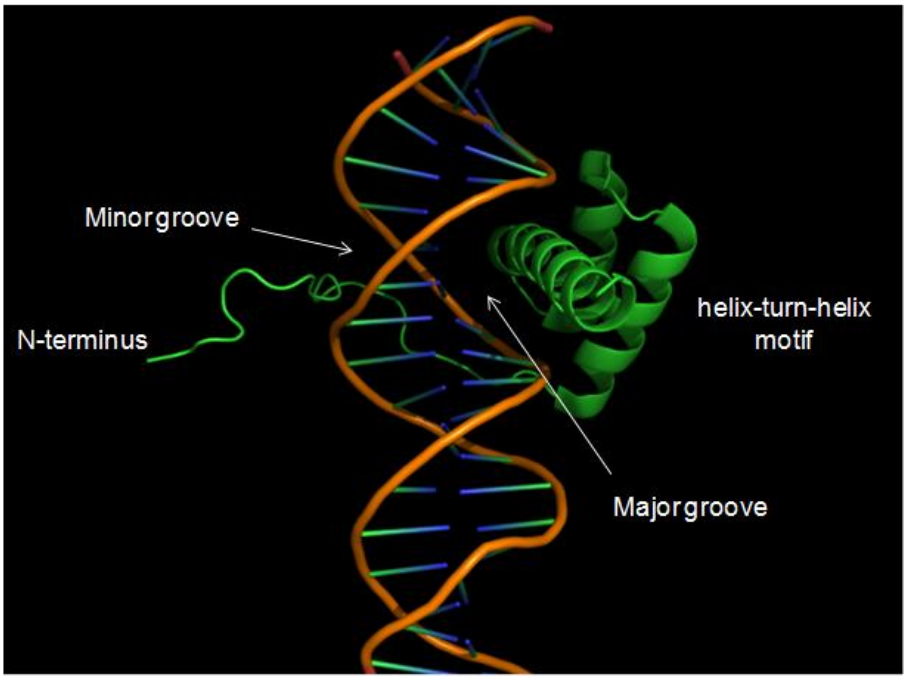


Figure 1.9: 3D structure of a homeodomain interaction with DNA.

The homeodomain exists as a helix-turn-helix motif and exerts its effect on the regulatory regions of its target genes through an interaction between its third α-helix and the major groove of the target DNA. A pentapeptide N-terminal motif also makes contact with the minor groove to contribute to high affinity binding.

This 3D view was created through PyMol software (available on www.pymol.org) using a PDB file of the crystal structure of HOXA9 complexed with PBX1 and DNA (available on http://www.rcsb.org/pdb/home/home.do), which was reported by (LaRonde-LeBlanc and Wolberger, 2003).

1.5.5 *HOX* genes in human development

Studying the pattern of *HOX* genes expression during different stages of development, even through *Hox* genes of mice models, is quite complex because their expression changes rapidly with time and stage. Moreover, multiple alternative transcripts add another dimension to their complicated roles due to possible different regulation of each transcript (Kim et al., 2000), in addition to having potentially distinct, related, or even opposite function (Wang et al., 2008a). However, it is widely accepted that during embryogenesis, HOX proteins control the expression of development-associated genes via binding to their regulatory elements in order to regulate cellular proliferation, migration and differentiation towards normal organogenesis process and body axis identification (Dorn et al., 1994, Krumlauf, 1994, Maconochie et al., 1996, Lemons and McGinnis, 2006). Controlling cell proliferation and differentiation in different regions of the embryo might be the most important roles of HOX proteins (Duboule, 1995b, Chen and Capecchi, 1999). Evidence of HOX proteins roles were derived from loss-of-function and gain-of-function studies on HOX homologues in mice embryos. Hox proteins were found to be essential for normal spinal cord formation (Carpenter et al., 1997, Tiret et al., 1998), while others were linked to neuronal differentiation (Gaufo et al., 2003). More studies of Hox-development relationship revealed their key roles in the development of the gut (Beck, 2002, Zacchetti et al., 2007), ear and hindbrain (Chisaka et al., 1992, Lumsden and Krumlauf, 1996), limb and vertebrae (Horan et al., 1995, Fromental-Ramain et al., 1996, Zakany and Duboule, 2007), ossification and joint formation (Villavicencio-Lorini et al., 2010), skin (Stelnicki et al., 1998), rib cage (Vinagre et al., 2010), lungs (Golpon et al., 2001), central nervous system (Akin and Nazarali, 2005), pancreas (Gray et al., 2011), reproductive organs (Du and Taylor, 2004), cardiovasculature (Gorski and Walsh, 2003) and even hair (Awgulewitsch, 2003). Moreover, HOX proteins were proven to be vital for normal hematopoietic development and differentiation, and have been detected to be actively expressed in primary human stem/progenitor cells (Sauvageau et al., 1994, Magli et al., 1997). Interestingly, some transcription factors, including some members of homeobox genes, were found to act as "messenger proteins" through crossing nuclear and even cellular membranes raising the possibility of having the ability to mediate cell-to-cell communication and organising of a surrounding cells' behaviour (reviewed by Prochiantz and Joliot, 2003).

Regulatory roles of HOX proteins are not restricted to embryonic pattering since many of them have been found to be not only detectable in adult tissues, but also organ-specific (maybe to maintain their identity), and expressed in a collinear pattern comparable with that of the developing embryo (Cillo et al., 2001, Neville et al., 2002, Yamamoto et al., 2003, Takahashi et al., 2004). In particular, adult tissues with ongoing cell differentiation have been found to retain high expression of some *HOX* genes. A clear example of this is the continuous expression of HOXA9, HOXA10, HOXA11 and HOXA13 in male and female reproductive systems to regulate spermatogenesis and endometrial remodelling (Lindsey and Wilkinson, 1996, Taylor, 2000,

Morgan, 2006, Vitiello et al., 2007). Another example is HOX gene expression in hematopoietic cells, in which different combinations of HOX proteins determine lineage identity and regulate cell proliferation (Eklund, 2007). Furthermore, HOX proteins are important players in some physiological processes such as wound healing (Mack et al., 2003) and lactation (Morgan, 2006).

1.5.6 Congenital disorders of *HOX* genes

Within the last ten to fifteen years, a number of *HOX* gene mutations have been identified to associate with human congenital disorders. Most of these disorders include features of abnormal limbs, which support the role of HOX proteins in planning the body structure. One of these disorders is the autosomal recessive Bosley-Salih-Alorainy syndrome, which is a result of a rare homozygous mutation in the *HOXA1* gene, and was first described in consanguineous individuals from Saudi Arabia and Turkey (Tischfield et al., 2005). A different type of *HOXA1* mutation was found to cause one more congenital disorder called athabascan brainstem dysgenesis syndrome (Holve et al., 2003, Bosley et al., 2008). A mutation that affects the 44th amino acid of HOXA2 homeodomain was identified to be a direct cause of an autosomal recessive microtia (Alasti et al., 2008). Deletion or point mutation of *HOXA13* are the reasons for the autosomal dominant hand-foot-genital syndrome (Mortlock and Innis, 1997), while a missense mutation in its homeobox is a cause of another condition called Guttmacher syndrome (Innis et al., 2002). Malfunctioning HOXC13 resulting from loss-of-function mutations was found to cause pure hair and nail ectodermal dysplasia (Lin et al., 2012). To date, mutated *HOXD13* might be the leading cause of human congenital disorders among *HOX* genes since it is linked to numerous conditions including Synpolydactyly type II (Akarsu et al., 1996), Brachydactyly types D and E (Johnson et al., 2003), Synpolydactyly type V and Brachydactyly-Syndactyly (Zhao et al., 2007). Mutations in *HOXD10* gene have also been found to cause some congenital disorders (discussed in a later section).

A malfunctioning HOX protein is not always necessary to cause a congenital disorder since an abnormal *HOX* gene expression is sufficient. For instance, abnormal over-expression of HOXB5 during embryonic life is associated with both the rare congenital malformation of the lower respiratory tract, bronchopulmonary sequestration, and the congenital cystic adenomatoid malformation (Volpe et al., 2000, Golpon et al., 2001, Volpe et al., 2003). Although most of the remaining *HOX* genes have not been yet linked to known congenital disorders, their reported roles in organogenesis might predict which organs can be affected most by loss-of-function.

1.5.7 *HOX* genes and cancer

Numerous types of cancer were detected as having an interesting alteration in a number of *HOX* genes. More interesting is the fact that some members of the HOX proteins family which were found to play an oncogenic role in a cancer are tumour suppressors in some other cancer. A possible reason for this is the presence or absence of some HOX proteins cofactors, as well as the variations in active pathways among different types of tissues. Moreover, such paradoxical observations support the idea that HOX proteins and their targets are cell and tissue type-specific. The definite reason and the impact of abnormal expression disturbance of some *HOX* genes on the molecular pathology of different cancers are still elusive in most of the cases. However, there is a general idea that the down-regulation or up-regulation of some *HOX* genes leads to improper activation of embryonic developmental pathways in mature organs, and subsequently, interrupting normal differentiation or cell growth cascades which might contribute in abnormal cell behaviour and neoplasia (Lawrence and Largman, 1992, Abate-Shen, 2002, Shah and Sukumar, 2010). It was noticed that the majority of the up-regulated *HOX* genes in a cancer of a particular part of the body are normally expressed during the development of that part, while the ones that are expressed normally in mature organs are down-regulated in cancer (Abate-Shen, 2002). Moreover, a relationship between congenital anomalies-causing genes and the risk of cancer development might be true (Anbazhagan and Raman, 1997). This implication of development-associated genes, such as *HOX* genes, in cancer has established the "oncology recapitulates ontology" hypothesis (Lechner et al., 2002, Lappin et al., 2006), in which genes that determine the differentiation and lineage of the progenitor cells do have some important roles in cancer initiating cells as well. The importance of HOX proteins in cancer pathology is tightly linked to their capacity to regulate important mechanisms, such as angiogenesis, cell proliferation, invasion, metastasis, survival and apoptosis. In general, there are abnormal changes in *HOX* genes expression in different tumours and cancers, including leukaemia, breast cancer, head and neck cancer, lung cancer, ovarian cancer, prostate cancer, cervical cancer, thyroid cancer, renal cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, melanoma, lymphoma and neuroblastoma. Examples of these cancers and their associated dysregulated *HOX* genes are explored briefly below.

Leukaemia. A and B *HOX* gene clusters were proven to have key roles in the maturation of hematopoietic stem cells, while their expression was found to decline in the differentiated progeny (Sauvageau et al., 1994, Lawrence et al., 1997, Pineault et al., 2002, Lawrence et al., 2005). However, *in-vivo* experiments showed that aberrant expression of particular members of these clusters, such as *HoxA9* or *HoxA10*, drive acute myeloid leukaemia (AML) (Thorsteinsdottir et al., 1997, Kroon et al., 1998). More members of the *HoxA* cluster are also implicated in AML and acute lymphoid leukaemia (ALL) development (Bach et al., 2010, Alharbi et al., 2013). Moreover, over-expression of *HOXA9* was linked to the worst prognosis outcome in AML patients (Golub et al., 1999). Such an abnormal expression of *HOXA9* and other *HOX*

genes in leukaemia might be driven by their regulators, CDX and mixed lineage leukaemia (MLL) fusion proteins (Hess, 2004, Bansal et al., 2006, Rawat et al., 2008, Smith et al., 2011). Moreover, cofactors of HOX proteins, such as MEIS1 and PBX3, were found to be essential for HOX oncogenic roles in leukaemia (Shen et al., 1999, Thorsteinsdottir et al., 2001, Li et al., 2013b). Although some HOX proteins appear to be oncogenic during leukaemia development, others might have a tumour suppressor function. For example, *HOXA4* and *HOXA5* were reported to be frequently inactivated by hypermethylation in a high percentage of the whole studied AML and CML samples, while their re-expression through DNA methylation inhibitor treatment induced cells differentiation (Strathdee et al., 2007). *HOXB1* was also suggested to be a TSG since enforcing its expression in AML cells impaired their growth and induced their apoptosis (Petrini et al., 2013).

Breast Cancer. Out of many *HOX* genes that showed indications of variable roles in this cancer *HOXA1*, *A7*, *A5*, *A9*, *B7* and *D10* might be the most significant. Two different studies have demonstrated a high level of *HOXA1* in breast cancer cells compared to normal cells (Chariot et al., 1995, Chariot and Castronovo, 1996). A further study showed that the up-regulation of *HOXA1* may be influenced by human growth hormone (hGH), and the increase of this gene contributes in oncogenic transformation by promoting cell growth and controlling the expression of the *Bcl-2* gene to abrogate apoptosis (Zhang et al., 2003). HOXA7 might have a similar role in promoting cell proliferation but through up-regulating the expression of estrogen receptor alpha (*ERα*) (Zhang et al., 2013b), which can initiate a growth-promoting cascade triggered by an interaction with estrogen (Chalbos et al., 1982, Fox et al., 2009), or even without estrogen activation signal (FOWLER et al., 2004). HOXB7 is also a progression-promoting HOX protein in breast cancer. It has a high expression level in breast cancer cells and enhances their growth factors-independent proliferation (Care et al., 1998). Moreover, HOXB7 plays a role in tumour cells resistance to chemotherapy (Jin et al., 2012), and mediates breast cancer cell invasion and epithelial-mesenchymal transition mechanism (Wu et al., 2006).

A potential role of HOXA5 in breast cancer was indicated by identification of its function as a positive regulating factor of the tumour suppressor gene, *TP53*. The low level of *TP53* mRNA in cancer cells was detected to result from a low expression level of HOXA5 (Raman et al., 2000), which was estimated to occur in around two thirds of breast carcinomas (Henderson et al., 2006), and caused mostly by promoter methylation (Raman et al., 2000). The mismatch repair pathway component, *MLH1* gene, was found to be a direct target of HOXA5 (Duriseti et al., 2006), which is further indication of a HOXA5 tumour suppression function in breast cancer cells, a function that has been reported previously in leukaemia.

In contrast to its oncogenic role in leukaemia, *HOXA9* is a TSG in breast cancer. This breast tissue development and lactation-associated *HOX* gene (Chen and Capecchi, 1999), was found to be epigenetically silenced via promoter methylation in up to ninety percent of all

the tested samples (Unger, 2002, Reynolds et al., 2006). The reason for this high inactivation frequency might be the ability of HOXA9 to repress the progression and metastasis of breast cancer cells through direct targeting and up-regulating the level of the TSG, *BRCA1* (Gilbert et al., 2010). Such findings were further supported and confirmed by a following study (Jin and Sukumar, 2010). A more recent study has identified regulatory elements of *HOXA9* in breast cancer that maintain its inactive form (Sun et al., 2013). The demethylating TET1 protein was proven to be an inducer of the expression of different members of *HOXA* cluster genes, including *HOXA9*, through demethylating their promoters. However, the function of TET1 is restrained by the chromatin remodelling factor, HMGA2 (reviewed by Fusco and Fedele, 2007), which has a high expression level in breast cancer cells and plays a role in their invasion and metastasis (Rogalla et al., 1997). In this relation, HMGA2 indirectly suppresses the expression of *HOXA9* in breast cancer. *HOXD10* is also a TSG in breast cancer, but will be discussed in more detail later.

Prostate cancer. *HoxB13*, the last gene identified in vertebrate *Hox* clusters (Zeltser et al., 1996), is essential for normal prostate development and maturation in a mouse model (Economides and Capecchi, 2003). Human HOXB13 protein is not only highly specific to prostate but also of a high-expression level in normal human prostate tissue (Jung et al., 2004), which supports the need to retain some developmental gene expression in tissues with continuous cell differentiation. The expression of HOXB13 in prostate tissues decreases in the case of cancer (Jung et al., 2004). Such irregular alteration in expression applies to other *HOX* genes as well, including *HOXD10* (Miller et al., 2003). The low HOXB13 in prostate cancer cells might be due to its ability to repress the signalling pathway initiated by androgen receptor activation, which impairs cell growth triggered through this cascade (Jung et al., 2004). Interestingly, tumour suppression functions of HOXB13 turns out to be oncogenic with an increase in expression if cells are androgen-independent (Kim et al., 2010b), which is an example of how the roles of HOX proteins respond to the alterations in the active pathways in a particular cell type. Other *HOX* genes found to associate with prostate cancer progression are some of the *HOXC* cluster genes (Miller et al., 2003), particularly the proliferation and progression-associated gene, *HOXC6* (Ramachandran et al., 2005, McCabe et al., 2008) and HOXC8, which might mediate an invasive and metastasis phenotype of prostate cancer (Waltregny et al., 2002).

Lung cancer. Re-expression of airway development associated-HOXA5 is seen in mature lung malignancies (an example that harmonises with the "oncology recapitulates ontology" theory). Tissues of NSCLC and SCLC express higher level of HOXA5 when compared to the non-cancerous tissues (Abe et al., 2006, Xiao et al., 2013); a finding that does not match what has been demonstrated in breast cancer and leukaemia cases. Moreover, such high expression of *HOXA5* transcripts in SCLC cells is reported to associate with multidrug resistance (Xiao et al., 2013). Other *HOX* genes found to be abundantly expressed in lung

primary tumours include *HOXA10* and *HOXB9* (Calvo et al., 2000), *HOXD10* (Abe et al., 2006), in addition to *HOXC5* and *HOXD13* (Lechner et al., 2002). Different studies presented controversial findings as regards the expression of some *HOX* genes in lung cancer, such as *HOXA9* and *HOXA7*, since they were reported to be up-regulated (Calvo et al., 2000), but also silenced through promoter methylation in another study (Rauch et al., 2007). This discrepancy might be due to the differences in the analysed sample type (e.g. cell line or tissue) and/or tumour stage.

Head and neck cancer. Some of the well studied misexpressed *HOX* genes in oral cancer include *HOXA1*, *HOXB7*, *HOXC6*, as well as the potential prognosis markers, *HOXA5* and *HOXA10*. Screening oral squamous cell carcinoma samples revealed a higher expression level of both *HOXA1* and *HOXB7* in cancers compared to normal oral tissues (Hassan et al., 2006, De Souza Setubal Destro et al., 2010, Bitu et al., 2012). Moreover, products of both genes were found to promote cancer cell growth *in-vitro*, concurrent with proliferation markers detected in cancer tissue samples, and to associate with poor prognosis (De Souza Setubal Destro et al., 2010, Bitu et al., 2012), which collectively supports their tumourgenic role in this type of cancer. *HOXC6* is another abnormally over-expressed *HOX* gene in head and neck cancer tissues and cell lines (Hassan et al., 2006, Moon et al., 2012). In addition, HOXC6 might play some role in HNC metastasis as its level is frequently higher in cancer tissues with lymph node metastasis than those with no metastasis (Hassan et al., 2006). Besides its function in inducing cell proliferation, HOXC6 is an upstream positive regulator of the anti-apoptosis *Bcl-2* gene, which has been reported to participate in chemotherapy resistance (Thomas et al., 2013). Although *HOXA5* is up-regulated in HNC compared to normal cells, its low level in this cancer might predict a shortened survival rate (Rodini et al., 2012). In contrast, the abundant expression of *HOXA10* in HNC cells might positively correlate with the disease-free survival rate (Yamatoji et al., 2010). In addition to the above mentioned *HOX* genes, *HOXD10* has previously been identified by our laboratory to have differential expression between normal oral keratinocytes and HNSCC cells (Hunter et al., 2006).

1.5.8 *HOX* genes as potential therapeutic targets in cancer

The expression of most *HOX* genes, and other developmental genes, is restricted in normal adult tissues. Therefore, when some *HOX* genes are abnormally expressed in cancer cells in some tissues, they might become relatively specific cancer targets, with low potential for adverse effects in other tissues/organs. The first report of blocking the functions of HOX proteins as an approach to tackle cancer was through direct interfering of their binding to their cofactors, which are essential for their function as previously mentioned. A small cell-permeable 18-residue peptide, named HXR9, was designed to antagonise the interaction between HOX cofactors, PBX proteins and their hexapeptide binding motif within HOX proteins (Morgan et al., 2007). HOX·PBX dimers have significantly greater binding affinity and specificity for target DNA sequences than the HOX monomer alone (Chang et al., 1995). Disrupting the formation of HOX·PBX dimer triggered apoptosis in melanoma cells both *in-vitro* and *in-vivo* (Morgan et al., 2007). Studying the effect of HXR9 was then tested in other cancers, including NSCLC (Plowright et al., 2009), ovarian cancer (Morgan et al., 2010), kidney cancer (Shears et al., 2008), in addition to myeloma (Daniels et al., 2010). The results are promising; HXR9 was able to significantly induce cancer cell apoptosis and block tumour growth (Shears et al., 2008, Plowright et al., 2009, Daniels et al., 2010, Morgan et al., 2010). However, such disruption of HOX·PBX interaction is not specific to a particular member of HOX proteins, which raise more challenges to enhance its specificity.

Some of current approaches to tackle cancer cells involve introducing siRNA or shRNA, which are capable of inhibiting specific tumour promoting genes or re-expressing the lost tumour suppressor genes. In some previous studies, HOXD10 was indeed capable of suppressing the growth and progression of breast cancer cells (Carrio et al., 2005). Based on this, researchers in the University of California are currently conducting a study using the latter approach to re-express the breast cancer tumour suppressor gene, *HOXD10*, in tumour cells. They transfected some mammary tumour macrophages with HOXD10-encoding vector, and assumed that those cells would mediate delivering *HOXD10* into breast tumour tissues, leading to restoration of HOXD10 level to play its tumour suppression role (Boudreau and Coussens, unpublished work). Such a study, and other similar published or on-going studies, might open new possibilities for taking advantage of the roles of HOX proteins in cancer and developing new therapies.

1.6 HOXD10

The human *HOXD10* gene, formerly known as *HOX4D* or *HOX4E*, is located in chromosome 2q31.1. This two-exon gene is transcribed into three different mRNAs transcripts. However, the 1,023 base pair (bp) variant is the only one that encodes a functional protein of 340 amino acids size (Holland et al., 2007, Flicek et al., 2013). Similar to the other HOX proteins, HOXD10 functions as a sequence-specific transcription factor, and is expressed early in the developing limb buds and is believed to be involved in the development and pattering of the limb and spinal cord (Johnson and Tabin, 1997, Lance-Jones et al., 2001, Shah et al., 2004, Tarchini et al., 2006).

1.6.1 Pathologies associated with abnormal *HOXD10* expression

Defects in *HOXD10* have been linked to three well-described congenital disorders, Charcot-Marie-Tooth (CMT), congenital vertical talus (CVT) and Wilms tumour. The motor and sensory neuropathy, or CMT, was reported more than a century ago and since then, many forms of it have been described and linked to mutations in more than thirty different genes (reviewed by Noriega and Ramos, 2013). It is considered the most common inherited neurologic disorder with a prevalence of 17 to 40 per 10,000 new births (Vallat et al., 2013). CMT is characterised by weakness of distal limbs resulting from progressive loss of muscle tissues with or without loss of touch sensation (Figure 1.10, A) (Pareyson, 1999). CVT, the common name of the "rocker-bottom foot" condition or congenital convex pes valgus, is an uncommon physical distortion that affects and dislocates the joints of the navicular and talus bones (Figure 1.10, B) (Shrimpton et al., 2004). It was reported for the first time in 1913 (Rocher, 1913), and has an approximate incidence of one in 10,000 (reviewed by Mckie and Radomisli, 2010). Through a whole-genome linkage analysis study using DNA samples from a CMT and CVT-affected family, followed by sequencing the candidate detected genomic site, an abnormality was detected in the *HOXD10* gene. The defect was identified to be a missense mutation resulting from a single-nucleotide substitution (thymine to adenine, (956T>A), and mapped to the recognition helix of HOXD10 homeobox (Shrimpton et al., 2004), which has a negative impact on the functionality of HOXD10 and affects its role during development leading to the manifestations of both CVT and CMT. Another study showed that this *HOXD10* mutation can drive the CVT disorder without CMT phenotype (Dobbs et al., 2006).

Wilms tumour, also referred to as nephroblastoma, is the most common type of childhood renal tumours, accounting for more than ninety percent of all cases (Pastore et al., 2006). Although it was reported for the first time in 1814, it was named after the surgeon who fully described it, Dr. Max Wilms (Rance, 1814, Wilms, 1902). Clinical features of this tumour includes an asymptomatic abdominal mass and haematuria (Figure 1.10, C) (Petruzzi and

Green, 1997). An association between Wilms tumour and an aberrant *HOXD10* transcript has been reported (Redline et al., 1994). Wilms tumours that have different genetic abnormalities, including WT1 deletion, are also associated with very high expression levels of *HOXD10* (Royer-Pokora et al., 2010). The function of HOXD10 in this cancer might be linked to its role in kidney development (Fu et al., 2012).

The expression level of *HOXD10* has been reported to be deregulated in different neoplasia, including gastric cancer, oesophageal squamous cell carcinoma, ovarian cancer, cervical cancer and endometrial adenocarcinoma, besides the previously mentioned, breast cancer, prostate cancer, lung cancer and oral cancer. In some cancers, including endometrial adenocarcinoma, prostate cancer and breast cancer, its expression varies with the histological grade of differentiation. Moreover, its function has been found to differ from one cancer to the other. For instance, in contrast with its suggested tumour suppression function in breast cancer (Carrio et al., 2005), ovarian cancer (Nakayama et al., 2013), endometrial adenocarcinoma (Osborne et al., 1998) and prostate cancer (Miller et al., 2003), it is reported to have an oncogenic role in cervical cancer (Lopez et al., 2006), and oesophageal squamous cell carcinoma (Takahashi et al., 2007), as well as head and neck and lung cancers. HOXD10 over-expression in the latter two cancers has been reported in many studies (Abe et al., 2006, Hassan et al., 2006, Hunter et al., 2006, Plowright et al., 2009, Rodini et al., 2012). However, its function and interacting molecules in those two types of cancers have not yet been described. Although HOXD10 was suggested to have a tumour suppressing role in gastric cancer associated with its loss of expression by promoter hypermethylation (Wang et al., 2012), a previous study reported a variable expression in patient samples (Rossi Degl'Innocenti et al., 2007).

Figure 1.10: Congenital disorders associated with a mutated HOXD10 gene.

(A) CMT. Characterised by weakness of distal limbs resulting from progressive loss of muscle tissues with or without loss of touch sensation. (B) CVT. An uncommon physical distortion that affects and dislocates the joints of the navicular and talus bones. (C) Wilms tumour. Clinical features of this tumour include an asymptomatic abdominal mass and haematuria. *The photos A, B and C were taken from (Pareyson and Marchesi, 2009), (Gore and Spencer, 2004) and (Petruzzi and Green, 1997), respectively.*

1.6.2 Cancer-related HOXD10 interactions

According to cell type or tissue type, HOXD10 appears to be implicated in different cancer-related networks, which in turn may mediate its differential roles between cancers. *HOXD10* expression is significantly reduced in breast cancer development, since it is highly expressed in the premalignant cells compared to low level in primary tumour cells, then extremely low levels detected in invasive breast carcinomas (Carrio et al., 2005). Similar findings were recently demonstrated in ovarian cancer (Nakayama et al., 2013). Losing HOXD10 in those two cancers is important for their progression, but maybe not their proliferation, since HOXD10 was found to down-regulate α3 integrin and inhibit breast cancer-related abnormal vasculogenesis (Myers et al., 2002), in addition to reduced metastasis in both cancers. These findings are supported by detection of a decrease in the expression of a number of genes involved in cell invasion, cell migration and extracellular matrix remodelling, such as *uPA*, *RhoC* and *MMP14*, when high expression of HOXD10 was maintained in breast cancer cells (Carrio et al., 2005). Further research revealed that HOXD10 is actually part of more complex networks in breast cancer, in which hyaluronan, which is an abundant glycosaminoglycan in the extracellular matrix in most tissues, binds to its receptor, CD44, leading to transcriptional activation of the Twist transcription factor mediated by c-Src kinase. Twist, in turn, up-regulates miR-10b expression, which is located within the *HOXD* genes cluster and exerts post-transcriptional inhibition on HOXD10 mRNA. Subsequently, the resulting *HOXD10* down-regulation induces the expression of RhoA/RhoC and activates the Rho-kinase pathway leading to promotion of cellular invasion and migration (Myers et al., 2002, Carrio et al., 2005, Ma et al., 2007, Bourguignon et al., 2010). However, miR-10b is not capable of down-regulating HOXD10 in some other cancers, such as HNSCC (Severino et al., 2013). Another important microRNA found to interact with HOXD10 is miR-7. Reddy and colleagues discovered that introducing high levels of HOXD10 in breast cancer positively regulates the expression of miR-7, which in turn inhibits p21-activated kinase 1 (Pak1) (Reddy et al., 2008). Pak1 plays a fundamental role in controlling cell motility and invasion (Ong et al., 2011). However, miR-7 might not always function as a tumour suppressor, rather an oncogenic, or "oncomiR", since it was found to play a key role in promoting EGFR-mediated lung oncogenesis (Chou et al., 2010). In addition to its identified direct target, *IGFBP3*, HOXD10 was found to up-regulate the expression of *PDGFRL*, *HCLS1*, *CXCL9* and *ANP32A*, while it has a suppressing role over *RAC2*, *NTS*, *KRT5* and *TUSC3* genes (Wang et al., 2012). Many of these genes were reported previously to be involved in cell proliferation, adhesion and apoptosis.

1.7 Project hypothesis and aims

Current data indicate that *HOXD10* is highly expressed in most HNSCC and NSCLC cell lines, while cells derived from metastases have low HOXD10 expression. These findings are in agreement with other studies reporting that *HOXD10* expression is inversely correlated with invasiveness. SNP analysis of the HNSCC lines by Hunter et al. showed no consistent alteration at the *HOXD10* locus, indicating that the alterations are more likely the result of transcriptional and/or post-transcriptional control. The distinct functions and suggested roles of HOXD10 in other cancers, as previously discussed, in addition to the current data and the need to identify new HNSCC and NSCLC biomarkers underpinned investigation of its role in both cancers simultaneously. Both cancer types demonstrate predominance of squamous cell carcinoma share aetiological factors, such as tobacco smoking. Moreover, a recent study has demonstrated that HNSCC and NSCLC share some gene expression signatures and their molecular subtypes are similar (Walter et al., 2013).

The hypothesis of this project is:

HOXD10 participates in HNSCC and NSCLC pathogenesis with over-expression providing proliferative advantage while low expression supporting cancer cell invasiveness and metastasis.

The objectives of this project are:

1- Evaluation of the expression level of *HOXD10*, and previously reported interacting molecules, in HNSCC and NSCLC cell lines.

2- Assessment of HOXD10 protein level in a panel of HNSCC and NSCLC cell lines and tumour tissue microarray.

3- Assessment of the phenotypic effects of *HOXD10* over-expression or knocking down on HNSCC and NSCLC cells.

4- Detection of novel downstream targets of HOXD10 in NSCLC and HNSCC cells.

CHAPTER 2: MATERIALS AND METHODS

2.1 Materials

2.1.1 Equipment

Main laboratory equipments used in this project are listed in (Appendix 1).

2.1.2 Reagents and chemicals

The analytical grade chemicals, reagents and consumables used during routine lab work were supplied by Sigma-Aldrich (Cambridge, UK) or Thermo Fisher Scientific Inc. (Loughborough, UK), unless otherwise stated.

2.1.3 Buffers and solutions

The in-house prepared buffers, solutions and dyes used for different experiments are listed with their ingredients in Appendix 2.

2.1.4 Antibodies

A goat polyclonal (sc-33005) antibody raised against human HOXD10 was provided by Santa Cruz through Insight Biotechnology (London, UK) to be used at a concentration of 1 µg/mL for the detection of HOXD10 protein in cell lysates. HOXD10 protein in tissue sections was detected using 20 µg/mL of rabbit polyclonal (ab90704) antibodies purchased from Abcam (Cambridge, UK). For the detection of AMOT-p80 protein in both cell lysates and tissue sections, an anti-AMOT-p80 rabbit polyclonal (ab83381) provided by Abcam was used at a concentration of 2 µg/mL for WB and 20 µg/mL for IHC. Beta actin in cell lysates was detected using a Sigma-Aldrich mouse monoclonal anti-β-actin antibody at a concentration of 0.5 µg/mL. For western blotting, secondary antibodies conjugated with horseradish peroxidase to detect anti-mouse, anti-rabbit and anti-goat antibodies were purchased from Amersham (Bucks, UK) and were used in a dilution of 1 in 3000. Biotinylated goat anti-rabbit IgG secondary antibody for IHC work was purchased from VECTOR labs (Peterborough, UK).

2.1.5 Peptides

Abcam HOXD10 synthetic peptide (ab90993), which was derived from within residues 100-200 of human HOXD10, was used to neutralise Abcam's anti-HOXD10 antibody (ab90704) during IHC quality control steps.

2.1.6 Antibiotics

For bacterial culture work, ampicillin purchased from Fisher was used at a concentration of 100 µg/mL. ACROS G418 sulfate (geneticin) was provided by Fisher, and was prepared in a sterile environment for the purpose of selecting neomycin-expressing mammalian cells.

2.1.7 Media for eukaryotic cell culture

In a Class II laminar flow cabinet, media for human cell culture were prepared by combining their supplements listed in Appendix 3.

2.1.8 Cells

All normal cells and cancer cell lines used in this study with their relevant information are listed in Table 2.1.

Table 2.1: Cell lines used in this study.

Histology	Name*	Source tissue	Culture medium	Reference
Normal keratinocyte	Oral keratinocyte 17 (**OK17**)	buccal mucosa	Keratinocyte Growth Medium (KGM)	(*isolated from healthy volunteers*)
	Oral keratinocyte 19 (**OK19**)			
	Oral keratinocyte 21 (**OK21**)			
	Oral keratinocyte 23 (**OK23**)			
Immortalized normal keratinocyte	FNB6 hTERT (**FNB6**)			(McGregor et al., 2002)
	OKF4	floor of the mouth		(Hu et al., 1991)
	OKF6/TERT-1 (**OKF6**)			(Dickson et al., 2000)
Dysplastic	**D4**	tongue		(McGregor et al., 2002)
	D6			
	D19			
	D20			
	D35			
Metastatic	**TR146**	neck lymph nodes		(Rupniak et al., 1985)
	BICR 22 (**B22**)**			(Edington et al., 1995)
Squamous cell carcinoma (SCC)	BICR 16 (**B16**)**	tongue		
	BICR 56 (**B56**)**			
	H357			(Prime et al., 1990)
	T4	floor of the mouth		(Hunter et al., 2006)
	T5	buccal mucosa		
Non-small cell lung cancer (NSCLC)	**A549**	lung	Roswell Park Memorial Institute (RPMI) medium	(Giard et al., 1973)
	NCI-H647 (**H647**)			(Carney et al., 1985)

* The names given in bold are the abbreviations used in the text.

** BICR stands for Beatson Institute for Cancer Research.

2.1.9 Primers and probes

The below pre-designed TaqMan® chemistry-based probes/primers listed in (Table 2.2) were purchased from Life Technologies (Glasgow, UK).

Table 2.2: TaqMan® pre-developed probes/primers.

Gene Name	Symbol	NCBI Accession #	Assay ID
Homeobox D10	*HOXD10*	NM_002148.3	Hs00157974_m1
β-2 microglobulin	*B2M*	NM_004048.2	Hs00984230_m1
microRNA-7	*miR-7*	NR_029605.1	000386
microRNA-10b	*miR-10b*	NR_029609.1	00218
microRNA-146a	*miR-146a*	NR_029701.1	000468
microRNA-155	*miR-155*	NR_030784.1	002623
small nucleolar RNA, C/D box 48	*RNU48*	NR_002745	001006

Primers listed in (Appendix 4) were designed in-house then synthesised by Sigma-Aldrich to be used for SYBR® Green I dye-based real-time quantitative polymerase chain reaction (qPCR).

2.1.10 Expression vectors and siRNA/shRNA

Pre-designed short hairpin RNA (shRNA) and short interfering (siRNA) to silence HOXD10 expression were purchased from BiOSETTIA (San Diego, USA) and OriGene (Maryland, USA), respectively. Cloned pre-miR-146a in pcDNA3 was purchased from Addene (Cambridge, USA). pcDNA3.1D/V5-His-TOPO vector purchased from Invitrogen (Glasgow, UK) was used to clone the coding sequence of any gene of interest. The pGL3-basic luciferase reporter vector and pRL-TK *Renilla* internal control vector (Promega, Southampton, UK) were used to study promoter activity of some genes.

2.2 Methods

2.2.1 Cell culture

2.2.1.1 Routine cell culture procedure

HNSCC and NSCLC cell lines in addition to normal oral keratinocytes were cultured in tissue culture flasks (CELLSTAR, Greiner, Germany) with their specific media in a 37 °C CO_2 incubator with five percent CO_2 concentration and 100% humidity. The culture media were refreshed two to three times per week. The cell growth was monitored daily by microscope until cells reached around eighty percent confluency when they were ready to be either passaged or harvested according to the following procedure. As a part of the standard tissue culture safety procedures, cell media were routinely tested for any Mycoplasma infection by the staff lab technicians. In case of cell types that demonstrated a slow growth rate, their media were refreshed by aspirating just half of the old media and compensated with an equal amount of fresh media to preserve a reasonable amount of the signalling molecules, growth factors and autocrine factors normally produced by the cultured cells in their media, which would enhance the overall cells' growth.

2.2.1.2 Cell passaging and harvesting

The cell culture medium was aspirated with a sterile pipette, 3 mL of sterile Phosphate Buffer Saline (PBS) were added to wash the cells and completely remove the remnants of the culture medium. To detach the monolayer cells from the flask's surface, 3 mL of a one percent Trypsin/Ethylenediaminetetraacetic acid (TE) solution were added to the washed cells followed by incubation in the 37 °C CO_2 incubator. Proteolytic activity of the trypsin enzyme breaks the cell-to-cell and cell-to-surface adhesion while the Ethylenediaminetetraacetic acid (EDTA) chelates Ca^{2+}, which is needed for calcium-dependent cell adhesion molecules (Hong et al., 2009), and eventually this combination creates a single-cell suspension. Detachment of cells from the flask's surface was monitored under the microscope on a minute-to-minute basis and enhanced by fine tapping on the sides of the flasks to avoid prolonged incubation of cells with trypsin, which might cause cell damage or lysis. To stop trypsin activity, an equal volume of a serum-containing growth medium was added, mixed, then the cell suspension was transferred into a universal tube of volume 15 mL to be centrifuged for five minutes at 200 g. After removing the supernatant, the cell pellet was either resuspended or harvested according to the following procedure. For passaging cells into another flask, 5 mL of fresh culture medium were added to the cell pellet, then mixed thoroughly prior to counting and re-plating between new culture flasks. In the case of harvesting cells for protein or RNA extraction, the cell pellet was washed by 3 mL of cold PBS, centrifuged as before, then the supernatant was aspirated completely, while the cell pellet was either used for the next procedure directly or stored at -20 °C.

2.2.1.3 Cell counting

A haemocytometer cell counting chamber (Neubauer, Cambridge, UK) was used for routine cell counting. Briefly, 10 µL of a 0.4% Trypan blue stain were added to 10 µL cell suspension, mixed, then 10 µL of the mix were dispensed at the edge of the cover-slip and allowed to run under the cover slip. Using a microscope with 200x magnification objective lens, the viable cells were counted in the four squares of the chamber with excluding the blue-stained dead cells. The average number was multiplied by 10,000 to give the actual number of cells per mL.

2.2.1.4 Cryogenic preservation and recovery

Cell suspensions containing 10^6 cells/mL were combined with a tissue-culture-grade dimethyl sulfoxide (DMSO) in a ratio of 10:1 followed by aliquoting into cryotubes, then placed in a freezing container (Mr. Frosty) and stored at -80 °C for 24 to 48 hours. Cryotubes were then stored in a liquid nitrogen container for long-term storage.

To recover frozen cells, a cryotube was reomved from the liquid nitrogen and promptly placed into a 37 °C water bath for a quick thawing. The vial's content was then transferred into a universal tube containing 3 mL of culture medium. After a three minutes centrifugation at 200 g, the cell pellet was resuspended in 5 mL medium before transferring into a culturing flask to be incubated in normal culturing conditions.

2.2.2 Quantitative real-time polymerase chain reaction (qPCR)

2.2.2.1 Total RNA extraction

The ISOLATE RNA Mini kit (Bioline, London, UK) was used to extract total RNA from cell pellets according to the manufacturer's instructions. To describe the procedure briefly, a cell pellet was resuspended in lysing buffer containing guanidine-thiocyanate, which inactivates RNases present in cell lysate to ensure the purification of an intact RNA. Seventy percent Ethanol (EtOH) was then added to enhance binding of RNA to the silica-based membrane of the spin column during centrifugation. The RNA was subsequently washed twice followed by elution with 50 µL RNase/nuclease-free H_2O. The concentration and purity of the isolated RNA were measured using a NanoDrop spectrophotometer. The machine's software uses the spectrophotometric absorbances at 230, 260 and 280 nm wavelengths to calculate the 260 to 280nm absorbance ratio ($A_{260/280}$) and 260 to 230 nm absorbance ratio ($A_{260/230}$) according to Beer-Lambert law (Trumbo et al., 2013). $A_{260/280}$ massively less than two indicates a possible contamination with protein, while $A_{260/230}$ less than 1.8 indicates a contamination with ploy phenols and/or polysaccharides (Farrell, 1998). Therefore, only RNA samples of $A_{260/280}$ and

$A_{260/230}$ more than or equal to 1.9 were used for complementary DNA (cDNA) preparation and qPCR analysis.

For more concentrated microRNA content in the total RNA extract, the Ambion *mir*Vana miRNA Isolation kit (Invitrogen) was used as per the manufacturer's guidelines. In this protocol, cell pellets were firstly disrupted in a denaturing lysis buffer then subjected to Acid-Phenol:Chloroform extraction for purification and removal of DNA. After that, Ethanol was added to samples before performing a filtering step using a cartridge containing a glass-fiber to immobilise the RNA. The filter was then washed two times before eluting the RNA using RNase/nuclease-free H_2O followed by assessment of its concentration and quality as described earlier.

2.2.2.2 Complementary DNA (cDNA) preparation

High Capacity cDNA Reverse Transcription kit (Life Technologies) was used to generate cDNA from the extracted single-stranded mRNA. In this procedure, 1 µg RNA was reverse transcribed according to the manufacturer's recipe:

Component	Amount
100 mM dNTPs	0.8 µL
10X RT buffer	2 µL
Multiscribe enzyme	1 µL
Random primers	2 µL
Rnase inhibitor	1 µL
Sample RNA (100ng/µL)	10 µL
Nuclease-free water	3.2 µL
Total volume	20 µL

Then, the reaction mix was loaded into a thermal cycler programmed with the following parameters:

Step	Temp (°C)	Time (min)
1	25	10
2	37	120
3	85	5
4	4	∞

The final product was either used directly for qPCR or stored at -20 °C.

In the case of miRNA reverse transcription, specific stem-loop primers (Life Technologies) for the target miRNA and the endogenous control RNU48 were used separately to generate the cDNA. For each sample, 15 ng RNA were reverse transcribed according to the manufacturer's instructions:

Component	Amount
100 mM dNTPs	0.15 µL
10X RT buffer	1.5 µL
Multiscribe enzyme	1 µL
Stem-loop primer	3 µL
Rnase inhibitor	0.19 µL
Sample RNA (3 ng/µL)	5 µL
Nuclease-free water	4.16 µL
Total volume	15 µL

Then, the reverse transcription reaction was carried out in a thermal cycler programmed as follows:

	Step 1	Step 2	Step 3	Step 4
Temperature	16 °C	42 °C	85 °C	4 °C
Time	10 min	120 min	5 min	∞

The final product was either used directly for qPCR or stored at -20 °C.

2.2.2.3 Designing qPCR primers

As previously mentioned, TaqMan® probes/primers were pre-developed assay reagents, while SYBR® Green I dye-based primers were designed in-house. To design a pair of primers compatible with SYBR® Green reaction method, the mature mRNA sequence of a target gene was obtained from Ensembl Genome Browser (http://www.ensembl.org/index.html) (Flicek et al., 2013). Primer sequences were chosen to span exon-exon junctions to avoid co-amplification of genomic DNA.

Primers were designed to the following properties: GC content in the 40% to 60% range, a primer melting temperature (Tm) between 58 °C to 65 °C, 20 bp to 32 bp length and an amplicon product size of 50 bp to 150 bp. Before ordering, primers' specificity and properties were assessed using the Primer-Blast algorithm (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) (Ye et al., 2012).

2.2.2.4 Running qPCR reactions

The ABI 7900HT Fast Real-Time PCR system was used to provide a quantitative detection of target genes using either a pre-designed TaqMan® probes/primers set or SYBR® Green I dye-based primers. Although the chemistry of both TaqMan® and SYBR® Green methods allow the detection of the PCR product as it accumulates during PCR cycles, they have different level of specificity. The TaqMan® method provides more specific detection since it consists of a set of primers for amplification and specific fluorogenic probe for detection. On the other hand, the detection of the DNA-binding SYBR® Green dye depends on the specificity of the designed primers only since it detects any amplified product. Moreover, singleplex PCR was performed for all SYBR® Green I dye-based experiments in which target genes and endogenous controls were amplified in separate wells, while in the case of TaqMan® probe-based experiments, multiplex PCR was possible as it allowed co-amplification and co-detection of both the target gene and the endogenous control in the same well as long as their probes were labelled with different dyes, such as FAM or VIC, which are detected at different wavelengths. In general, the quality control of each qPCR experiment was monitored by including a positive and negative control samples with known housekeeping and target genes levels. TaqMan® and SYBR® Green Master Mix were prepared as per the manufacture's guidelines:

	TaqMan®	SYBR® Green	Final concentration
Master Mix	5 µL	5 µL	1X
cDNA	0.5 µL	0.5 µL	50 ng/µL mRNA or 1 ng/µL for miRNA
Gene's probes/primers	0.5 µL	-	50 nM
Endogenous control probes/primers	0.5 µL	-	50 nM
Forward primer	-	0.5 µL	50 nM
Reverse primer	-	0.5 µL	50 nM
Nuclease-free water	3.5 µL	3.5 µL	-
Total volume	10 µL	10 µL	-

The reaction mixes were prepared in triplicates for each sample and target then loaded into a 96-well reaction plate. After brief centrifugation of the plate, it was loaded onto the instrument. Then, SDS 2.4 software (Life Technologies) was used to set up the two-step TaqMan® qPCR run by applying the following recommended thermal cycling parameters:

Stage	Temp (°C)	Time (min:sec)	Purpose
Hold	50	2:00	AmpErase® UNG Activation
Hold	95	10:00	AmpliTaq Gold® DNA Polymerase Activation
40 Cycles	95	0:15	Denaturing
	60	1:00	Annealing/Extension

In the case of SYBR® Green-based reactions, the following thermal cycling conditions were applied:

Stage	Temp (°C)	Time (min:sec)	Purpose
Hold	95	0:20	AmpliTaq® Fast DNA Polymerase, UP Activation
40 Cycles	95	0:1	Denaturing
	60	0.20	Annealing/Extension

2.2.2.5 Assessment of the quality of SYBR® Green-based qPCR reaction

SYBR® Green binding dye is a non-specific dye, which not only detects the target amplicon, but also any double-stranded DNA present whether it was genomic DNA contamination, primer dimers, or extra PCR product resulting from misannealed primers (King, 2010). Therefore, it is important to verify the amplification specificity of each new set of primers designed for a SYBR® Green-based qPCR reaction. One of the approaches to examine this is by analysing the PCR products by an agarose gel electrophoresis. Detecting a single DNA band per reaction of a size that matches the predicted amplicon size reflects a specific PCR reaction. A further method is the dissociation curve analysis, which might be more precise than the agarose gel electrophoresis method in identifying multiple amplicon amplification (Kochan et al., 2008). In this method, the amplified products and any double-stranded DNA sequences are dissociated at 95 °C for one minute. Then, from 55 °C to 95 °C, the temperature increases by 0.5 °C/30 seconds. During this, SYBR® Green binding dye binds to any annealed DNA sequence, while the resulting fluorescence signals are detected. When the reaction stops at 95 °C, the collected fluorescence data are plotted against their corresponding temperatures. The presence of a single major peak at a high temperature reflects specific amplicon amplification, while distinct melting peaks might indicate multiple PCR products. The dissociation curve analysis is one of the SDS 2.4 software functions within the ABI 7900HT Fast Real-Time PCR system.

2.2.2.6 qPCR data analysis

After the end of each qPCR run, the built-in RQ Manager software (Life Technologies) was used to extract raw qPCR data in order to be analysed using the relative quantification comparative delta threshold cycle (ΔC_T) method (Schmittgen and Livak, 2008), where C_T stands for the cycle number at which the fluorescence passes the detection threshold. To ensure accurate qPCR data analysis, the C_T triplicates of each gene target or endogenous control were accepted only if they were within 0.5 C_T of each other, as per the recommendation of a previous study (Tania et al., 2006). Moreover, the above analysis method requires that the endogenous control C_T

measurements are biologically and technically reproducible and have similar values across all the tested samples. Hence, the reproducibility of the endogenous control was assessed by determining the standard deviation (SD) of all samples' endogenous control C_T values to ensure that it falls below the cut-off of <1 cycle, which is the recommended cut-off to accept the stability of the used endogenous control (Hruz et al., 2011). B2M and RNU6 were used routinely as housekeeping genes when qPCR experiment was carried out to study the expression of any mRNA expression level while RNU48 was the endogenous control used for all microRNA studies. The values of the detected target were normalised to the endogenous control readings by substituting the averaged C_T values of the target gene from the averaged C_T values of the endogenous control as follows:

$$\Delta C_T = C_{T\,endogenous} - C_{T\,target}$$

The relative expression of the gene of interest to the endogenous control in each sample was then calculated using the following equation:

$$Relative\ expression = 2^{-\Delta CT}$$

qPCR runs were designed to contain a technical triplicate for each gene and endogenous control in each sample. After performing three biological replicates, the mean and the standard deviation of the technical triplicates were computed followed by computing the standard error of the mean expression level (S.E.M) for the means of the three biological replicates using the Excel software (Microsoft, USA). Finally, the analysed data were presented in a Bar Chart.

2.2.3 Detection of protein expression in cell lysates

2.2.3.1 Protein extraction

The cell lysis buffer for protein extraction was made by, firstly, dissolving a one tablet of EDTA-free Protease Inhibitor Cocktail (Roche, West Sussex, UK) in 1.5 mL dH_2O as per the manufacturer guidelines. Then, 100 µL of this protease inhibitor solution were added to 900 µL of 1X RIPA lysis buffer (Fisher). Based on its approximate size, the cell pellet was resuspended in 50 µL to 100 µL of the cell lysis solution and incubated on ice for 20 minutes to allow complete cell lysis with minimising the effect of the cellular proteases. The protein content was separated from cell debris by centrifugation the lysate for ten minutes at 13,000 g before transferring the supernatant into a new 1.5 mL tube and discarding the pellet.

2.2.3.2 Protein quantification

The colorimetric Bicinchoninic Acid Protein Assay (BCA) kit (Fisher) was used to measure the concentration of the protein extract in each sample as described previously (Smith et al., 1985). Reagent A was added to reagent B in a 50:1 ratio as per the manufacturer's guidelines, then aliquoted into a 96-well plate in a quantity of 200 µL per well. Then, 10 µL of (one in ten) diluted samples in addition to six serially diluted protein standards with known concentrations were added separately and in triplicates to the aliquoted BCA working reagent. The reaction was allowed to take place at 37 °C for 30 minutes. During the incubation time, the proteins of the standards and samples reduce the copper II in the working reagent to copper I, which subsequently interacts with the bicinchoninic acid and forms a purple-blue complex. This complex absorbs light at 570 nm, which can be detected by a spectrophotometer plate reader. The absorbance value is directly proportional to the protein quantity. Using the Excel programme (Microsoft), the absorbance measurements of the six standards were plotted against their known concentrations in order to construct a reference standard curve. To calculate the unknown protein concentration in the tested samples, their absorbance was applied to the equation of the standard curve and the results were multiplied by the dilution factor, which is ten, to give the actual protein concentration in every sample.

2.2.3.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting (WB)

SDS-PAGE was used to separate proteins from each other according to their sizes. Gel casting cassette (Bio-Rad, Hemel Hempstead, UK) was assembled according to the manufacturer's instructions followed by filling two-thirds of the gel cassette with twelve percent Resolving acrylamide gel, which contained 25% Lower Tris, 25% acrylamide gel, one percent ammonium persulfate (APS) and 0.05% tetramethylethylenediamine (TEMED) (Appendix 2). Then, 2 mL of 100% isopropanol were added to remove air bubbles and to even the edge of the gel surface. After 45 minutes, the isopropanol was removed and washed out completely with deionised water (dH_2O) before adding a stacking acrylamide gel, which contained 16% acrylamide gel, 35% Upper Tris, 0.3% TEMED and six percent APS. A comb of 1.5 mm cell size was placed in the stacking gel form the sample-loading cells. Samples to be tested were prepared by combining enough sample volume containing 50 µg protein with a sample loading dye (Appendix 2) in a total volume of 20 µL. Prepared samples were then placed in a 95 °C heat-block for five minutes in order to denature proteins completely before leaving them at room temperature for five minutes to cool down. During this time, the gel cassette was assembled with the clamping frame and placed in the system's tank which was filled with the SDS-PAGE running buffer (Appendix 2). Five microliters of EZ-Run prestained recombinant protein ladder (Fisher) were regularly loaded in the first lane to be used as a reference for proteins' molecular

weight. Then, the prepared 20 µL of each sample were transferred into the loading wells. The electrodes were fitted before allowing the electrophoresis to take place at 120 volts for around one hour to induce protein separation. The next stage was to transfer the proteins from the acrylamide gel onto a nitrocellulose membrane suitable for immunoblotting. To carry out the transferring process, six sponges, six filter papers and a nitrocellulose membrane were firstly soaked briefly in the transfer buffer (Appendix 2), then sandwiched with the gel between the covers of the transfer cassette from left to right as follows: the negative charge cover, three sponges, three filter papers, gel, nitrocellulose membrane, three filter papers, three sponges, then the positive charge cover. After fitting the assembled cassette in the electrotransfer apparatus box (Bio-Rad), it was filled with the transfer buffer before allowing protein transfer to take place for 45 minutes at 60 volts. After that, the membrane was transferred into a universal tube containing 5 mL of five percent skimmed milk blocking buffer (Appendix 2) and incubated for one hour at room temperature on a roller mixer. This step was necessary to block the unoccupied sites on the membrane by the proteins of the skimmed milk to prevent unspecific Ab binding during the following steps. Then, the blocking buffer was removed and replaced with fresh blocking buffer containing the primary Ab in the concentration recommended by the manufacturer. After an overnight incubation on a roller mixer at 4 °C, the membrane was washed three times using a one percent Tris Buffered Saline with Tween-20 (TBST) washing buffer (Appendix 2) to remove any unspecific Ab binding. Next, the membrane was placed in a fresh blocking buffer containing the manufacturer recommended concentration of the anti-primary secondary Ab conjugated with horseradish peroxidase (HRP). After a one-hour incubation at room temperature, the membrane was washed three times as before followed by removal of the excess washing solution and addition of developing ECL Prime detection reagent (Amersham), which was freshly prepared as per the manufacturer's instructions. The acridan-based substrate in the developing reagent interacts with the HRP of the secondary Ab and generates detectable chemoluminescence. The reaction was stopped by removing the substrate before exposing Amersham Hyperfilm X-ray films to the membranes in a dark room in order to detect the chemiluminescence at different intervals which ranged between two to 30 minutes. Finally, the films were developed in a film processor machine to produce exposed black bands, which reflected the detected proteins.

In addition to the above standard SDS-PAGE and WB protocol applied in this study, Abs that failed to detect their target proteins were subjected to some optimising or troubleshooting steps. For example, using different Ab or protein concentrations was the first attempt to test the Ab sensitivity, in addition to changing the Ab incubation temperature from 4 °C to room temperature. Moreover, TBS was substituted with PBS during washing buffer preparation since some Abs might work better in different buffers. The percentage of the added Tween was also decreased from 0.05% to 0.01% when faint bands were detected, which might reflect a weak Ab to protein binding or affinity. Furthermore, Bovine Serum Albuminutes (BSA) was an alternative

for the skimmed milk when a different blocking buffer was suggested. Polyvinylidene difluoride (PVDF) membrane was also used occasionally instead of the nitrocellulose membrane since they have different protein binding properties, which might affect Ab binding.

2.2.4 Detection of protein expression in tissue sections

Immunohistochemistry (IHC) was applied to detect a protein of interest in tissue samples using a specific Ab. Formalin-fixed paraffin-embedded (FFPE) tissue sections mounted on glass slides were first treated for a short-time with the organic solvent: xylene, for deparaffinisation, followed by incubation at 100% absolute ethanol to remove the xylene and hydrate the tissues. Three percent H_2O_2 in methanol was then used to block and quench the cellular endogenous peroxidase during incubation for 30 minutes. After a brief rinse in PBS, tissue sections were subjected to antigen retrieval (AR) procedure. Two different methods were used for this task according to their compatibility with the primary Ab. For an acidic AR, slides were immersed in a sodium citrate buffer (Appendix 2) and placed in a microwave for eight minutes , while basic AR was performed in a steam pressure cooker using a Tris-EDTA buffer (Appendix 2) (Lin and Prichard, 2011). AR was necessary to unmask the epitopes of the target antigen by either inducing protein unfolding or reversing the formaldehyde-induced protein cross-links or by some other proposed mechanisms (Fowler et al., 2011). In order to minimise the nonspecific staining, tissue sections were incubated at room temperature for 30 minutes with a normal blocking serum prepared from the species in which the secondary antibody was made. The primary Ab was combined with the same blocking serum in a suitable concentration (see section 2.1.4), then added to the tissue sections and incubated at 4 °C overnight.

On the following day, slides were washed twice with PBS, and then incubated for 30 minutes with a biotinylated secondary antibody supplied in a Vectastain Elite ABC kit (VECTOR Ltd, Peterborough, UK). Formation of Avidin and Biotinylated horseradish peroxidase macromolecular Complex (ABC) took place by preparing the ABC solution as per the manufacture's guidelines followed by adding it to the washed slides. After an incubation period for 30 minutes, tissue sections were washed with PBS as before then incubated in a peroxidase substrate solution for five minutes to allow stain development. The reaction was stopped by rinsing the slides in distilled water. Finally, using a counterstain machine, slides were immersed in hematoxylin, 0.1 acid alcohol (Appendix 2), scott’s tap water (Appendix 2), ethanol, and then xyelen. Counterstained slides were mounted using a Distrene-Plasticiser-Xylene (DPX) permanent mounting medium.

2.2.4.1 Scoring of immunohistochemistry (IHC) staining

A modified semi-quantification QuickScore method (Schiessl et al., 2009, Kiesslich et al., 2010, Lash et al., 2012) was used to score the IHC staining. The QuickScore categories are based on both the proportion of the stained cells (extent) and strength of the staining (intensity). The intensity corresponding to no staining, weak, moderate and strong staining, are given a score from zero to three, respectively, while the extent of the stained cells is rated as follows: 1=1% to 33%, 2=34% to 66%, and 3=67% to 100%. QuickScore is a result of multiplying the extent score by the intensity score, which yields values ranging from zero to nine.

2.2.5 DNA cloning into plasmid vectors

2.2.5.1 Cloning of a coding sequence

2.2.5.1.1 Designing the PCR primers

The coding sequence or the open reading frame (ORF) sequence of a gene of interest was initially obtained from Ensembl Genome Browser database. Specific primers were then designed to amplify the ORF but with the addition of extra sequences needed for introducing the ORF into a mammalian expressing vector; pcDNA3.1D/V5-His-TOPO (Invitrogen). As per the manufacturer's instructions, the forward primer has to have a (5'-CACC-3') base pair sequence on its 5' end to be fused during the PCR cycles to the 5' end of the cloned ORF. This sequence is necessary for the later cloning step and will not be translated with the gene's sequence since it precedes the start codon (Appendix 5).

2.2.5.1.2 PCR reaction

The Phusion high-fidelity DNA polymerase kit (New England BioLabs, Hitchin, UK) was used for any qualitative PCR reactions. The designed primers were added to the PCR reaction components as per the manufacturer's guidelines, while cDNA of normal Keratinocyte cells was used as a reaction template. Then the amplification was carried out in a PCR thermocycler with the following parameters:

Stage	Temperature	Time	No. of cycles
Initial denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	35
Annealing	63 °C	30 sec	
Extension	72 °C	1.5 min	
Final extension	72 °C	10 min	1
	4 °C	∞	

2.2.5.1.3 Assessment of PCR products using agarose gel electrophoresis

To confirm whether the PCR generated the required amplicon, the size of the PCR products were identified in a one percent agarose gel before extracting the DNA for further confirmation. The agarose gel was prepared by dissolving 500 mg agarose in 50 mL Tris Acetate-EDTA (TAE) buffer (Appendix 2), heated in a microwave to homogenise the mixture, before adding 1 µL Ethidium Bromide (EtBr). The hot gel mixture was then poured into a sealed gel cast (Bio-Rad) with a comb and left for 30 minutes to solidify. The gel was then placed in the electrophoresis apparatus filled with TAE buffer. PCR products were combined with a 6X gel loading dye in a 5:1 ratio to provide an instant visualisation of the electrophoresis process. Following removal of the comb, 30 µL of each sample were loaded into the wells of the gel, while 5 µL of GeneRuler™ 1 kb DNA Ladder (Fisher) were loaded in the first well as a reference for the DNA molecular weight. Electrophoresis then took place at 100 volts for one hour. EtBr in the agarose gel binds strongly to the DNA, absorbs Ultraviolet (UV) light and converts the acquired energy into a visible orange light (Scaria and Shafer, 1991). Light was detected using a gel imaging system and a photo for the detected DNA bands was taken to compare the size of the band to the DNA Ladder. The gel fragment that contains the DNA of the correct size was then excised and transferred into a 1.5 mL microcentrifuge tube. ISOLATE II PCR and Gel kit (Bioline) was then used to purify the DNA from the gel as per the manufacturer's instructions before measuring its concentration using a NanoDrop spectrophotometer.

2.2.5.1.4 Insertion of a gene's coding sequence into a mammalian expression vector

As per the manufacturer's instructions, 1 µL of the linearised pcDNA3.1D/V5-His-TOPO vector, which contains 10 ng of plasmid DNA, was combined with 10 ng of the purified coding sequence and incubated at room temperature for 30 minutes. The vector's Topoisomerase I (TOPO) enzyme ligates the CACC base pairs at the ORF 5'-end with its complementary sequence on the vector in addition to ligating the ORF 3' blunt-end with the vector's other end to form a circular sequence (Appendix 5).

2.2.5.2 Cloning of a gene's promoter

2.2.5.2.1 Detecting a promoter region

Successful detection of the promoter region of a gene of interest, if not already published, using *in-silico* analysis is a complex task that requires a careful and a step-wise strategy. Therefore, several online algorithms that rely on different prediction parameters were used to predict the same upstream promoter region before cross matching their output for more confirmation. Firstly, using the official gene's symbol or GenBank accession ID, its promoter was predicted using three well known algorithms: PROMOSER (Halees et al., 2003) (http://biowulf.bu.edu/zlab/PromoSer/), HsPD (http://rulai.cshl.edu/cgi-bin/CSHLmpd2/promExtract.pl?species=Human) (Zhang, 2003) and PromoterInspector (Cartharius et al., 2005) (http://www.genomatix.de/solutions/genomatix-software-suite.html). The predicted sequences were merged into one overlapped sequence then aligned against the reference DNA sequence on the University of California Santa Cruz (UCSC) Genome Browser (Meyer et al., 2013) (http://genome.ucsc.edu/cgi-bin/hgBlat) in order to add more DNA sequence upstream and downstream from the pasted sequence. For more confirmation, the new expanded sequence was extracted and re-assessed on other promoter prediction algorithms, such as Promoter Prediction Server (Knudsen, 1999) (http://www.cbs.dtu.dk/services/Promoter/) and Promoter Scan (Prestridge, 1995) (http://www-bimas.cit.nih.gov/molbio/proscan/). The DNA sequence that was frequently predicted by most of these algorithms as a promoter was considered the promoter region of the gene of interest.

2.2.5.2.2 Designing the PCR/cloning primers

The detected promoter region was planned to be cloned into a pGL3-basic luciferase reporter vector for a Dual Luciferase Reporter Assay (described in section 2.2.9). PCR primers (or cloning primers) were designed to amplify specifically the target DNA sequence, but with the addition of upstream and downstream restriction site sequences that should allow cloning it into the vector's multiple cloning site (Appendix 6). PCR reactions and products purification were performed as previously described in sections (2.2.5.1.2) and (2.2.5.1.3), respectively.

2.2.5.2.3 Mutagenic PCR

This approach was used to mutate a short sequence of DNA, which was part of a promoter sequence in this study in particular, before cloning it into a pGL3 vector. Briefly, two PCR reactions were carried out to amplify a wild-type DNA sequence as previously described in section (2.2.5.1.2) but with using one cloning primer (forward or reverse) and one mutagenic primer (reverse or forward). The cloning primers were designed to add restriction sites to the amplified DNA sequence while the mutagenic primers were designed to be complementary to the template's sequence apart from nucleotides which will appear in the amplified product instead of the wild-type corresponding nucleotides. All the mutagenic and cloning primers were designed to have roughly the same Tm and GC content properties. The two products of the first run were detected in an agarose gel to confirm their sizes, then purified as described before in section (2.2.5.1.3). In the second PCR reaction, 100 ng DNA of each amplified product were mixed and used as a template for the cloning primers only. During this reaction, the sequences of the mutagenic primers, which have been already incorporated into the amplified products, will act as a set of internal primers since they are complementary to each other leading to ligating the two sequences (Figure 2.1). Examining the PCR product in an agarose gel and purifying it were performed as described before.

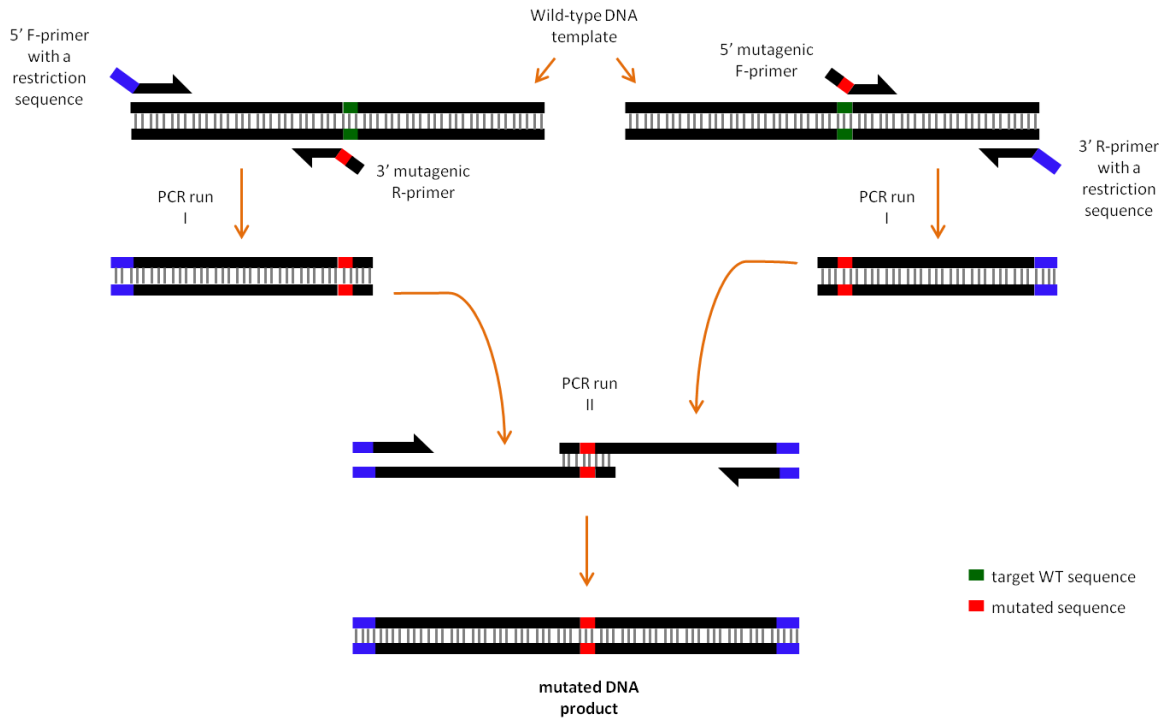


Figure 2.1: A diagram of a two-step mutagenic PCR.

2.2.5.2.4 Promoter DNA to vector ligation

Promoter sequences in addition to pGL3 vector were double digested with the same set of restriction enzymes using the manufacturer recommended recipe and conditions (New England Biolabs). The digested DNA sequences were resolved on a one percent agarose gel before purifying the digested DNA as previously described in section (2.2.5.1.3). The concentration of the purified products was measured using the NanoDrop spectrophotometer. Then, in a 1:3 molar ratio, the digested linear plasmid was added to the digested DNA to be ligated and form a circular plasmid. T4 ligase enzyme and the other components of the ligation reaction were combined as per the manufacture's guidelines (New England Biolabs).

2.2.6 Amplification and purification of a cloned vector

2.2.6.1 Bacterial transformation

In order to have a sufficient amount of the cloned vector for the subsequent applications, a cloned plasmid was transformed into *Escherichia Coli* (*E. coli*) chemically competent bacteria, which were then amplified using basic bacterial culture. Using the standard heat-shock transformation method (Hanahan, 1983), 2 µL of the previously prepared DNA/vector ligation reaction were added to one vial of chemically competent *E. coli* (New England Biolabs), incubated on ice for 30 minutes, followed by heat-shocking the bacteria at 42 °C in a water bath for 30 seconds before re-incubation on ice for five minutes. Next, 250 µL of Super Optimal broth with Catabolite repression (SOC) medium were added to the transformed *E. coli*, then the vial was incubated at 37 °C for one hour in a shaker incubator running at 180 revolutions per minute (rpm). Using the antibiotic resistance acquired by the introduced plasmids, successfully transformed *E. coli* were selected to grow on pre-prepared selective culture plates by spreading 100 µL of the bacterial culture on pre-warmed antibiotic-containing culture plates (Appendix 7). Culture plates were then incubated overnight for 18 hours at 37 °C to allow bacterial growth and colonies formation. On the following day, further amplification of the transformed bacteria was performed by picking a segregated colony and inoculating it into 5 mL of miniprep culture broth containing 100 µg/mL antibiotic (Appendix 8) followed by an overnight incubation at 37 °C in a shaker incubator running at 200 rpm.

2.2.6.2 Purification of plasmid DNA from miniprep cultures

The ISOLATE II Plasmid Mini kit (Bioline) was used to purify the cloned vector DNA from the previously prepared miniprep bacterial cultures. Briefly, bacteria were harvested by centrifugation at 10,000 g for one minute followed by removing the supernatant. Then, the pellet was resuspended and lysed under alkaline conditions. The lysis reaction was neutralised with a buffer containing an RNase enzyme to break down cellular RNA. The plasmid DNA was then adsorbed onto a silica membrane in the presence of a high-salt buffer, while RNA, cellular proteins and metabolites were removed by centrifugation at a high speed centrifugation (13,000 g) for five minutes. Finally, the plasmid DNA was washed twice and eluted in 50 µL nuclease-free H_2O.

2.2.6.3 DNA sequencing

To confirm a successful cloning of the correct DNA sequence, samples from the purified DNA were sent to the Core Genomic Facility at the University of Sheffield Medical School for sequencing on an Applied Biosystems 3730 DNA analyser. Results were then extracted using Finch TV software (Geospiza Inc., Seattle, USA) (http://www.geospiza.com/Products/finchtv.shtml) and aligned against the reference human genome using the online nucleotide BLAST (BLASTn) algorithm (http://blast.ncbi.nlm.nih.gov/).

2.2.6.4 Maxiprep amplification and extraction

To generate concentrated and ultra-pure plasmid DNA, further amplification of the transformed bacteria was performed using a maxiprep bacterial culture followed by purification using EndoFree Plasmid Maxi Purification kit (QIAGEN, Manchester, UK). Firstly, in a 500 mL clean conical flask, 200 mL LB broth medium was combined with enough selective antibiotic to give a final concentration of 100 µg/mL. After brief mixing, 3 mL of the previously prepared miniprep bacterial culture were added. The inoculated maxiprep culture was incubated overnight in a 37 °C shaker incubator running at 200 rpm. Purification was performed after 18 hours according to the kit manufacturer's guidelines. Such purification method was important to extract plasmid DNA free of bacterial lipopolysaccharide endotoxin, which can decrease the transfection efficiency in addition to its role in inhibiting cell proliferation (Butash et al., 2000).

2.2.7 Gene over-expression and silencing

2.2.7.1 In-vitro transient and stable transfection of a plasmid DNA

In this study, two main types of plasmids were transfected into some cells for different purposes: an over-expressing plasmid encoding an ORF of a gene of interest or a plasmid encoding shRNA to silence a target gene's mRNA. The transfection procedure for both approaches was the same. Briefly, cells to be transfected were seeded in 6-well culture plates in a number sufficient to reach around sixty percent confluency at the time of transfection. The number of cells varied between 50,000 and 250,000 cells/well depending on the cell's size and rate of growth. On the following day, 0.1 to 1 µg DNA in 1 µL were diluted in 144 µL Opti-MEM (Life Technologies), then 5 µL of FuGENE HD transfection reagent (Promega) were added and mixed gently. The mixture was incubated at room temperature for 30 minutes to allow the formation of a DNA/transfection reagent complex. During this time, cell media were removed, followed by washing the cells with 500 µL of Opti-MEM medium twice. Fresh 450 µL of the Opti-MEM medium were added into each well before adding the transfection complex drop-wisely. The plates were gently swirled to mix the complexes and incubated at 37 °C for 6 hours followed by addition of 600 µL of twenty percent serum-containing medium. After 24 hours, 1 mL of fresh normal media were added to each well and re-incubated in normal culturing conditions. After 48 or 72 hours from the transfection, cells were either harvested for RNA and protein expression analysis or used for functional assays. It is worth mentioning that, in case of gene over-expression, the amount of the transfected DNA was identified by studying its level when different amounts of plasmid DNA were transfected. The reference level was the basal expression level of this gene in any similar type of cells detected to naturally express it in a high level. The lowest amount of DNA found to be capable of driving an increase in the gene of interest to a level similar to the reference level was considered a suitable amount. Moreover, the transfection procedure was validated by using a green fluorescent protein (*GFP*)-expressing vector. *GFP*-encoding vector DNA was transfected into the target cells for 48 hours. Using a UV light microscope, detecting a GFP expression in sixty percent or more of the transfected cells reflects an efficient transfection procedure and protein expression.

In the case of stable transfection, cells that incorporated the transfected DNA into their genome were selected using an antibiotic. All plasmids used in this study were encoding a *Neomycin* resistance sequence. Therefore, the first step was to identify the optimal G418 antibiotic concentration that would kill all cells and spare those that acquired permanent resistance. To do this, non-transfected cells were subjected to different concentrations of G418 antibiotic to identify the lowest concentration sufficient to kill all the cells within 14 days. Transfected cells were then grown for 14 days in a medium containing similar G418 concentration. Cells that continued to grow in the presence of the antibiotic were expanded and assessed on RNA and protein levels for the changes in the expression of the transfected gene.

2.2.7.2 In-vitro transient transfection of siRNA

Short interfering RNAs (siRNA) are widely used for gene-specific silencing *in-vitro* and *in-vivo*. In this study, transfection reaction for each well was prepared by diluting 40 nM of siRNA in 100 µL Opti-MEM medium followed by adding 10 µL Oligofectamine transfection reagent (Life Technologies), then incubated for 30 minutes at room temprature. During this time, cells media were removed, followed by washing cells with 500 µL Opti-MEM medium. Fresh 500 µL of the Opti-MEM medium were then added into each well before adding the transfection complex drop-wise. The plates were gently swirled after that to mix the complexes and incubated at 37 °C for 6 hours followed by addition of 600 µL of 20% serum-containing medium. After 48 hours, the cells were either harvested for RNA and protein expression analysis or used for functional assays. It is known that many of genes silencing procedures affect cell viability. Therefore, the amount of siRNA transfected into target cells was optimised by a dose response strategy through comparing the percentage of cells remained viable after transfecting them with different amounts of siRNA. Percentage of viable cells out of all the transfected cells was calculated using the number of cells stained blue by Trypan blue stain. The appropriate amount of siRNA was determined by its ability to silence the target gene by sixty percent or more of its basal level without affecting the viability of fifty percent or more of the transfected cells.

2.2.8 Assays of cell function

2.2.8.1 Transwell cell migration assay

This experiment assessed cell migration capacity towards chemoattractants molecules through a porous membrane during a specified period of time. Stably transfected cells or cells transiently transfected for 48 hours were starved for 24 hours by removing the old medium, washing with 5 mL PBS and adding a serum-free growth medium. Depriving cells of growth factors contained in the serum is assumed to be capable of synchronising all cells to the same cell cycle stage (Kues et al., 2000), although cells synchronisation was not tested in this study. On the next day, cells were harvested as described before but using a serum-free growth medium. Cell pellets were then resuspended in DMEM containing 0.1% weight/volume (w/v) BSA before counting. A sufficient volume that contains 50,000 to 150,000 cells according to cell size and rate of growth was transferred to a new universal tube, centrifuged and then resuspended in 400 µL of fresh medium containing 1 µg/mL mitomycin C, which is an anti-proliferation agent. Following this, the Transwell components (BD Falcon, Oxford, UK) were assembled by placing cell culture inserts (upper chamber) in a 24-well culture plate (lower chamber) containing 500 µL of two percent serum-containing medium (Figure 2.2). The 400 µL cell suspension was then divided between two inserts, and the assembled system was incubated at normal cell culturing conditions for 16 hours. The growth factors in the lower chamber work as chemoattractants for the cells in the

upper chamber, which induce cells to migrate towards them through the insert's 8 µm microscopic pores and stick to the bottom surface of the upper chamber's porous membrane. However, if the incubation continued for long time, sticking cells might detach. Therefore, after the 16-hour incubation, cells activity was interrupted by removing media from both chambers. Cotton swap was used to remove completely non-migrated cells from the upper chamber. After washing both chambers with 500 µL PBS, the migrated cells were fixed by adding 500 µL of 100% methanol into both chambers for ten minutes. The methanol was then aspirated and a washing step took place followed by adding 0.1% crystal violet dye (Appendix 2) to stain the cells. After 20 minutes, the dye was removed and chambers were washed with PBS as before. Under x20 microscopic magnification and in the presence of 500 µL PBS in every chamber, the porous membrane was divided virtually into four areas. Three random digital images were taken using a microscope-attached camera. Cells were counted in each picture using the ImageJ software (available on: http://rsb.info.nih.gov/ij/). Each experiment was designed to include three technical replicates and was repeated three biological times. Excel programme (Microsoft) was used to compute the mean and the standard deviation of the numbers of the transfected and negative control cells. Also, it was used to compute the mean of the three biological repeats and the standard error of the mean (S.E.M) and present the data by a Bar Chart. The total numbers of the transfected and negative control cells were analysed statistically by applying the parametric student's *T*-test to compute the significance *p*-value.

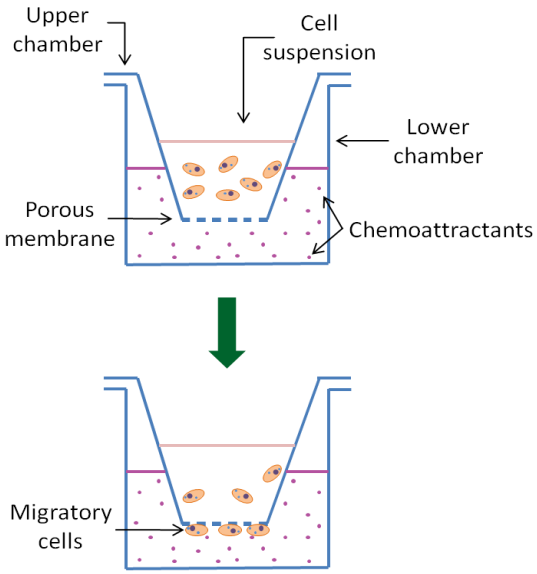


Figure 2.2: A diagrammatic representation of Transwell migration assay.

2.2.8.2 Transwell cell invasion assay

This experiment was carried out to examine the ability of the transfected cells to invade a Matrigel-coated membrane. The same procedure of the Transwell migration assay was applied apart from coating the top surface of the porous membrane with a thin layer of Matrigel (BD Biosciences, Oxford, UK). Matrigel is composed of basement membrane extra cellular matrix (ECM) proteins, primarily laminins and collagen IV, isolated from the Englebreth-Holm-Swarm mouse sarcoma (Guan, 2005). The day before setting up the experiment, Matrigel was diluted on ice in DMEM in a 1:45 ratio then directly 100 µL of the diluted Matrigel were added to the upper chamber (Figure 2.3). The system chambers were then assembled and incubated overnight at 37 °C incubator. This incubation allowed the Matrigel to form a semi-gel barrier which makes it resemble the native ECM in its physical and chemical properties. Invasive cells that could produce proteolytic enzymes to digest the Matrigel were then detected on the bottom side of the porous membrane and analysed as mentioned earlier in the Migration assay.

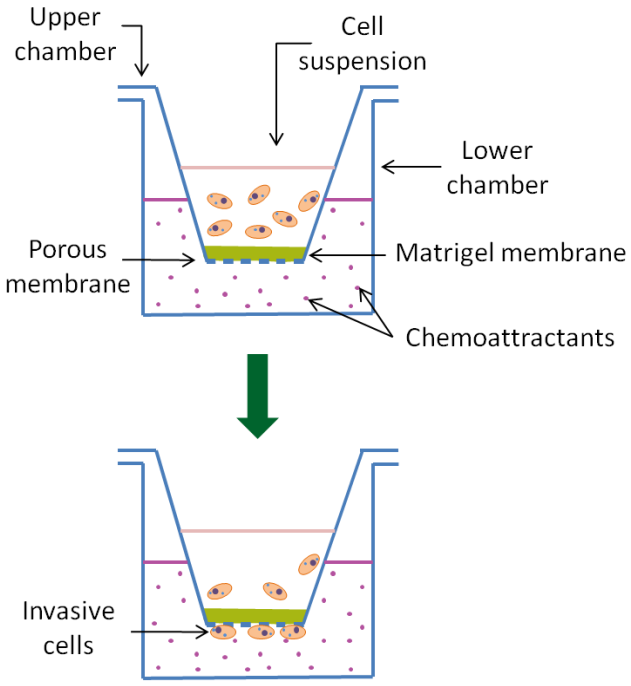


Figure 2.3: A diagrammatic representation of Transwell invasion assay.

2.2.8.3 Cell adhesion assay

This assay was used to test cell ability to adhere to a surface coated with firbronectin within one hour. Briefly, 100 µL of diluted fibronectin (1:100 PBS) were added to a 96-well plate and incubated at 4 °C overnight to allow enough time for the fibronectin to coat the well surface. On the following day, wells were washed with PBS before adding 100 µL of a one percent (w/v) BSA in DMEM solution into each well to block any fibronectin unoccupied parts. During the incubation time, stably or transiently transfected cells were harvested in serum-free medium and counted as previously described. Wells were washed with PBS and 30,000 cells were seeded into each well followed by a one hour-incubation at 37 °C. Non-adherent cells were removed by a PBS washing followed by adding 100 µL normal medium (Figure 2.4). Then, 20 µL of MTS reagent (Promega) were added to each well as per the manufacture's guidelines. Mitochondrial dehydrogenase, in living cells, reduces the MTS tetrazolium compound resulting in the production of a coloured formazon. After a one-hour incubation time, the developed colour was measured at 490 nm wavelength using a microplate spectrophotometer. The intensity of the detected colour was directly proportional to the number of adherent cells. Raw data of the technical replicates were exported to the Excel programme (Microsoft) in order to average the readings of the transfected and negative control cells and calculate the standard deviation. The mean of the three biological repeats was also computed in addition to the standard error of the mean (S.E.M) and represented by a Bar Chart. The total readings of the transfected and negative control cells were analysed statistically by applying the parametric student's *T*-test to compute the significance *p*-value.

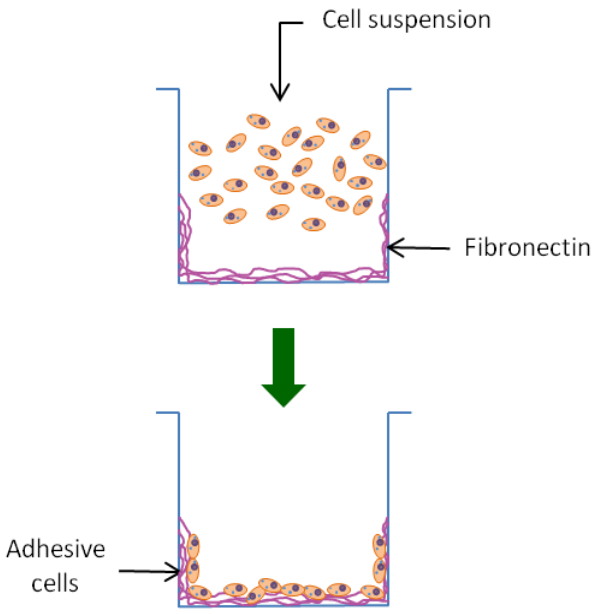


Figure 2.4: A diagram of cell adhesion assay.

2.2.8.4 Cell proliferation assay

This assay was applied to measure the increase or decrease in cell growth over four consecutive days. Briefly, cells were harvested and counted as described before. Using a 96-well plate, three wells were assigned for each time point; 1 hour, 24 hours, 48 hours, 72 hours and 96 hours, then 5000 cells were seeded into each well. At each time point, dead cells were removed by a PBS wash followed by addition of 100 µL fresh media. Then, 20 µL of MTS reagent (Promega) were added as per the manufacture's guidelines. After a one-hour incubation time, the developed colour was measured at 490 nm wavelength using a microplate spectrophotometer. The intensity of the detected colour was directly proportional to cells' number. Raw data of the technical replicates and biological repeats were analysed as previously described in the invasion assay. The total readings of the transfected and negative control cells at the time point (96 hours) were analysed statistically by applying the parametric student's *T*-test to compute the significance *p*-value.

2.2.9 Dual luciferase reporter assay

For studying a potential interaction between a gene's promoter and a transcription factor, the dual-luciferase reporter (DLR) assay is the most appropriate assay. In this experiment, the pGL3 reporter vector with cloned promoter region previously described in section (2.2.5.2), was transfected into cells already stably transfected with a transcription factor, such as HOXD10. If such a transcription factor can bind to the promoter, it will either induce or suppress the expression of the encoded firefly luciferase enzyme. The proportional amount of expression was then measured by detecting light in the 550 to 570 nm wavelength range, generated when Luciferin substrate molecules are added (Figure 2.5). However, to confirm the transfection efficacy and to normalise the firefly luciferase expression among all the tested cells, cloned pGL3 vector was co-transfected with a pRL-TK *Renilla* internal control vector which contains the herpes simplex virus thymidine kinase (HSV-TK) promoter. From its substrate, *Renilla* luciferase produces light of 480 nm wavelength. Since such a simultaneous transfection of two reporter vectors might cause trans effects between the promoter elements (Farr and Roman, 1992), the potent constitutive active pRL-TK *Renilla* reporter vector was co-transfected with pGL3 firefly luciferase reporter vector in a ratio of 1:10. Forty eight hours following transfection, cells were lysed as per the manufacturer's instructions. Using a Plate-Reading Luminometer (Promega) and the substrates provided in the Dual-Luciferase Reporter Assay kit (Promega), cell lysates were analysed to detect the expression level of firefly and *Renilla* luciferase. The data were then exported into an Excel file for normalisation and statistical analysis.

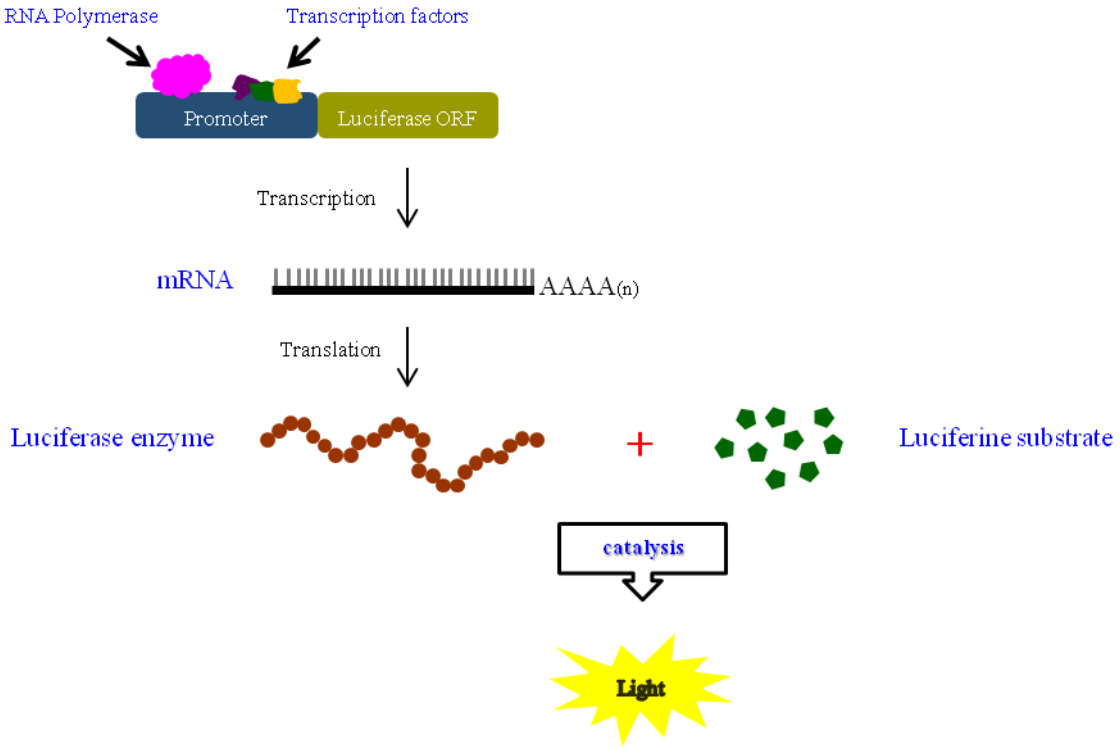


Figure 2.5: A diagrammatic sketch of the principle of luciferase reporter assay.

2.2.10 Microarray

Agilent One-Colour Microarray-based Gene Expression Analysis protocol (version 6.5, May 2010) (Agilent, Wokingham, UK) was applied to perform the microarray experiment in this study. Below are brief descriptions of the main steps.

2.2.10.1 RNA preparation and quality control

Total cellular RNA was isolated as previously described in section (2.2.2.1). In order to reduce biases in microarray analysis caused by poor RNA quality, extracted RNA was assessed for its concentration, purity and integrity using the NanoDrop Spectrophotometer and the Agilent 2100 Bioanalyzer instrument (Agilent). The below conditions, which are recommended by the manufacturer and some other studies (Zhang et al., 2004), had to be met in order to consider any RNA sample for the subsequent steps.

Requirements	Range	Reason
RNA concentration	10 - 200 ng/µL	To be within the range of the labelling kit
260/280 nm absorbance ratio	≥ 1.9	To exclude protein contamination
260/230 nm absorbance ratio	≥ 1.9	To exclude chemical contamination
18S/28S Ribosomal ratio	≥ 2	To exclude degraded RNA
RNA Integrity Number (RIN)	≥ 9	

2.2.10.2 Sample labelling and hybridisation

For each sample, serial dilutions of Agilent One-Colour Spike-In Mix (Agilent), which contains 10 pre-designed positive control transcripts, were mixed with a total RNA input amount of 100 ng/dilution. Following a multi-step procedure, sample RNA together with the internal control transcripts were labelled with Cyanine 3-CTP (Cy3) dye during an amplification reaction using the Agilent Low Input Quick Amp kit. The generated fluorescent complimentary RNA (cRNA) products were then purified using Qiagen's RNeasy Mini kit. Next, the yield of the linearly amplified cRNA and the Cy3 specific activity were quantified using the Microarray Measurement function of the NanoDrop spectrophotometer. To reduce its structure complexity, 600 ng cRNA of each sample were fragmented into approximately 50-200 bases long. The fragmentation step

took place at 60 °C for 30 minutes. Such fragmentation improves cRNA specificity and binding efficiency to the oligo arrays. After this, 25 µL of the fragmentation reaction were combined with 25 µL hybridisation buffer and immediately loaded onto the gasket slide and assembled to the microarray slide, which contains 8x60K whole Human Genome Oligo arrays. Assembled slide chambers were placed in a hybridisation rotator set to rotate at 10 rpm inside a 65 °C hybridization oven (Figure 2.6). After 17 hours, hybridised slides were disassembled, washed and then scanned at 3 µm resolution using Agilent C Microarray Scanner (Agilent) pre-set with the default settings for a 8x60K G3 Microarray Format.

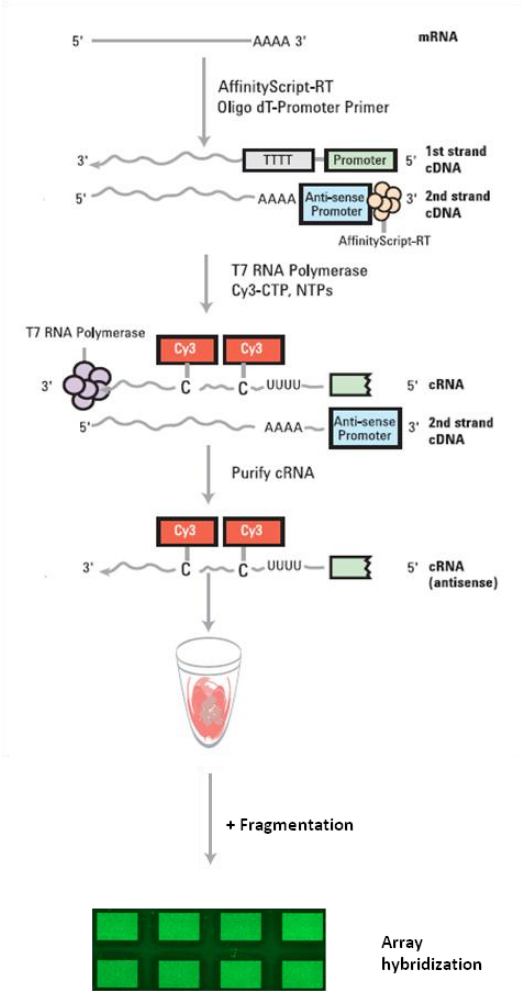


Figure 2.6: Steps of the Agilent one-colour microarray sample labelling and hybridisation.

The sketch was adapted from the manufacturer's manual.

2.2.10.3 Data acquisition and analysis

After detecting the fluorescence signals of the hybridised slides, the Agilent Feature Extraction (FE) Software (version 10.7) was used to process the generated image. A pre-defined protocol (GE1_107_Sep09) that matches the slide type was initiated to import a set of parameter values and settings, which allow the software to align the default grid to the image. Following identification of the image spots, the software automatically performs a background correction and dye normalisation by subtracting the background intensities from the foreground intensities and scales average signal intensity for each sample to the average signal intensity for all samples in order to flag and reject the outliers and low quality probes. This was followed by an automatic computing of feature log ratios (Agilent's Processed Signal value) and their *p*-values. The FE preliminary processing ends with production of a QC report for each array image, which include statistical results with threshold values useful to evaluate the reproducibility and reliability of the microarray, and an exportable raw data file in a "txt" format, which contains all the parameters, statistical calculations and the annotation information associated with the Agilent microarray used in the experiment.

Raw "txt" data files were then analysed using Agilent GeneSpring GX software (version 12.1) (http://www.genomics.agilent.com/article.jsp?pageId=2141). Briefly, intensity values were subjected to a $\log_2$ transformation and normalised to the 75th percentile-shift normalisation. The imported data files were then grouped and assigned to the experiment parameters and conditions. For instance, data files of cells transfected with HOXD10 were assigned to the condition "Overexp", while the control-transfected cells were assigned to the condition "Low". Both conditions were grouped under the parameter, "over-expression", which allows internal comparison in the same group or even a comparison with another parameter. Next, the software automatically computed the arrays' Principal components analysis (PCA) scores and presented them in clusters in a 3D scatter plot to allow a QC check before the next analysis. Based on visual analysis, outlier cluster of each triplicate array was excluded from the subsequent analysis. To identify differentially expressed genes, statistical analysis was performed by either using *T*-test, when comparing two conditions under one parameter, or one-way ANOVA statistical test (Welch), when comparing a pair of two conditions under two parameters simultaneously with applying Benjamini-Hogberg multiple testing correction method to correct the computed *p*-values. Finally, the list of the identified genes was filtered using the built-in fold change function before exporting the list into an Excel file.

2.2.10.4 Gene Ontology (GO) and pathway analysis

GO enrichment analysis was conducted in the web-based functional annotation GeneRanker tool (part of the Genomatix software suit; http://www.genomatix.de/cgi-bin//GeneRanker/gene_ranker_web_gui.pl?s=80a3f65115a178c0239f569315c23021;TASK=gene_ranker;generanker_input=1). Exploring genes' interactions and mapping them to different pathways were carried out in the web-based Genomatix Pathway System (GePS) tool (part of the Genomatix software suit; http://www.genomatix.de/cgi-bin//GeneRanker/gene_ranker_web_gui.pl?s=80a3f65115a178c0239f569315c23021).

2.2.1 Statistical Analysis

Non-parametric Kruskal-Wallis test (one way ANOVA) was performed on the IHC tissue sample scoring using SPSS 21 software (http://www.ibm.com/software/uk/analytics/spss/). Parametric student's *T*-test was used to calculate all other significance values, unless otherwise stated.

2.2.2 Ethics

Ethics approval for the use of biopsy tissues in this study was obtained from The West Paisley LREC (ref: 08/S0709/70). Ethical approval for the normal oral keratinocyte primary cultures used was obtained from Sheffield Research Ethics Committee 09/H1308/66.

CHAPTER 3: DIFFERNTIAL AND ALTERED HOXD10 EXPRESSION IN HNSCC AND NSCLC CELLS

3.1 Introduction

Microarray study is a widely used approach to screen the global gene expression profile of different types of cells. However, the resulting differential gene expression cannot be taken for granted unless verified by independent means. qPCR analysis is the method of choice to validate microarray-generated data (Provenzano and Mocellin, 2007). In their microarray study, Dr Hunter and colleagues identified high level of *HOXD10* transcripts in primary tumour cells compared to very low level detected in normal oral keratinocytes (OKs) and cells derived from lymph node metastases, while oral pre-malignant lesion (OPL) cells showed inconstant level of HOXD10. To verify these findings, some normal oral keratinocytes, OPL, primary HNSCC, metastatic HNSCC and NSCLC cell lines were cultured according to the standard procedures described earlier in order to extract their total RNA for a qPCR study. Since high cell density can affect RNA transcription level (Hannan et al., 2000), cells were harvested for RNA and protein extraction when they reached 70% to 80% confluency. RNA of normal lung epithelial cells (LEC) was courtesy of Dr Colin Bingle.

Manipulation of the expression level of a gene of interest is a commonly used method to study its functions and also to identify its interacting molecules. Moreover, such a sudden change in the level of a gene might cause cells to gain or lose some phenotype such as migration, invasion, adhesion and proliferation. The mechanisms and cellular signalling controlling these phenotypes are believed to overlap. For example, cell migration, which can be described as the ability of cells to change their position within tissues or organs by a directed movement (Kramer et al., 2013), is a complex mechanism that involves organised mechanical and biochemical changes in cell morphology in addition to an on-and-off cell adhesion to the surrounding matrix (Lauffenburger and Horwitz, 1996). The ability of cells to produce ECM-digesting proteases might not be necessary for their motility, while it is a very important feature for their invasion, which is defined as the penetration of tissue barriers through degrading and remodelling the ECM (Friedl and Wolf, 2003, Kramer et al., 2013).

Our gene of interest, *HOXD10*, is one of the many human transcription factors that share a primary function, which is binding specifically to a DNA sequence and recruiting RNA polymerase II in order to initiate or repress the expression of different genes (reviewed by Vaquerizas et al., 2009). The downstream effect of this mechanism of gene regulation mediates cell function and fate. Hence, changing natural HOXD10 expression levels in some HNSCC and NSCLC cells may reveal some of its functional roles in addition to confirming its effect on *miR-7* and *IGFBP3*, which were identified to be downstream targets of HOXD10 in breast and gastric cancers, respectively.

The effect of gene manipulation on cell function can be studied through a variety of *in-vitro* assays that mimic, to some extent, the *in-vivo* phenotypes. Here, the effect of HOXD10 manipulation on cells migration, adhesion, proliferation and invasion is reported. Moreover, the

effect of HOXD10 manipulation on *miR-7* and *IGFBP3* expression levels was screened. It is worth mentioning that most of the functional assays were performed early in this project using transient HOXD10 manipulation procedures, except for cell invasion assay, which was carried out at a later stage using stably transfected cells. This was due to demonstrating an inconsistency in cell invasion results when HOXD10 was transiently manipulated, while stable over-expression or knocking down showed more reproducible results. The functional assays carried out using transient transfection were briefly repeated using stably transfected cells. The results were consistent and highly similar. The previously reported HOXD10 target molecules were also re-assessed in stably transfected cells. The results were no different from what have been demonstrated in the case of transient HOXD10 manipulation.

3.2 HOXD10 mRNA and protein level in a panel of HNSCC and NSCLC cell lines

Prior to performing a qPCR analysis, quality of the extracted RNA was assessed using the NanoDrop spectrophotometer. Only RNA samples with $A_{260/280}$ and $A_{260/230} \geq 1.9$ were used for cDNA preparation and qPCR analysis (Figure 3.1).

Validation of the initial microarray data by qPCR demonstrated that HOXD10 is indeed over-expressed in primary HNSCC and NSCLC cell lines compared to low levels in (OKs) and LECs, respectively. However, the high expression is not seen in cells derived from lymph node metastases, B22 and TR146. Expression of HOXD10 in the OPLs was variable, although some cell lines were found to have high HOXD10 expression level, for example, D35 (Figure 3.2, A1 and A2). For significance analysis, HOXD10 levels in each cell type were grouped for statistical analysis using the non-parametric Kruskal-Wallis test. The high HOXD10 level in HNSCC cells was statistically significant compared to all the other cell types (Figure 3.2, B1). Similar significance was detected in the case of NSCLC compared to normal cells (Figure 3.2, B2). Moreover, in a smaller panel of random cell lines, HOXD10 protein level matched its mRNA level (Figure 3.3).

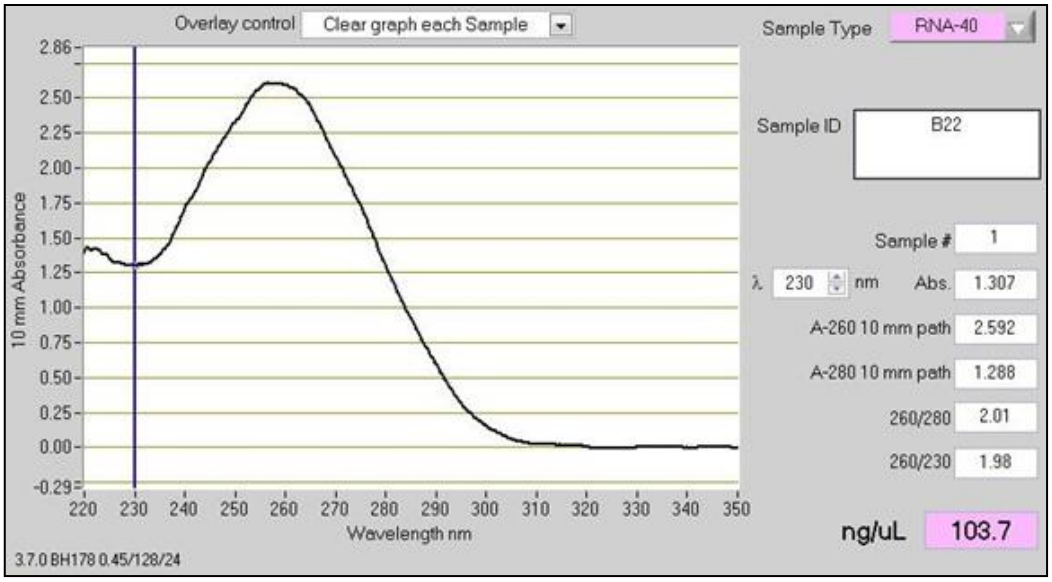


Figure 3.1: A screenshot for a result of RNA sample analysis using the NanoDrop spectrophotometer.

The concentration and purity of any extracted RNA sample were measured using the NanoDrop spectrophotometer. RNA samples of $A_{260/280}$ and $A_{260/230}$ more than or equal to 1.9 were used for cDNA preparation and qPCR analysis.

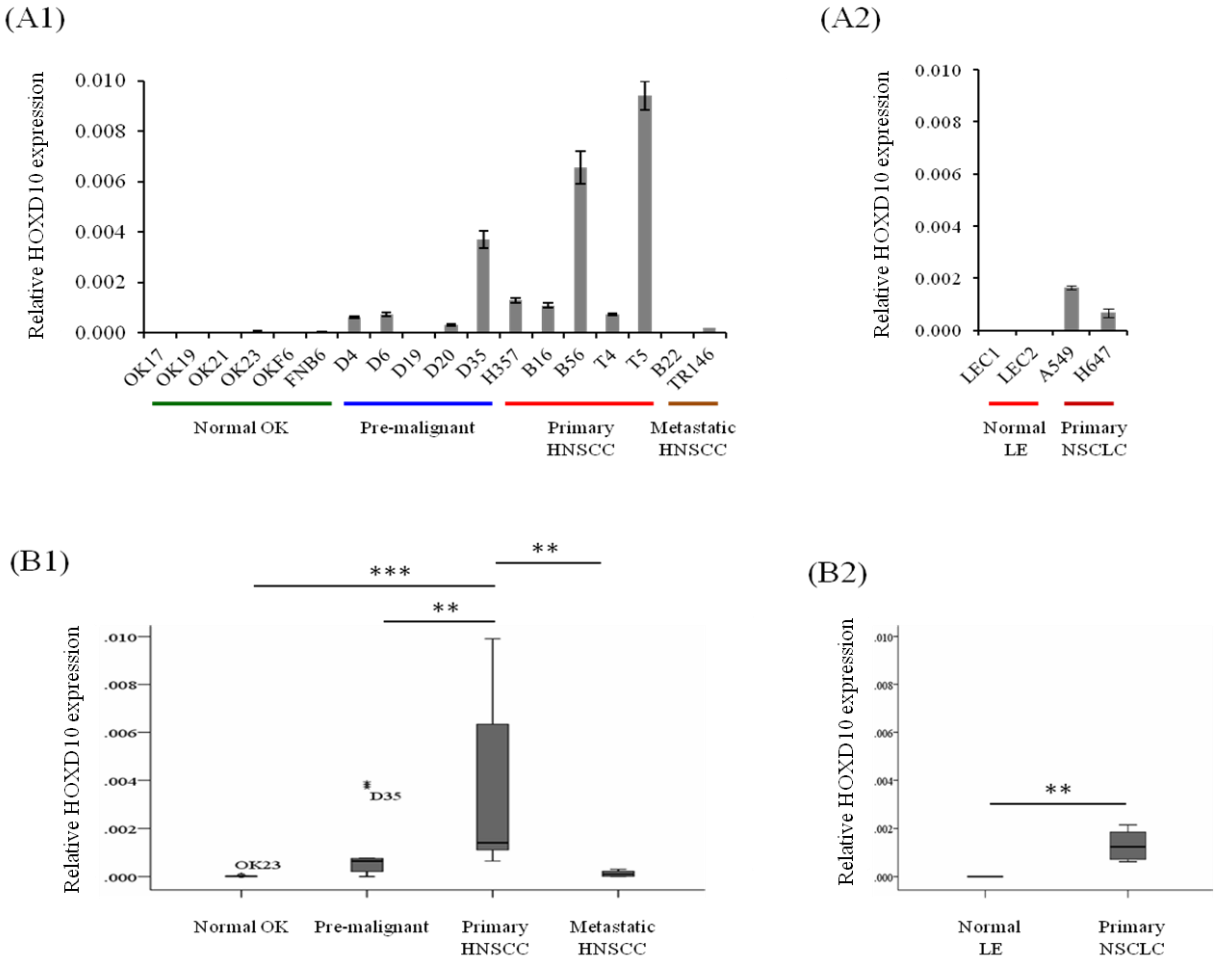


Figure 3.2: HOXD10 mRNA level in a panel of HNSCC and NSCLC cell lines.

(A1) The level of HOXD10 expression is high in HNSCC tumour cells, variable in pre-malignant cells, compared to a very low amount detected in normal cells and metastatic cell lines. (A2) NSCLC cells also showed high level of HOXD10 expression compared to normal lung epithelial cells (LEC). (B1) HOXD10 levels in each cell type were grouped for statistical analysis using the Kruska-Wallis test. The high HOXD10 level in HNSCC cells was statistically significant compared to all the other cell types. (B2) Similar significance was detected in case of NSCLC compared to normal cells (** *p*-value<0.01, *** *p*-value < 0.001, *T*-test (primary tumour vs. each group), data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

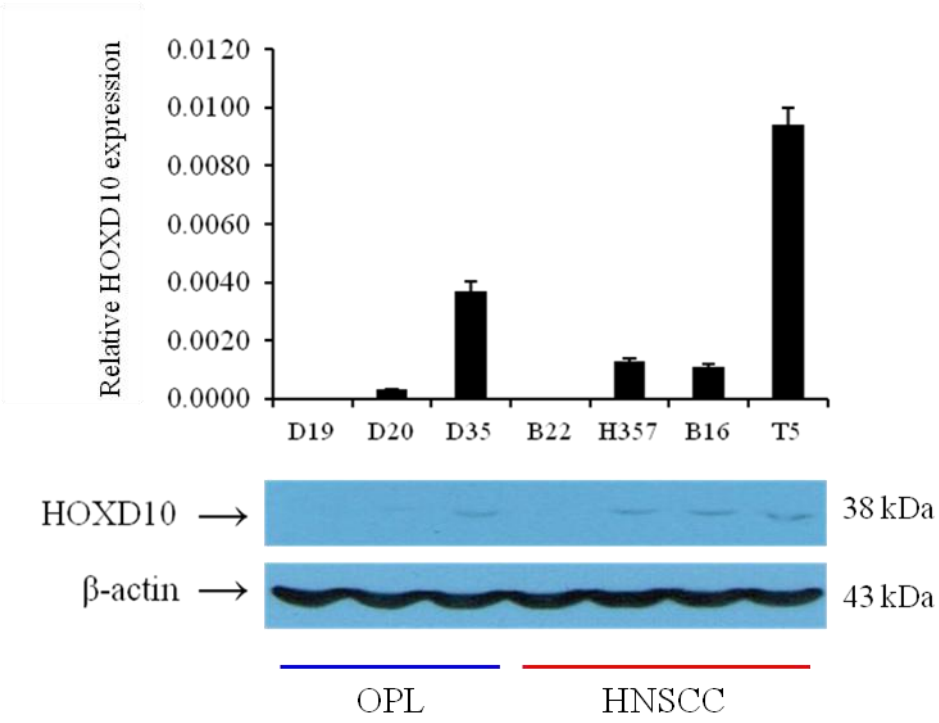


Figure 3.3: HOXD10 protein level in a panel of cell lines.

A bar graph of HOXD10 mRNA level in a panel of HNSCC cells as determined by qPCR (upper panel). Western blot analysis of HOXD10 protein in the same panel of HNSCC cells (lower panel). Beta actin protein expression is shown as a control for loading.

3.3 HOXD10 protein level in HNSCC tissue sections

The previously used Santa Cruz's HOXD10 Ab in WB did not work well with IHC protocol. Therefore, other commercially available antibodies were tested for both WB and IHC. Only one Ab out of seven was capable of detecting HOXD10 specifically in tissue sections:

Table 3.1: The Abs tested to detect HOXD10 protein in cell lysate and tissue sections.

Manufacturer	Host and clonality	Cat number	Tested applications	Optimised for
Santa Cruz	Goat polyclonal	sc-33005	WB, IHC	WB
Santa Cruz	Rabbit polyclonal	sc-66926	WB, IHC	-
Abcam	Rabbit polyclonal	ab90704	WB, IHC	IHC
Abcam	Rabbit monoclonal	ab138508	WB	-
Assay Biotech	Rabbit polyclonal	C16318-1	WB, IHC	-
Novus Biologicals	Rabbit polyclonal	47570002	WB	-
Millipore	Rabbit polyclonal	ABE128	WB	-

Prior to considering the Abcam anti-HOXD10 Ab to detect HOXD10 protein in HNSCC tissue sections, its specificity was assessed using a pre-adsorption method. An overnight incubation of 2 µg of the HOXD10 Ab with 20 µg of HOXD10 peptide, which compose part of HOXD10 protein, resulted in neutralising the binding of HOXD10 Ab. As can be seen in Figure 3.4, the pre-adsorbed Ab was not able to detect HOXD10 protein, which is normally highly expressed in prostate tissues (Redline et al., 1992). It is worth mentioning that this HOXD10 Ab was tested to detect HOXD10 protein in cell lysates but no bands were detected.

Analysis of HOXD10 protein expression in a tissue microarray (TMA) constructed from a cohort of normal oral mucosa and tissues from different phases of HNSCC development demonstrated increased expression of HOXD10 in primary tumours compared to normal tissues, and loss of expression in metastases, which is similar to expression changes observed *in-vitro* (Figure 3.5). In addition, HOXD10 protein in different grades of dysplasia ranged between high and low level.

Further IHC analysis in another TMA constructed from a cohort of 27 matched HNSCC primary tumours with their metastases confirmed the pattern with expression lower in the metastases of 23/27 patients (85%) (Figure 3.6).

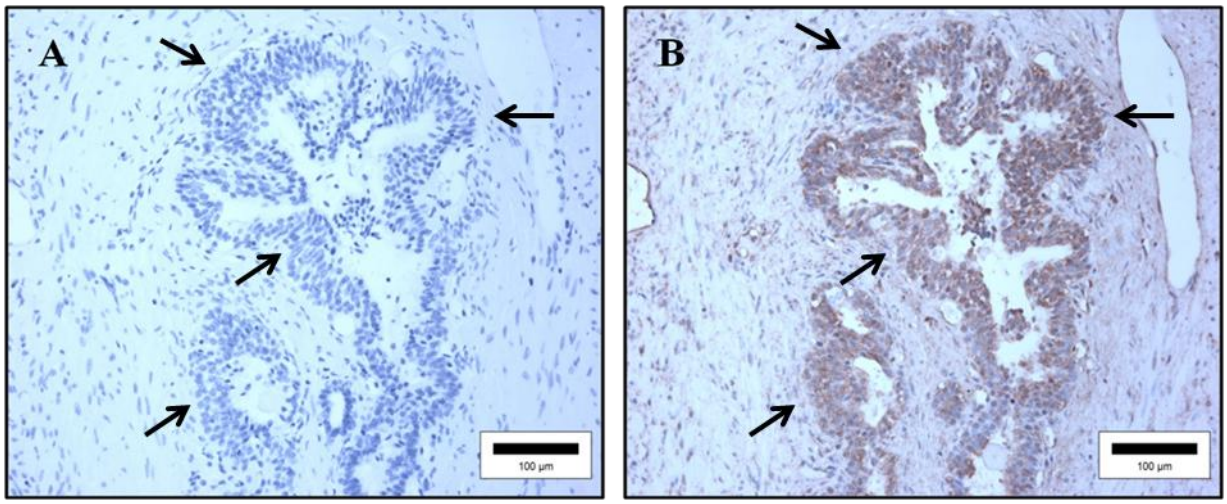


Figure 3.4: Photomicrographs of IHC staining for HOXD10 in prostate tissue sections.

(A) The added HOXD10 peptide pre-adsorbed HOXD10 Ab and blocked it from detecting HOXD10 protein in prostate tissue sections. (B) Non-neutralised Ab bound to HOXD10 proteins leading to the formation of a strong staining.

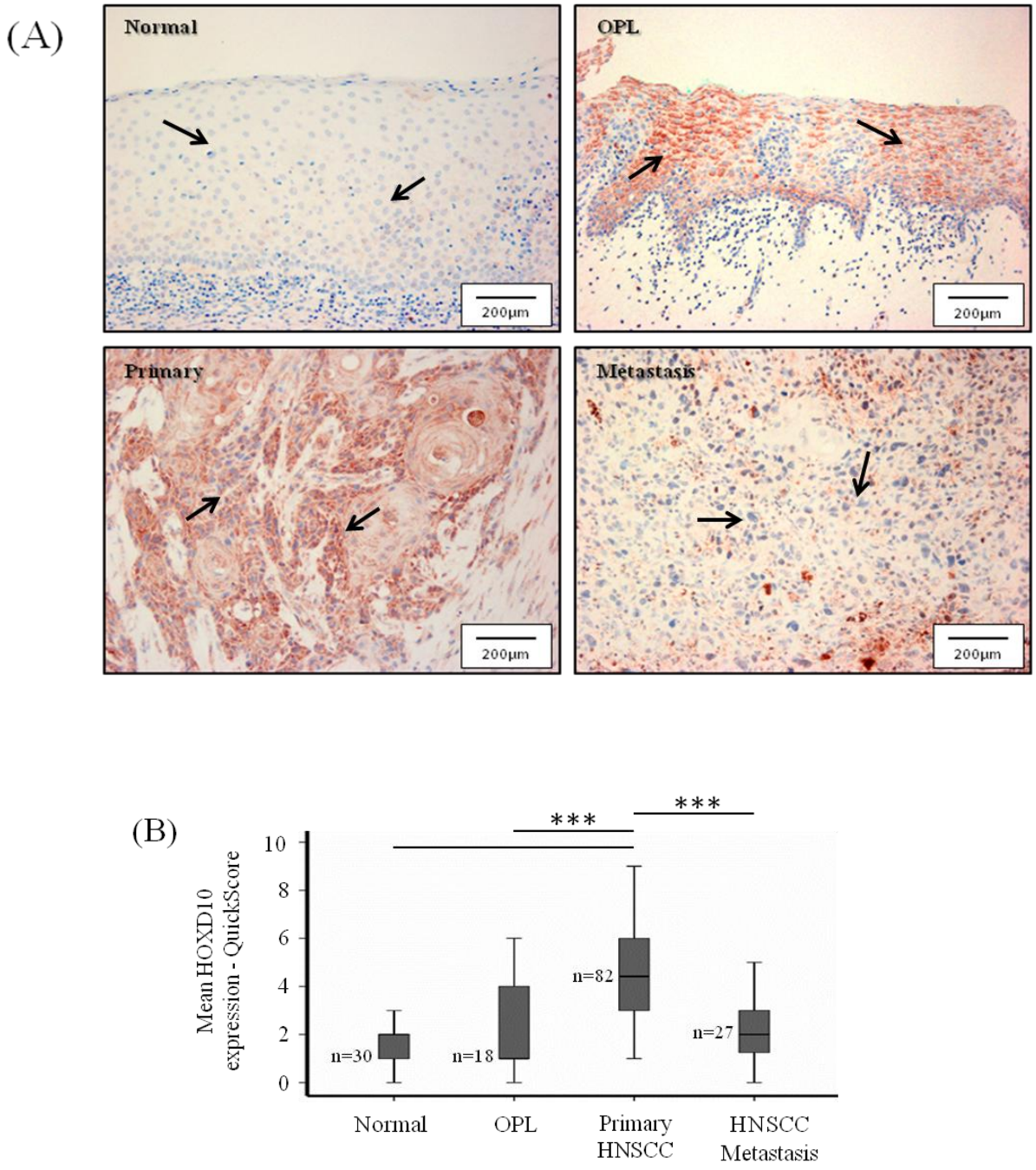


Figure 3.5: HOXD10 protein level in a cohort of HNSCC tissue sections.

(A) Stain intensity for normal oral mucosa and different HNSCC stages are represented by some representative photomicrographs. (B) HOXD10 protein found to be significantly expressed in tumour tissues variable in dysplastic, but very low in normal tissues and metastases. Kruskal-Wallis (One-way ANOVA) test was used to compute the statistical significance of the scores variance of all HNSCC stages against the score of normal tissues (*** $p < 0.001$, non-parametric Kruskal-Wallis test, n=number of samples).

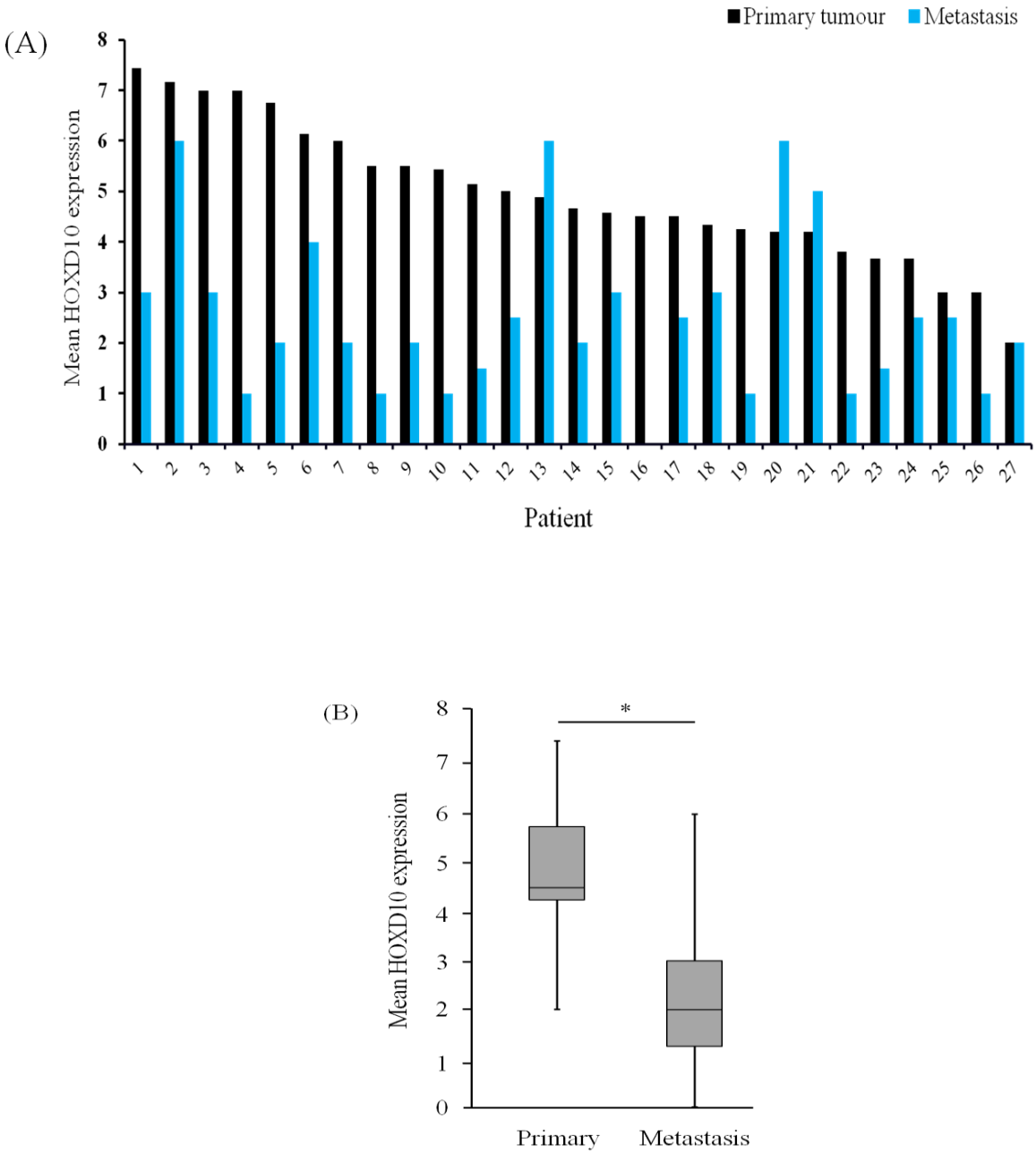


Figure 3.6: A screen for HOXD10 protein level in primary tumours vs. metastasis.

Cancer tissue sections were prepared from patients who have both primary HNSCC and metastatic disease. (A) Twenty three patients out of 27 showed higher HOXD10 level in their primary tumours compared with lower expression level detected in metastatic sites. (B) Statistical analysis using the parametric student's *T*-test for the whole two groups (Primary tumours vs. Metastasis) showed a significant difference between their HOXD10 levels (* $p < 0.05$).

3.4 Basal expression level of HOXD10 versus its previously reported target molecules

As previously mentioned, HOXD10 was found to regulate and to be regulated by *miR-7* and miR-10, respectively, in breast cancer. Moreover, *IGFBP3* was also identified as a direct target of HOXD10 in gastric cancer. In this study, HOXD10 expression and the expression of IGFBP3, miR-7 and miR-10b were compared through qPCR analysis in a panel of HNSCC and NSCLC cells. The results show no clear correlation between the basal level of HOXD10 expression and its targets, *miR-7* and *IGFBP3*, in the tested cell panel (Figure 3.7). Moreover, the expression of HOXD10 in different cell types showed no correlation with its regulator, miR-10b, suggesting independent regulation of expression (Figure 3.8).

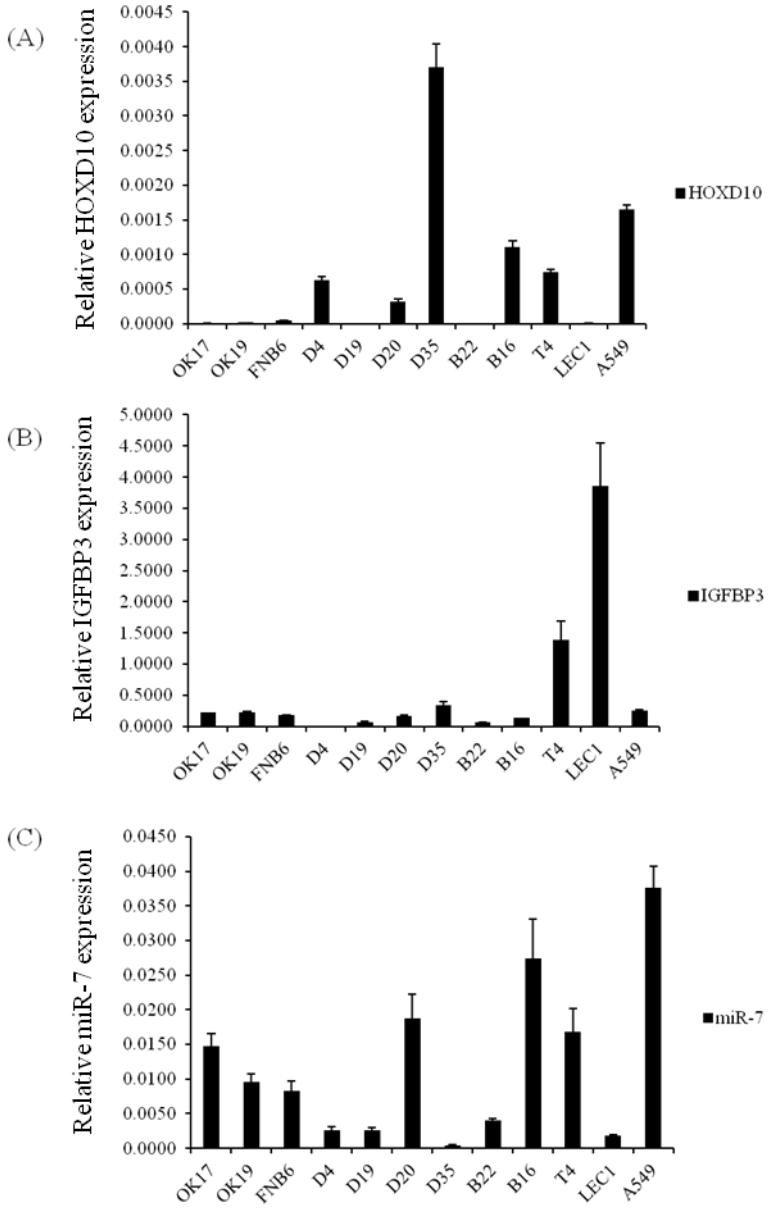


Figure 3.7: HOXD10 basal expression level vs. its previously reported downstream targets, *IGFBP3* and *miR-7*, in normal oral keratinocytes and a panel of HNSCC cell lines.

(A) qPCR analysis shows that the level of HOXD10 across different types of HNSCC cells and normal cells does not correlate with the detected amount of *IGFBP3* (B) or *miR-7* (C) (Data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

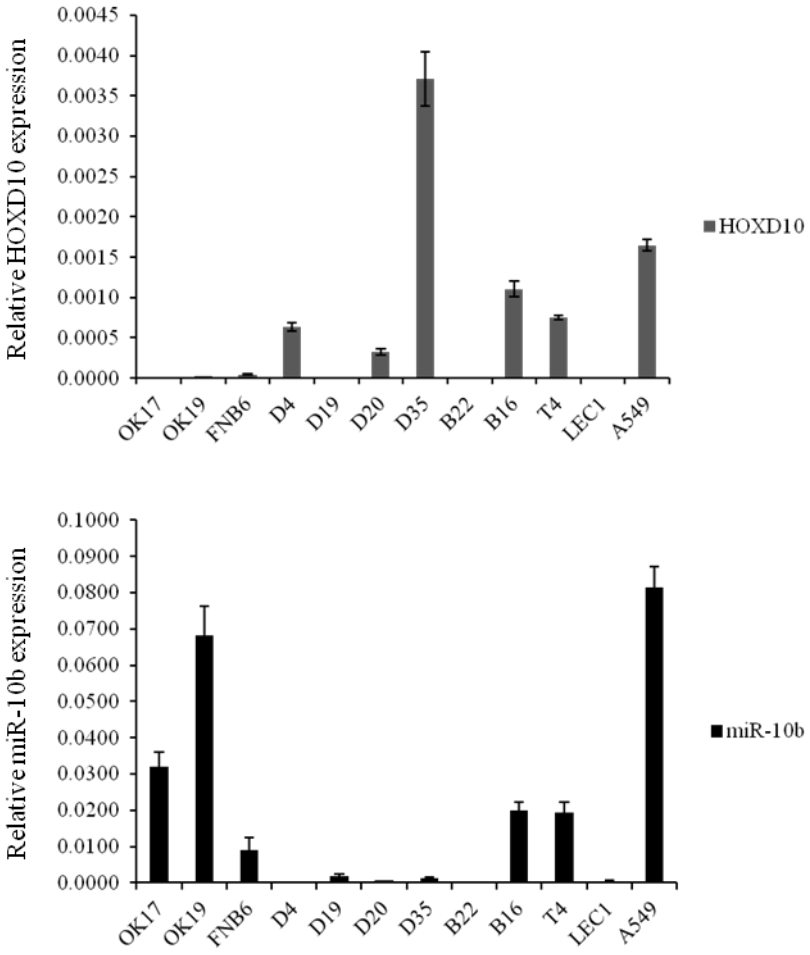


Figure 3.8: HOXD10 basal expression level vs. previously reported regulator, miR-10b, in normal oral keratinocytes and a panel of HNSCC cell lines.

According to qPCR analysis, no strong negative correlation or any association between HOXD10 and miR-10b levels was demonstrated (Data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

3.5 *In-vitro* HOXD10 over-expression and knocking down

HOXD10 coding sequence was extracted from Ensembl Genome Browser (Flicek et al., 2013), amplified, purified, then cloned into pcDNA3.1 mammalian expression vector (pcDNA3.1-HOXD10) using the following cloning primers: forward primer, 5'-CACCATGTCCTTTCCCAACAGCT-3', reverse primer, 5'-CTAAGAAAACGTGAGGTTGGCGG-3'. Three naturally low-HOXD10 expressing cells, NSCLC H647 primary tumour cells, HNSCC B22 metastatic cells and D19 OPL cells, were randomly chosen for a transient transfection with either HOXD10-encoding pcDNA3.1 or pcDNA3.1 control vector. The transfection procedure was optimised using a GFP-encoding vector. A GFP expression in sixty percent or more of the transfected cells reflected a successful transfection (Figure 3.9, A). On the other hand, pre-malignant cells, D35, HNSCC primary tumour cells, T5 and NSCLC primary tumour cells, A549, were subjected to a decrease in their cellular HOXD10 level using HOXD10 siRNA. During optimisation of HOXD10 knock down, it was noticed that targeted knockdown of HOXD10 mRNA by ninety percent or more of its normal level using a high amount of siRNA duplexes, such as 60 nM/well or more, significantly decreases cell viability and killed most of the transfected cells within 48 hours to 72 hours compared to control cells, which were not affected when transfected with a similar amount of the universal scrambled negative control siRNA duplex. Moreover, morphology of the dying cells was a mix of apoptosis (cell shrinkage) and necrosis (cell expansion) (Ziegler and Groscurth, 2004). It was found that 40 nM/well is the optimal concentration that caused a significant knocking down of HOXD10 by sixty percent of its normal level without causing a major reduction in cell viability (Figure 3.9, B). Total RNA and protein were extracted from the transfected cells for qPCR and WB analysis, respectively. Cells transfected with pcDNA3.1-HOXD10 plasmid showed stronger intracellular expression of HOXD10 mRNA and protein than cells transfected with pcDNA3.1 control plasmid (Figure 3.9, C). Knocking down of HOXD10 in those cells was confirmed by both qPCR and WB procedures (Figure 3.9, D).

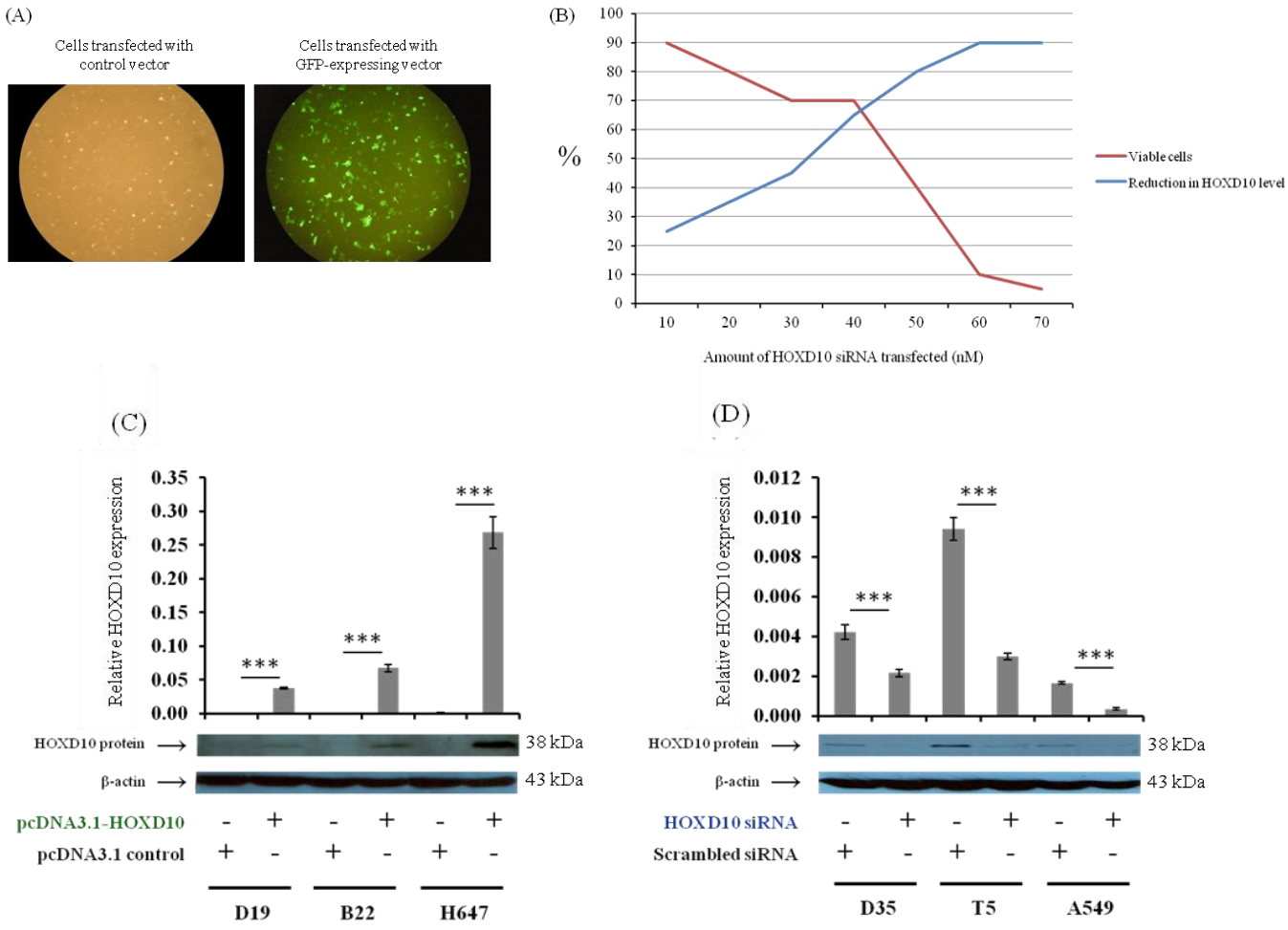


Figure 3.9: Transient manipulation of HOXD10 expression.

(A) GFP protein expression level was used to optimise the transfection procedure to drive the expression of the transfected gene in sixty percent or more of the transfected cells. (B) Percentage of viable cells subjected to HOXD10 silencing was compared to the amount of the introduced HOXD10 siRNA and the level of HOXD10 knocking down. Viability of cells reduced by more than fifty percent when HOXD10 level was knocked down to seventy percent or more of its basal level using 50 nM or more of HOXD10 siRNA. 40 nM HOXD10 siRNA was an appropriate amount enough to knock HOXD10 down by sixty percent of its basal level with maintaining an enough percentage of viable cells. (C) qPCR analysis showed very high expression level of HOXD10 transcripts in cells transfected with pcDNA3.1-HOXD10 for 48 hours compared to cells transfected with pcDNA3.1 control vector (upper panel). At the protein level, the anti-HOXD10 Ab was specific but not sensitive for very low levels of HOXD10 protein as it was barely detected when 100 µg total protein of each cell lysate were analysed using WB protocol (lower panel). (D) High-HOXD10 expressing cells demonstrated a significant decrease in their HOXD10 mRNA transcripts and protein when they were transfected with 40 nM of an anti-HOXD10 siRNA for 48 hours compared to cells transfected with similar amount of a scrambled siRNA control (*** p<0.001, parametric T-test, qPCR data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

3.6 Differential *IGFBP3* and *miR-7* expression levels following HOXD10 manipulation

Although neither *IGFBP3* nor *miR-7* showed a clear correlation between their basal level and HOXD10 basal level in a panel of HNSCC and NSCLC cell lines, it was investigated whether they would be affected by changing HOXD10 levels. In a contrast with expectations, the level of *miR-7* was decreased in two out of three cell types following HOXD10 over-expression (Figure 3.10, A). On the other hand, targeted knockdown of HOXD10 level using HOXD10 siRNA caused a decrease in *miR-7* in only one out of three cell type with no significant effect seen in the remaining two (Figure 3.10, B). Over-expression of HOXD10 resulted in significant increase in the expression level of *IGFBP3*, with a reciprocal effect seen in two high-HOXD10 expressing cell types out of three when HOXD10 was decreased (Figure 3.11).

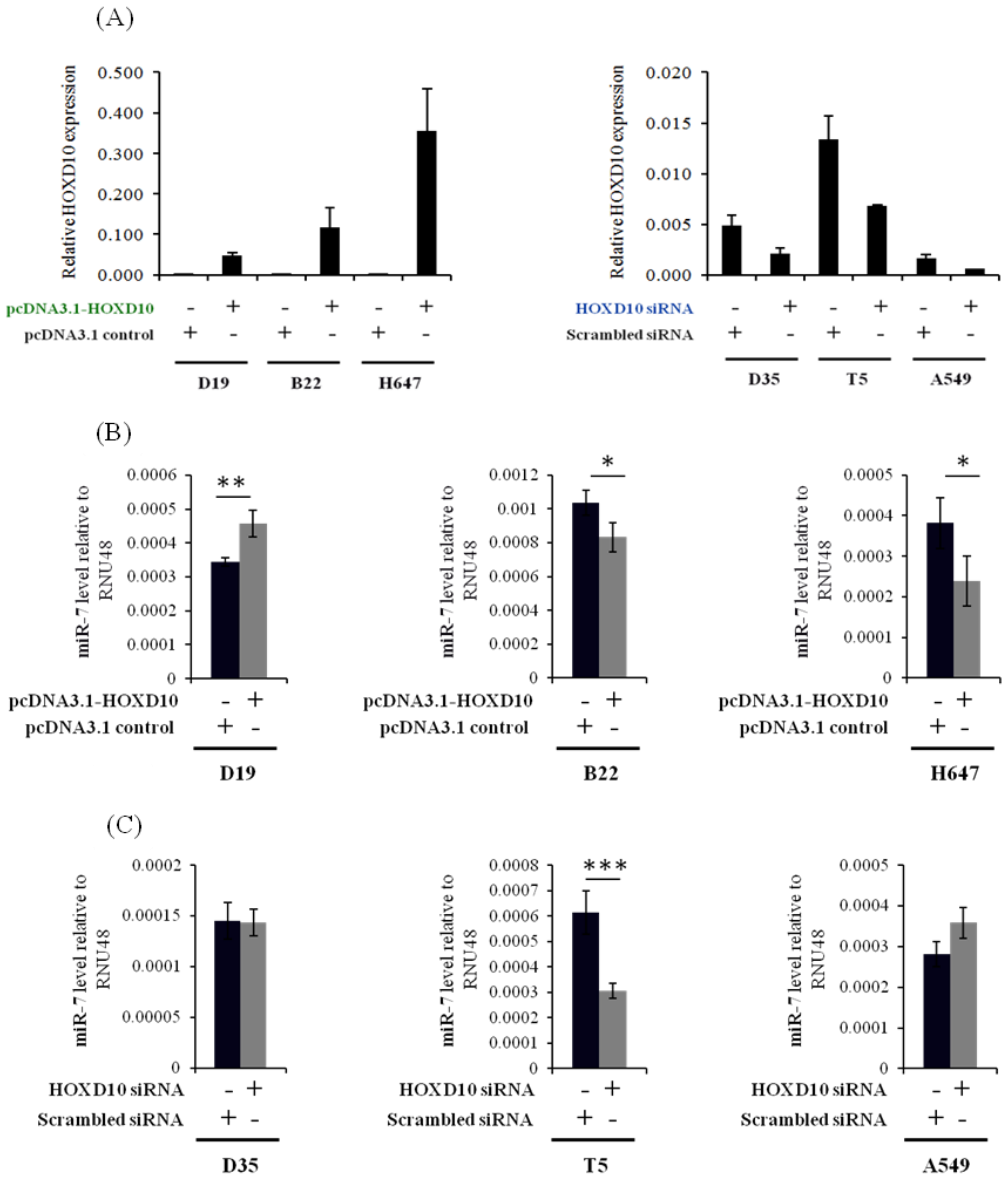


Figure 3.10: The effect of HOXD10 level manipulation on *miR-7* expression.

(A) HOXD10 over-expression and silencing was confirmed by qPCR analysis of HOXD10 mRNA level. Although HOXD10 has significantly affected the level of *miR-7* in most of the examined cell lines, the effect is not consistent neither in the case of HOXD10 over-expression (A) nor in silencing (B) (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

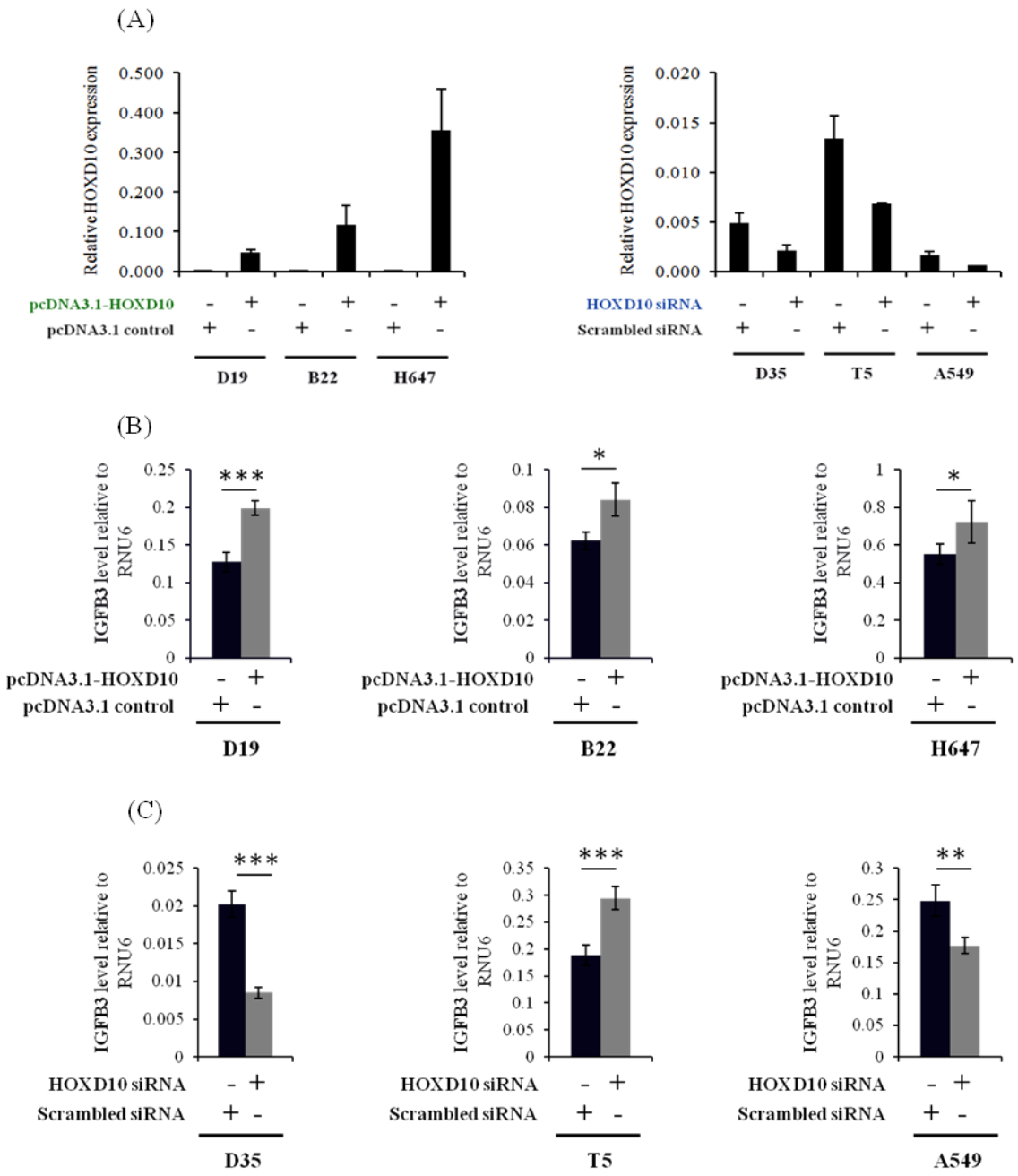


Figure 3.11: The effect of HOXD10 level manipulation on *IGFBP3* expression.

(A) HOXD10 over-expression and silencing was confirmed by qPCR analysis of *HOXD10* mRNA level. (B) Although the basal level of *IGFBP3* has not shown a correlation with the basal level of *HOXD10* previously, introducing high level of HOXD10 into low-HOXD10 expressing cells caused a significant increase in their *IGFBP3* level. (C) Decreasing HOXD10 on high-HOXD10 expressing cells has led to decreasing *IGFBP3* in two cell types out of three (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

3.7 HOXD10 over-expression induces cell proliferation and vice versa

Low-HOXD10 expressing cells, D19, B22 and H647, were tested for their ability to proliferate after introducing high levels of HOXD10 compared to cells transfected with a control plasmid. On the other hand, the cell growth of high-HOXD10 expressing cells, D35, T5 and A549, was assessed following knocking down of their HOXD10 level using HOXD10 siRNA. Forty eight hours after plasmid DNA or siRNA transfection, cells were harvested, counted using Trypan blue stain to exclude dead cells, and then seeded in an equal number for different time points to measure the change in their growth through the MTS assay. For each cell type, the colour intensity developed by the growing transfected cells was statistically compared to control-transfected cells at time point "96 hours" using the parametric student's *T*-test with a *p*-value below 0.05 considered significant. Results show that low-HOXD10 expressing cells responded positively to high levels of HOXD10 by increasing their proliferation, while decreasing HOXD10 level in high-HOXD10-expressing cells showed a reciprocal effect (Figure 3.12). It is worth mentioning that there was no association between the basal levels of HOXD10 and doubling time/normal proliferation of the cell lines under investigation since some low-HOXD10 expressing cells, such as B22 cells, proliferate faster than some high-HOXD10 expressing cells, such as T5 cells. Also, the low-HOXD10 expressing D19 cells and high-HOXD10 expressing D35 cells have similar doubling time.

3.8 HOXD10 over-expression promotes cell migration and vice versa

Over-expression of HOXD10 enhances B22, D19 and H647 cells motility compared to cells transfected with control plasmid (Figure 3.13, A). Knockdown of HOXD10 in high-HOXD10 expressing cells, T5, D35 and A549, significantly impaired their migration capacity compared to control cells with un-changed HOXD10 expression level (Figure 3.13, B).

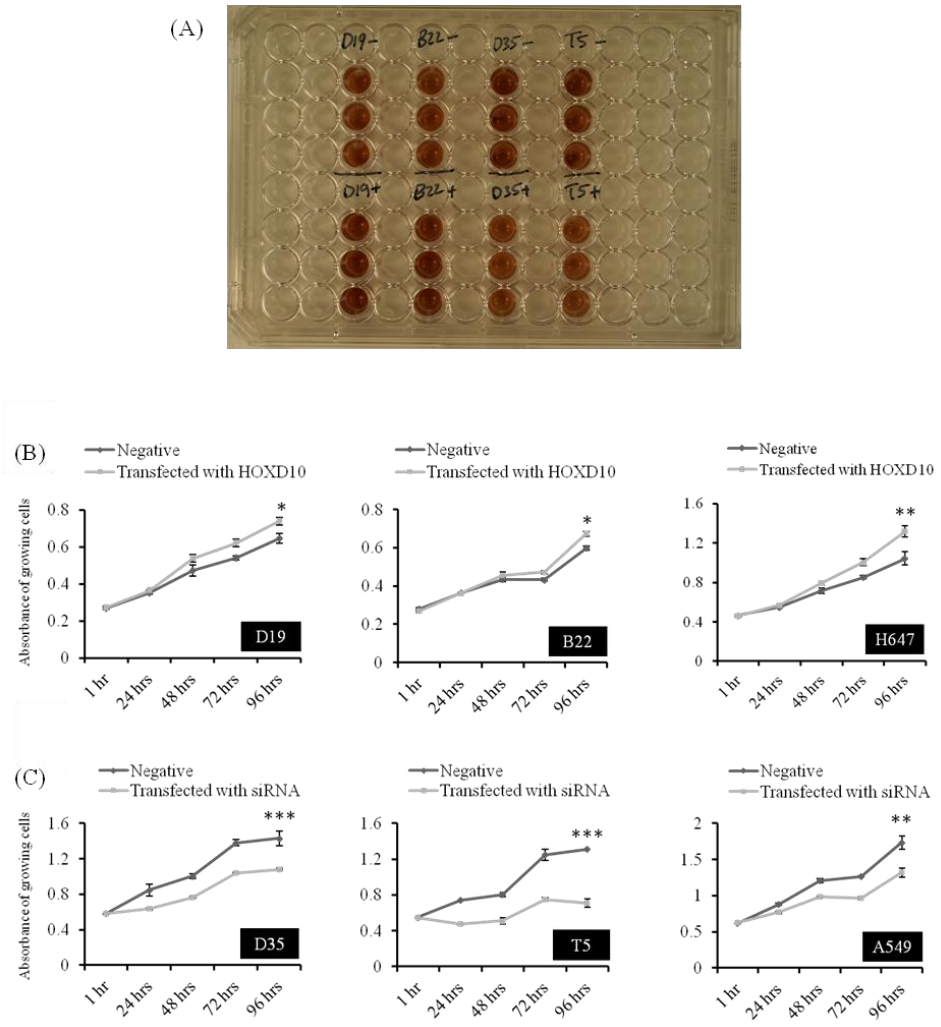


Figure 3.12: The effect of HOXD10 transient over-expression or silencing on cell proliferation.

(A) In a 96-well plate, cells subjected to HOXD10 over-expression (D19+ and B22+) in addition to control cells (D19- and B22-) were assessed for the change in their proliferation using an MTS assay, in which the intensity of the developed colour reflects the number of the growing cells. Same procedure was applied to cells subjected to HOXD10 silencing (T5+ and D35+) in addition to their controls (D35- and T5-). (B) Increasing the level of HOXD10 in D19, B22 and H647 cells caused a statistically significant increase in their proliferation. (C) D35, T5 and A549 cells, which naturally express high level of HOXD10, showed a considerable reduction in their growth following HOXD10 targeted knockdown (*p < 0.01, **p < 0.01, ***p < 0.001, parametric T-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

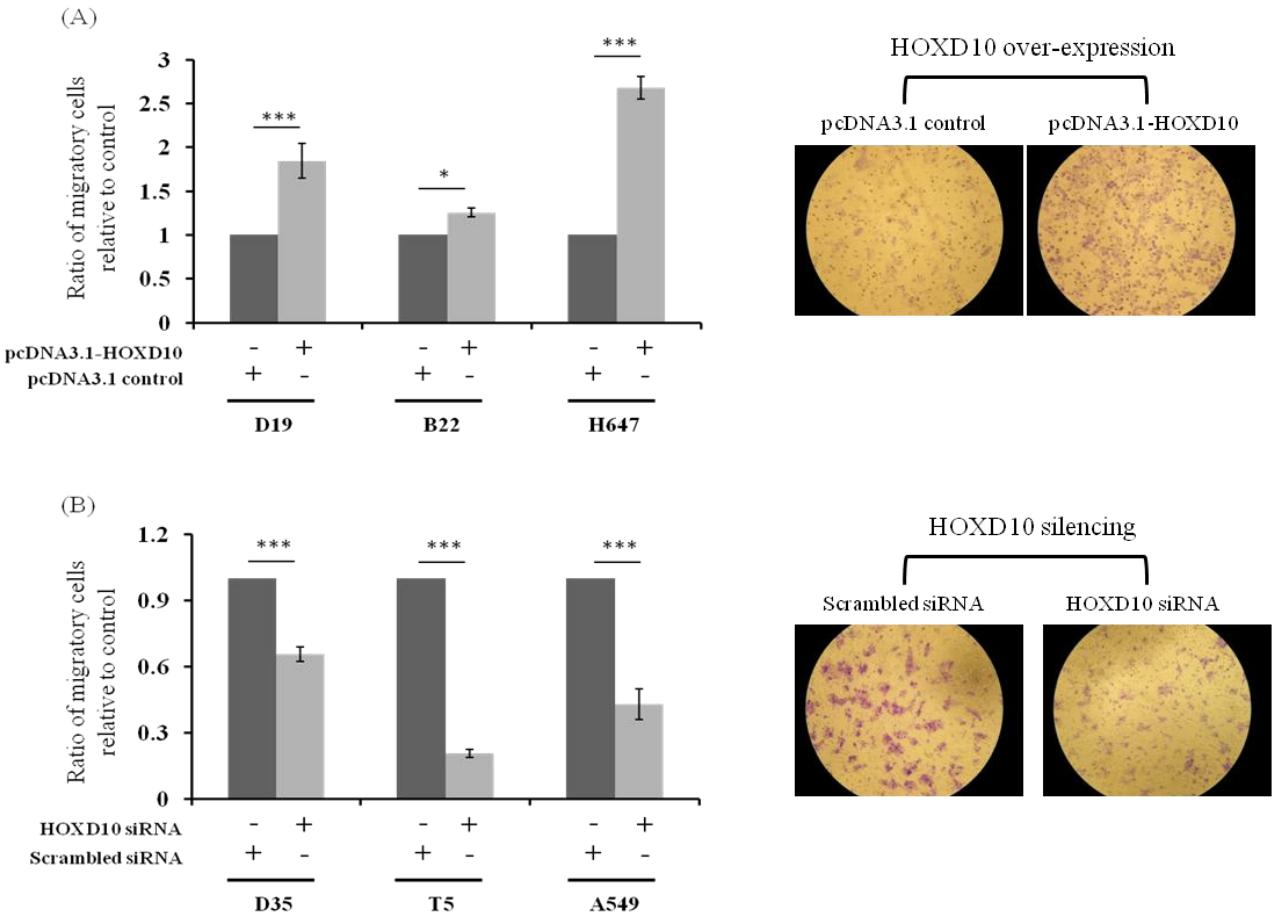


Figure 3.13: The effect of HOXD10 transient over-expression or silencing on cell migration.

(A) The number of cells transfected with HOXD10 demonstrated increased migration compared to cells transfected with the control plasmid as can be seen in the bar graph and the representative image of migrating cells stained with crystal violet. (B) Targeted HOXD10 knockdown for 48 hours reduced the ability of cells to migrate through the porous membrane compared to cells transfected with a scrambled siRNA control (*p < 0.05, *** p < 0.001, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

3.9 HOXD10 over-expression enhances cell adhesion and vice versa

Having showed that changing HOXD10 expression level was able to modulate cell migration, impact of such manipulation on cell adhesion was investigated. Fibronectin is one of the major ECM glycoproteins that interact with the cell surface and other ECM components to mediate cell adhesion and many other cellular and physiological processes (Hörmann, 1982, Pankov and Yamada, 2002, Sottile and Hocking, 2002). Low-HOXD10 expressing cells, D19, B22 and H647, already transfected with pcDNA3.1-HOXD10 or control plasmid, in addition to high-HOXD10 expressing cells, D35, T5 and A549 already transfected with HOXD10 siRNA or control siRNA were seeded on a fibronectin-coated surface. Cells that were able to bind to fibronectin were measured using the MTS colorimetric method. Statistical analysis of the results showed that HOXD10 significantly increases the number of cells capable of binding to fibronectin, while knocking it down negatively affects cell adhesion (Figure 3.14).

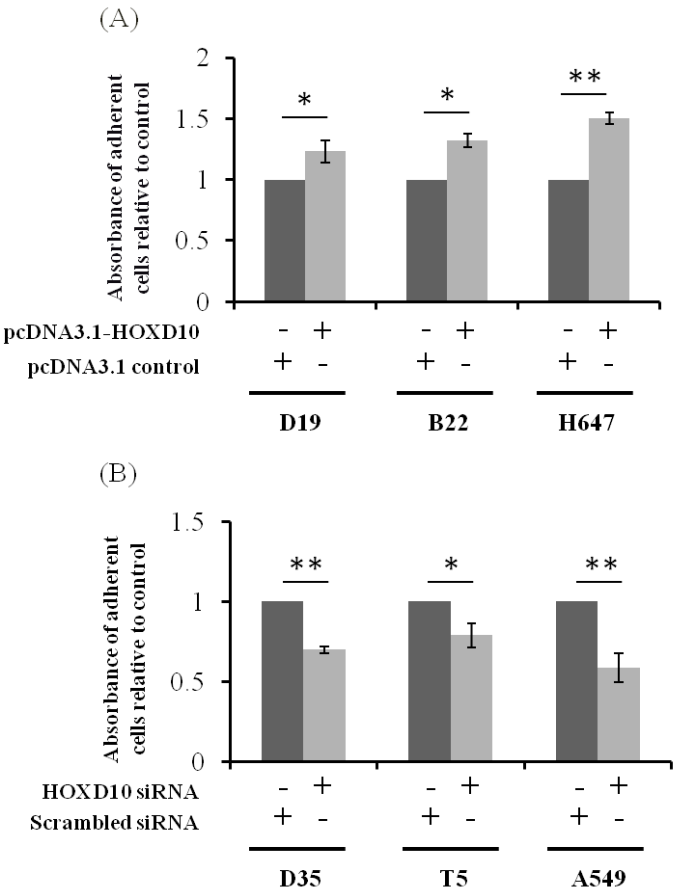


Figure 3.14: The impact of HOXD10 transient over-expression or silencing on cell adhesion.

(A) Over-expression of HOXD10 induced cell adhesion. (B) HOXD10 knockdown impaired cells' capacity to adhere to the surface coated with fibronectin (B) (* $p < 0.05$, ** $p < 0.01$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

3.10 HOXD10 over-expression impairs cell invasion and vice versa

Stably transfected cells were used for testing the effect of HOXD10 manipulation on cell invasion since the results of the transient transfection were not reproducible. B22 and D19 cells were stably transfected with either pcDNA3.1-HOXD10 sequence or pcDNA3.1 control vector, while T5 and D35 cells were stably transfected with either an anti-HOXD10 shRNA or a control shRNA vector (Figure 3.15). It was not surprising to see that stable transfection caused a moderate HOXD10 over-expression or silencing compared to the potent effect caused by transient transfection. Creating stably transfected H647 and A549 cells was not possible because they were highly resistant to the selective G418 antibiotic, even to 1.5 µg/mL G418, thereby, these cells were excluded from this particular study. Stably transfected cells were continuously maintained in a 400 µg/mL G418 antibiotic-containing medium before and during the experiment.

Inducing constant expression of high HOXD10 level in B22 and D19 cells negatively affects their invasion through a Matrigel membrane (Figure 3.16, A), while sustained decrease of HOXD10 endows D35 and T5 cells with an invasive ability *in-vitro* (Figure 3.16, B).

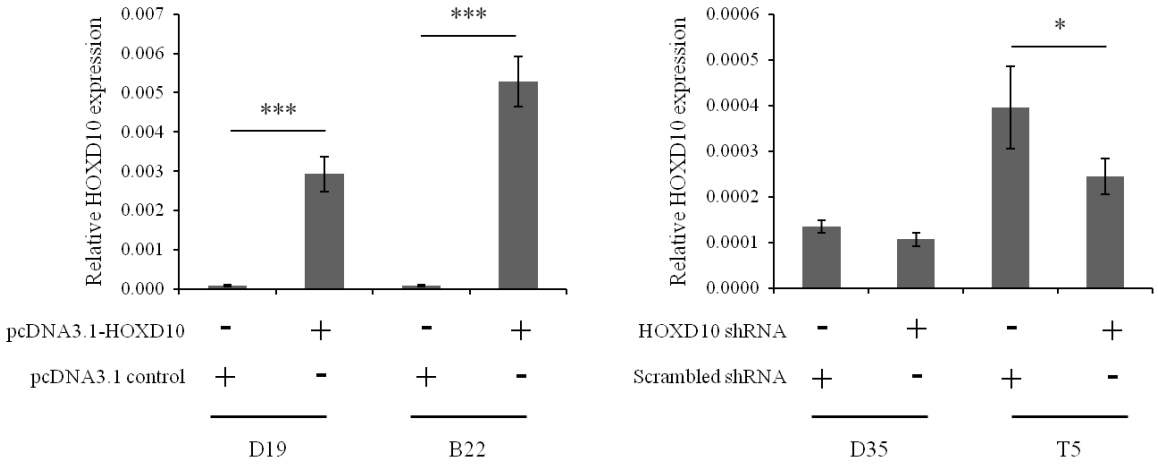


Figure 3.15: Stable manipulation of HOXD10 expression.

Naturally low-HOXD10 expressing cells B22 and D19 were stably transfected with either pcDNA3.1-HOXD10 or pcDNA3.1 control vector, while anti-HOXD10 shRNA or scrambled shRNA control were transfected into two naturally high-HOXD10 expressing cells, D35 and T5. The effect of stable transfection on HOXD10 is moderate compared to the transient transfection. (* $p < 0.05$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

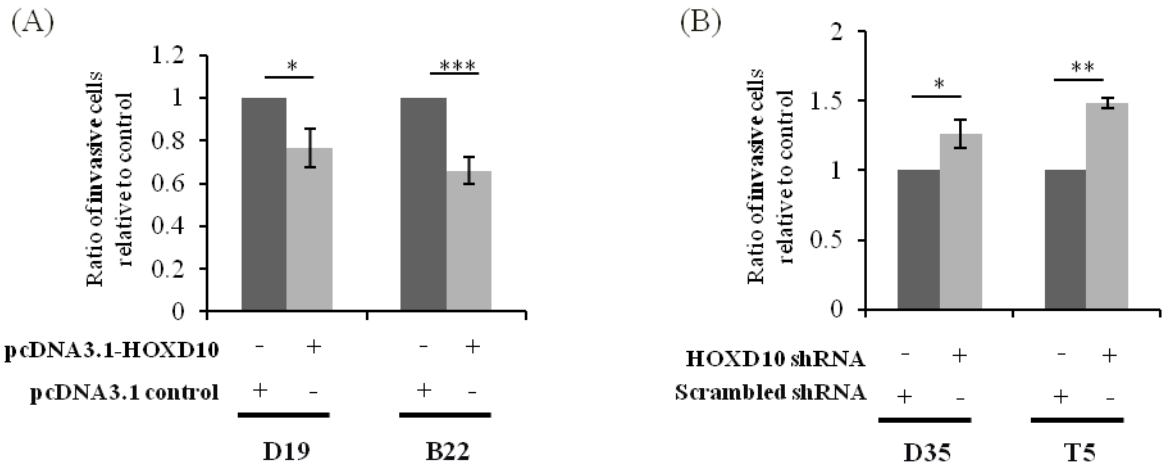


Figure 3.16: The effect of HOXD10 level manipulation on cell invasion.

(A) The number of cells transfected with HOXD10 detected to be lower than the number of cells transfected with the control plasmid. (B) Targeted knockdown of HOXD10 by shRNA led to increasing the ability of cells to invade the Matrigel layer compared with cells transfected with a scrambled siRNA control (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

3.11 Discussion

Although the high expression level of HOXD10 transcripts has previously been observed in head and neck and lung cancer cells, the present study not only confirmed these findings but also explored, for the first time, the differential expression of HOXD10 in pre-malignant and metastatic HNSCC cells. Also, a significant decrease in HOXD10 expression in HNSCC metastases relative to their paired primary tumours was demonstrated. Such findings agree with previous data shown in breast cancer and ovarian cancer that high HOXD10 expression is significantly decreased at the invasive front and during cell metastasis (Ma et al., 2007, Nakayama et al., 2013). Detection of HOXD10 in some OPL cells suggests that the increase in HOXD10 expression could occur early in the development of HNSCC and NSCLC. The variable expression of HOXD10 in pre-malignant cells might also be linked to their degree of tumourgenic progression. Also, in this study, the differential expression of HOXD10 during HNSCC development phases was confirmed at the protein level using human tissue samples. Moreover, an agreement between HOXD10 mRNA and protein level in a panel of cell lines is demonstrated. During the assessment of HOXD10 protein expression, a marked variability in specificity and sensitivity of seven different HOXD10 antibodies were demonstrated. Only two Abs were able to detect HOXD10 specifically in cell lysates and tissue sections. It is not clear why the other HOXD10 Abs failed to work properly in WB and IHC experiments, although some of them were used in some previously published work. Examples of the discrepancies associated with these Abs were detecting the same level of HOXD10 protein in different types of cells with differing mRNA expression, demonstrating no change in HOXD10 protein level when HOXD10 was over-expressed or silenced and detecting a HOXD10 protein of a size that is significantly different from its predicted or previously published size. However, the two HOXD10 Abs identified in this study showed a good quality WB and IHC results. Each antibody was subjected to different quality control measurements, specificity assessments, and required several WB and IHC optimisation steps. For example, the Ab used during WB protocol was validated by its ability to detect protein bands of a correct size and intensities that match qPCR results. In case of IHC, the specificity of HOXD10 antibody was confirmed by detecting HOXD10 protein in tissues known to express it in a high level, such as prostate (Redline et al., 1992). In addition, the pre-adsorption experiment has added more confidence to interpretation of the specificity of this Ab. Using different Abs to detect the same protein in different protocols is common and has been reported in different studies (Pradidarcheep et al., 2008, Schnabel et al., 2013). Multiple reasons might affect the binding of an Ab to its antigen in different protocols. For example, the WB denaturing and electrophoresis steps break the disulfide bonds of proteins tertiary structures and generate bands of concentrated linearised proteins, while in IHC, proteins are denatured to expose the epitopes as much as possible but part of the protein tertiary structure might still intact. Moreover, formalin fixation of tissue sections might cross links antigens and masks the epitopes, which reduces immunoreactivity (Dapson, 2007). Such differences in protein structure and accessibility to the epitope might cause the differential Ab

efficacy because some Abs work well when the epitope is just a sequence of amino acids, while some others bind better if the epitope is in its native conformation (Willingham, 1999).

The lack of association between HOXD10 and miR-10b levels matches the findings of a recently published head and neck cancer study (Severino et al., 2013). Such finding supports the lack of association between HOXD10 and previously reported HOXD10 targets, *miR-7* and *IGFBP3*. Since such interaction was demonstrated in different types of cancer, breast and gastric cancers, this suggests the presence or absence of some modulating molecules in these cancers that might function differently in head & neck and lung cancers. Moreover, as previously discussed in the roles of HOX genes in cancer, same genes might have different expression level and different interacting molecules depending on the type of cancer. However, our results do not completely exclude the interaction between HOXD10 and these molecules in HNSCC and NSCLC, but it can be suggested that such an interaction might be cell type-dependent and the level of this interaction could be affected by other molecules.

Over-expression of HOXD10, transiently or stably, was confirmed by qPCR and western blotting. Similarly, targeted knockdown of HOXD10 was also validated. The knocking down of HOXD10 by sixty percent or more of its basal level was confirmed by qPCR and WB. Targeted knockdown of HOXD10 by ninety percent or more of its basal level in all the three naturally high-HOXD10 cells, D35, T5 and A549, caused a significant reduction in their viability, which suggest a potential role of HOXD10 in the survival or apoptosis resistance in these cells.

Prior to studying the effect of HOXD10 manipulation on cell behaviour, it was necessary to verify that this approach was also capable of affecting the level of HOXD10 downstream targets. Differential expression of reported HOXD10 targets, namely *IGFBP3* and *miR*-7, was confirmed in four out of six cell lines. It was demonstrated that increasing HOXD10 level reflected positively on IGFBP3 expression in three different cell lines, with a reciprocal effect seen in two cell types out of three. In the case of miR-7 expression level, although there was a statistically significant response to HOXD10 manipulation in four cell types out of six, it did not reflect an expected behaviour of a direct target of HOXD10, at least in HNSCC and NSCLC cells. Collectively, these data proved that the introduced HOXD10 was fully functional and able to affect the expression of its downstream targets even if such an effect did not entirely match the previously published reports in other cell types. Moreover, the previously reported downstream targets of HOXD10 in breast and gastric cancers, *miR-7* and *IGFPB3*, respectively, might have different regulating elements in HNSCC and NSCLC cancers.

The results of the proliferation assay reflect a cell growth-promoting function of HOXD10 in HNSCC and NSCLC, which might be one of the reasons behind its high level in primary tumour cells. The effect of HOXD10 on cell migration in HNSCC and NSCLC did not agree with its previously reported effect in some other cancers. For example, in breast cancer, HOXD10 was identified as a migration suppressor factor. However, this might be due to the presence of

some active networks in breast cancer cells different from those of HNSCC and NSCLC cells, especially that the basal level of the previously reported HOXD10 target microRNA in breast cancer, *miR-7* and HOXD10 regulator, miR-10b, have shown no correlation with the basal level of HOXD10 in our study and in a recent study (Severino et al., 2013). The result of the adhesion assay supports the findings of our migration results knowing that cell adhesion is an essential mechanical part of cell motility (Lauffenburger and Horwitz, 1996).

Studying cell invasion was not a straightforward task. A fluctuation in the effect of HOXD10 manipulation on cells' ability to invade Matrigel was demonstrated. After a stepwise checking on the flow of the experiment, it was concluded that HOXD10 manipulation must be permanent to exclude the possibility that cells might need longer time to activate or organise their invasion machinery. Creating cell lines stably transfected with *Neomycin* ORF-containing plasmids was possible in the case of B22, D19, D35 and T5 cells, while it failed in the case of A549 and H647 cells due to their insensitivity to G418 antibiotic. Stable transfection of cells with HOXD10-cloned vector or anti-HOXD10 shRNA, in addition to their control plasmids, was confirmed at the RNA and protein level prior to carrying out the invasion assay. Assessment of cell migration, proliferation and adhesion using stably transfected cells produced similar results to transient transfection. In contrast, sustaining a high or low level of HOXD10 by stable transfection was necessary to study HOXD10 effect on cell invasion, which might suggest that the cellular networks of cell invasion that involve HOXD10 are more complex, and might respond to a cumulative level of HOXD10 or a potential threshold limit. The ability of HOXD10 to impair cell invasion in a response to its sustained over-expression agrees with its anti-invasion role demonstrated previously in breast and ovarian cancers, and recently in some other diseases, such as glioma (Hu et al., 2013) and hepatocellular carcinoma (Li et al., 2013a). This was further supported by demonstrating an increase in the invasion of high-HOXD10 expressing cells even when their HOXD10 level was continuously suppressed by fifty percent or more.

Taken together, these findings indicate a role for HOXD10 in promoting the development of HNSCC and NSCLC tumours to the point of metastasis. At that stage, the anti-invasion role of HOXD10 must be neutralised by suppressing HOXD10 expression by a regulatory mechanism other than promoter methylation. Identifying this regulatory mechanism is beyond the scope of this project. The next task was to identify some of the important molecules that mediate HOXD10 in HNSCC and NSCLC development and progression.

CHAPTER 4: MICROARRAY ANALYSIS

4.1 Introduction

This microarray study was performed in order to identify genes that associate with HOXD10 in HNSCC cells, which might highlight some of the pathways that involve this transcription factor. Two naturally low-HOXD10 expressing cells, HNSCC metastatic B22 cells and D19 OPL cells were selected to evaluate the change in gene expression caused by introducing high level of HOXD10 over-expression. Moreover, the impact of HOXD10 knockdown by siRNA was studied in the normally high-HOXD10 expressing cells, HNSCC primary tumour T5 cells and D35 OPL cells. The experiment was designed to detect firstly the genes affected by HOXD10 over-expression in D19 and B22 cells compared to their control transfected cells, and also the genes affected by HOXD10 silencing in D35 and T5 cells compared to their control transfected cells. Then, the genes affected by HOXD10 over-expression were compared to genes affected by HOXD10 silencing. Genes that were affected by HOXD10 over-expression with a reciprocal effect seen in the case of knocking down were considered for further analysis and validation.

4.2 Pre-analysis Quality Control

Before carrying out the microarray experiment, the quality and purity of the extracted RNA samples were assessed by the NanoDrop spectrophotometer then confirmed by the Agilent 2100 Bioanalyzer (Figure 4.1, A). All the tested samples were of a highly purity and un-degraded RNA. The QC reports of the Agilent Feature Extraction Software showed good quality hybridisation represented by the level and distribution of the detected signals (Figure 4.1, B), and Spike-In Linearity Plot, which reflects the accuracy and reproducibility of the probes signals (Figure 4.1, C).

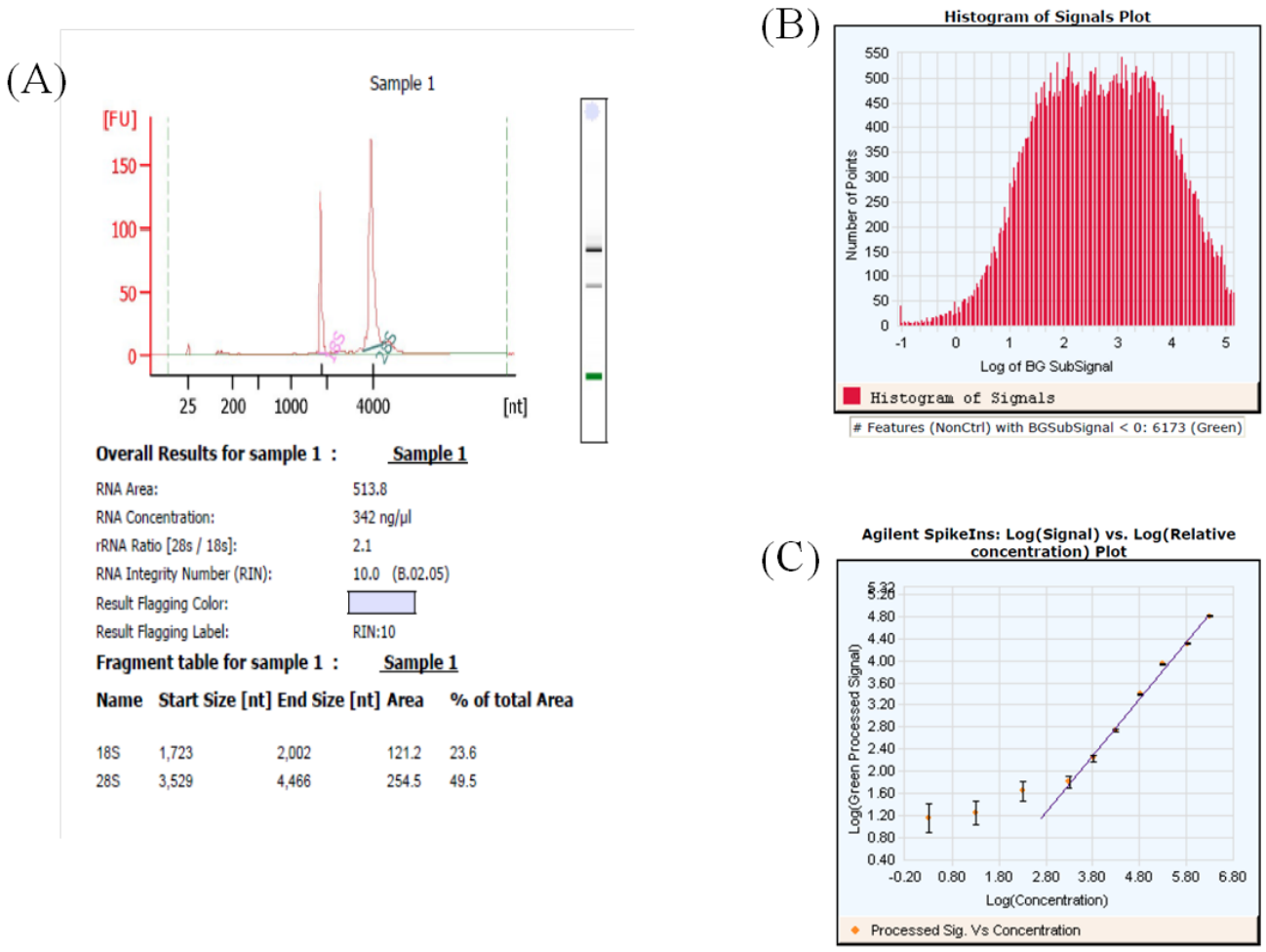


Figure 4.1: Examples of monitoring microarray pre-analytical QC.

(A) An electropherogram showing that total RNA samples were analysed by the Agilent Bioanalyzer 2100 to measure their integrity and purity according to their RNA integrity number (RIN) score and the ratio of the ribosomal RNA (rRNA) 28S to 18S. RIN score of 8-10 and 28S/18S ratio of >1.9 were necessary for an RNA sample to be considered for the microarray procedure. (B) Following subtraction of the background intensity, FE software plots the number of spots against the log of the processed signal to create a Histogram of Signals Plot. A bell-shape curve confirms a normal distribution of the (Cy3) signals intensity across the array. (C) Spike-in Linearity Plot shows the linear increase in the detection level from the lowest detection limit up to the optimal saturation point using the detection levels of the serially diluted spike-in internal positive controls. Small error bar occurs when the signal level is close to the saturation point, while more visible error bars indicate that the signal is close to the background noise. FE extracted data are considered reliable when falling within the signal range, while the signal increases linearly with the concentration of the target.

4.3 Microarray data analysis

Analysis of raw data started with importing "txt" data files generated by the Feature Extraction software into Agilent GeneSpring software followed by automatic computing of the 75th percentile-shift normalisation. The normalised data files were then grouped and assigned to the experiment parameters and conditions as follows: B22 and D19 cells stably transfected with pcDNA3.1-HOXD10 plasmid were assigned to the condition "Over-exp." and named B22 +ve and D19 +ve. B22 and D19 cells stably transfected with pcDNA3.1 control plasmid were assigned to the condition "Low" and named B22 -ve and D19 -ve. The two conditions were under the parameter "HOXD10 over-expression". T5 and D35 cells transiently transfected with HOXD10 siRNA were assigned to the condition "siRNA" and named T5 +ve and D35 +ve. T5 and D35 cells transiently transfected with scrambled siRNA control were assigned to the condition "High" and named T5 -ve and D35 -ve. The two conditions were under the parameter "HOXD10 silencing".

Although each cell type was examined in technical triplicate, it is possible that one of the array triplicates might show some variation due to different factors, which might cause a bias during data processing. Hence, the PCA scores for all the normalised data were computed, grouped, and displayed in a 3-dimensional (3D) scatter (Figure 4.2). Based on visual analysis, any outlier cluster of each triplicate was excluded from the study.

As in any complex statistical analysis, it is rarely the case that a single hypothesis test is always to be preferred for computing the statistical significance. A selection of the most appropriate statistical test might depend on some factors, such as the method of grouping the compared samples. For example, in our study, unpaired *T*-test was used to identify the genes with significant differential expression between two groups by comparing the means of their intensities to compute their *p*-values, where the *p*-value is an estimate of the probability that this gene detected to be differentially expressed by chance. Using this test, gene expression in HOXD10-manipulated D19 and B22 cells (group 1) was compared to genes expression in the control cells (group 2). Then, in a separate analysis, genes expression of HOXD10-manipulated D35 and T5 cells (group 3) was compared to genes expression in D35 and T5 control cells (group 4). The outputs of both analyses were cross matched to identify the common genes affected by HOXD10 manipulation. However, ANOVA test identified genes with differential expression when the groups being compared are defined by one or more conditions. It allowed performing an analysis of variance for all genes across all the groups with comparing differences in the means, which result in a simultaneous evaluation of their differential expression significance. Both statistical methods were tested in this microarray study since it is recommended to apply more than one method for analysing the microarray data (Jeffery et al., 2006).

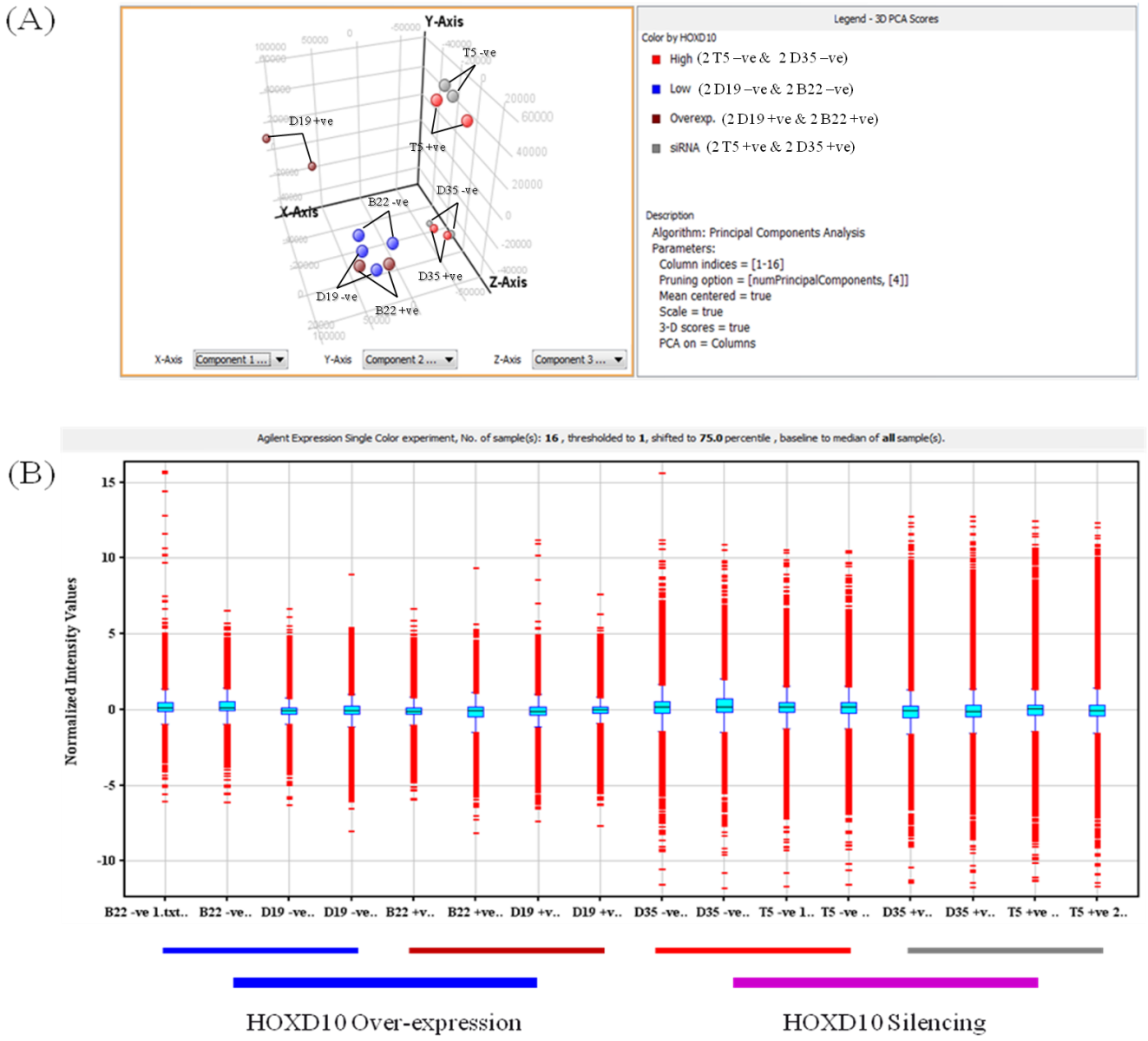


Figure 4.2: A scatter and Box & whisker plot for samples' PCA scores and normalised values, respectively.

Agilent GeneSpring software was used to compute PCA scores to identify the global trends of gene expression patterns for the replicates of each tested cell line, and to present the data in a 3D scatter plot (A). Based on visual analysis, any outlier cluster of each triplicate array was excluded. (B) A distinct parameter was then assigned to the normalised files of each two cell types subjected to comparable transfection procedure. Parameter "HOXD10 over-expression" included B22 and D19 cells transfected with either pcDNA3.1-HOXD10 or pcDNA3.1 control, while parameter "HOXD10 silencing" included T5 and D35 cells transfected with either HOXD10 siRNA or siRNA control.

4.4 Using *T*-test to identify HOXD10-associating genes

Using the straight forward workflow of GeneSpring software, the two conditions of the parameter "Over-expression", Over-exp. and Low, were compared and the significant difference across their lists of genes was statistically evaluated by the moderated unpaired *T*-test with a *p*-value cut-off of 0.05. The resulting list of genes (3255 genes) contains the genes that were up-regulated or down-regulated because of HOXD10 over-expression in D19 and B22 cells according to the used statistical analysis. After this, the same procedure was applied to the conditions, siRNA and High, of the parameter "Silencing". The resulting list of genes (1348 genes) contains the genes that were up-regulated or down-regulated following of HOXD10 silencing in T5 and D35 cells according to the used statistical analysis. Finally, the generated lists of the two parameters were crossed-matched using the GeneSpring's built-in tool "Venn diagram view" (Figure 4.3). According to the above analysis, only 81 genes were detected in the two compared lists. However, GO analysis has shown no significant enrichment.

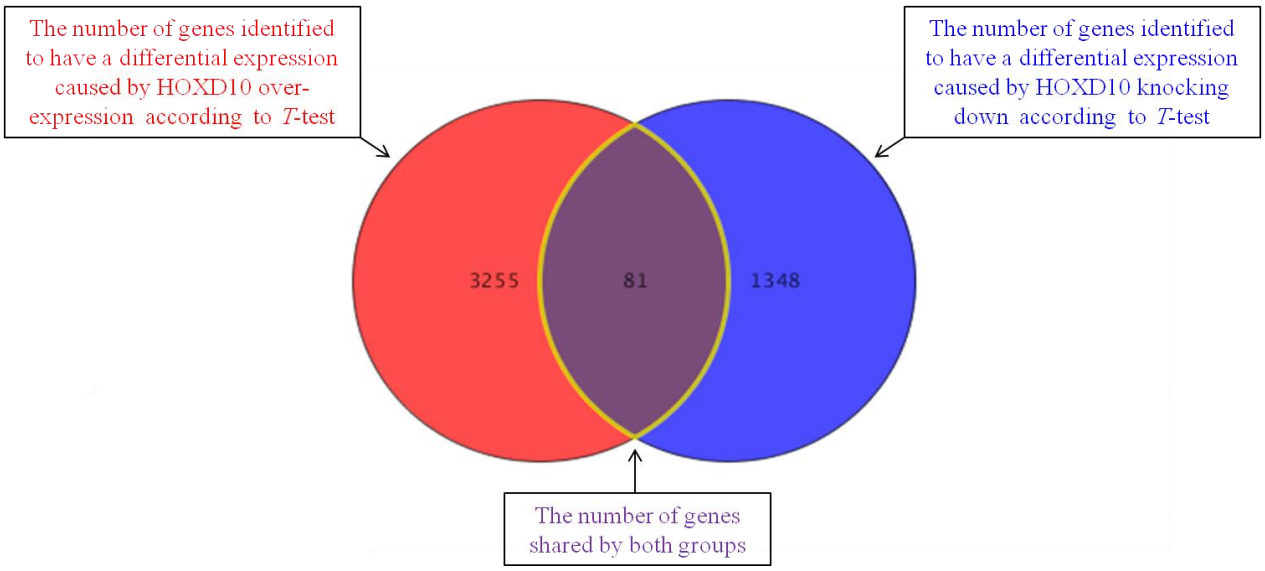


Figure 4.3: Venn diagram of the array analysis using *T*-test.

The four different cell types, D19, B22, D35 and T5, were assigned and analysed according to the type of HOXD10 manipulation they were subjected to. In the "Over-expression" group, expressed genes of D19 and B22 cells stably transfected with pcDNA3.1-HOXD10 were compared statistically to those of D19 and B22 stably transfected with a control plasmid using *T*-test, which resulted in identifying a differential change in the expression of 3255 genes with a 0.05 *p*-value cut-off. Differentially expressed genes in D35 and T5 cells transiently transfected with HOXD10 siRNA were compared to D35 and T5 cells transiently transfected with a control siRNA, which resulted in identifying a differential change in the expression of 1348 genes according to *T*-test with a 0.05 *p*-value cut-off. Eighty one genes were identified to be shared between the "Over-expression" and "Silencing" groups.

4.5 Filtering the list of HOXD10-associating genes identified by *T*-test

The 81 genes were exported with their annotations to an Excel file, and then subjected to some filtering criteria to exclude genes that show no association with HOXD10. Firstly, genes that show no reciprocal effect caused by HOXD10 over-expression versus silencing were excluded. Next, the promoter regions of the remaining genes were detected as described in section (2.2.6.2.1), and screened for HOXD10 binding sites using PROMO online algorithm (Messeguer et al., 2002) (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). This resulted in the nomination of 19 genes for further analysis (Table 4.1).

4.6 qPCR analysis to validate the candidate genes that correlate with HOXD10 expression level

To confirm the previous filtering procedure, qPCR analysis was used to assess the level of the 19 genes in the transfected cells and in a panel of HNSCC and NSCLC cell lines in addition to two different normal oral keratinocytes. For considering a gene to have an acceptable association with HOXD10 it was essential to see both a reciprocal effect on its expression level, in the case of HOXD10 over-expression versus silencing, and an acceptable match between its basal level and HOXD10 basal level in different HNSCC and NSCLC cell types. It is worth mentioning that all the primers used for the qPCR analysis to validate the results of the microarray study were designed for a SYBR® Green based qPCR reaction. Therefore, the specificity of each set of primers designed to detect a gene of interest was assessed by both a dissociation curve analysis and a one percent agarose gel electrophoresis. Any set of primers that failed to produce a single major peak or DNA bands that do not agree with qPCR results was excluded and substituted with a new set of primers (Figure 4.4). Nevertheless, an agarose gel electrophoresis cannot confirm that the band visualised is definitely related to the amplified target mRNA even if the size detected agrees with the expected amplicon size. Moreover, the one percent agarose gel and the 1 kb DNA ladder were not always suitable for detecting amplicons of small sizes. Sequencing the amplified DNA product after extracting it from the gel would validate specific target detection. However, this step was not applied in this study.

qPCR analysis confirmed that that change in the expression of most of the 19 genes (13 out of 19) caused by HOXD10 manipulation matched the microarray analysis, exemplified by USP6 (Figure 4.5). However, only one gene, CYP2J2 (Cytochrome P450, family 2, subfamily J, polypeptide 2), showed a correlation between its basal level and HOXD10 basal level in a panel of cell lines (Figure 4.6). Two putative HOXD10-binding sites were detected in CYP2J2 promoter and to validate them, DLR assay was used. However, it was more reasonable to re-analyse the microarray data using different statistical test to nominate more targets of HOXD10 before proceeding to the DLR assay.

Probe name	Accession number	Gene symbol	Effect of HOXD10 over-expression	Fold change	*p* - value (*T*-test)	Effect of HOXD10 silencing	Fold change	*p* - value (*T*-test)
A_24_P101704	NM_004560	ROR2	+	2.2	**1.92E-05**	-	1.6	3.42E-03
A_33_P3421209	NM_020066	FMN2	+	5.4	**8.37E-04**	-	3.6	8.25E-04
A_33_P3352970	NM_001570	IRAK2	-	1.4	**1.37E-03**	+	1.5	2.21E-02
A_23_P57417	NM_005940	MMP11	+	1.6	**1.40E-03**	-	1.4	3.51E-02
A_24_P245298	NM_003809	TNFSF12	+	1.4	**2.55E-03**	-	1.5	4.80E-02
A_23_P64129	NM_006410	HTATIP2	+	1.4	**3.07E-03**	-	1.4	5.45E-03
A_23_P36562	NM_002205	ITGA5	+	2.5	**1.08E-02**	-	1.8	4.59E-02
A_23_P39955	NM_001615	ACTG2	+	1.8	**1.20E-02**	-	2.6	1.60E-04
A_23_P343935	NM_022051	EGLN1	+	1.5	**1.33E-02**	-	1.4	4.48E-02
A_33_P3366195	NM_014862	ARNT2	+	1.9	**1.46E-02**	-	1.3	4.36E-02
A_24_P376391	NM_015103	PLXND1	+	1.4	**1.85E-02**	-	3.2	5.45E-05
A_33_P3374443	NM_024003	L1CAM	+	1.5	**1.90E-02**	-	1.9	7.91E-03
A_33_P3221303	NM_016602	CCR10	+	1.7	**2.08E-02**	-	2.2	3.94E-02
A_33_P3242663	NM_024749	VASH2	+	4.7	**2.72E-02**	-	2.3	2.63E-02
A_24_P238040	NM_004505	USP6	+	2.7	**3.11E-02**	-	2.3	1.28E-02
A_24_P219785	NM_005184	CALM3	+	1.2	**4.31E-02**	-	1.5	1.80E-02
A_23_P103486	NM_000775	CYP2J2	+	1.6	**4.70E-02**	-	1.5	4.16E-02
A_33_P3395369	NM_001634	AMD1	-	1.1	**4.84E-02**	+	2.1	2.55E-02

Table 4.1: The 19 associating genes with HOXD10 identified by *T*-test.

Genes affected by HOXD10 manipulation in B22, D19, D35 and T5 cells were analysed by GeneSpring software using the *T*-test with a significance cut-off of less than 0.05. The list of genes was filtered to include the genes that were affected by HOXD10 over-expression with a reciprocal effect caused by HOXD10 silencing. The list was filtered again to include genes which have putative binding sites for HOXD10 in their promoter region. The resulting 19 genes are listed above.

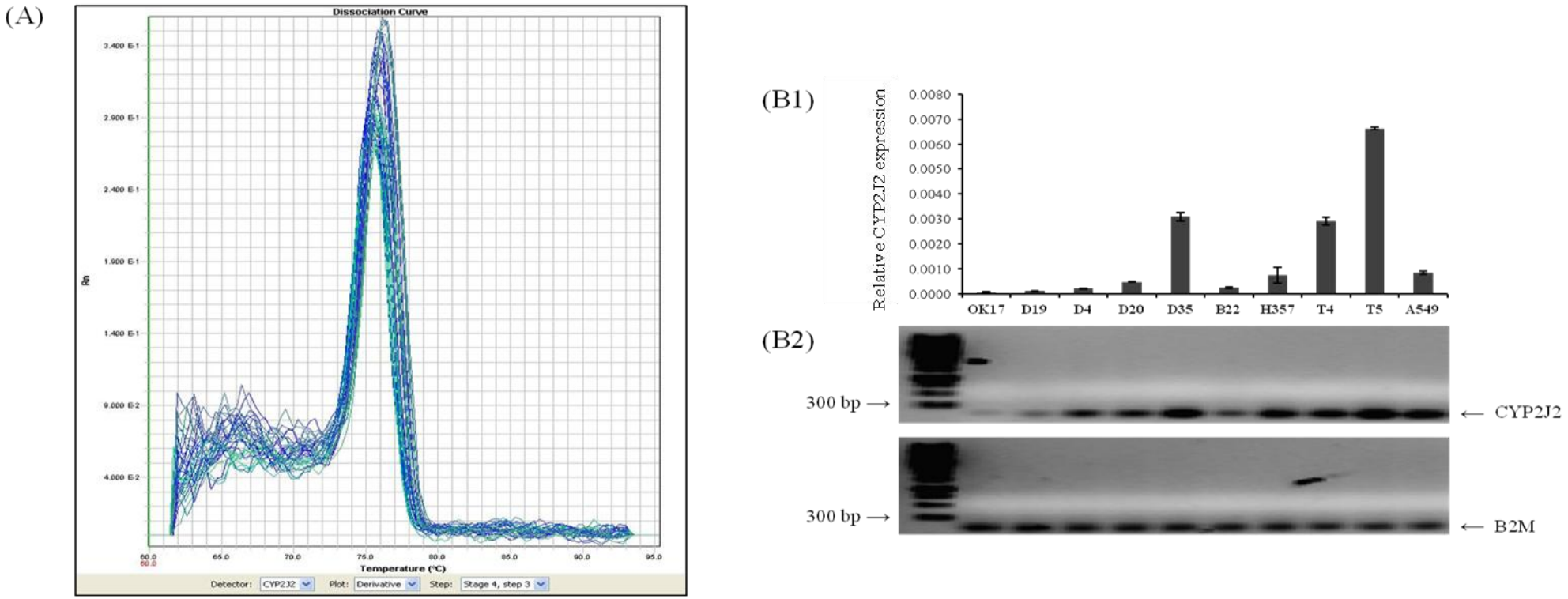


Figure 4.4: A dissociation curve and an agarose gel electrophoresis following a SYBR® Green qPCR run.

(A) The demonstrated single peak reflects a specific CYP2J2 amplicon amplification. After analysing CYP2J2 by qPCR (B1), an agarose gel electrophoresis was performed for both the target gene (CYP2J2) and the house-keeping control gene (B2M) to detect the amplified DNA. The in-house designed primers to detect CYP2J2 were expected to generate an amplicon of 85 bp, while 81 bp is the size of the amplicon should be generated by the commercial B2M primers. The result of the electrophoresis shows that both sets of primers amplified products of correct sizes (B2). The intensity of CYP2J2 bands matched qPCR result, while the high similarity among B2M bands reflects a reliable reference gene for qPCR analysis.

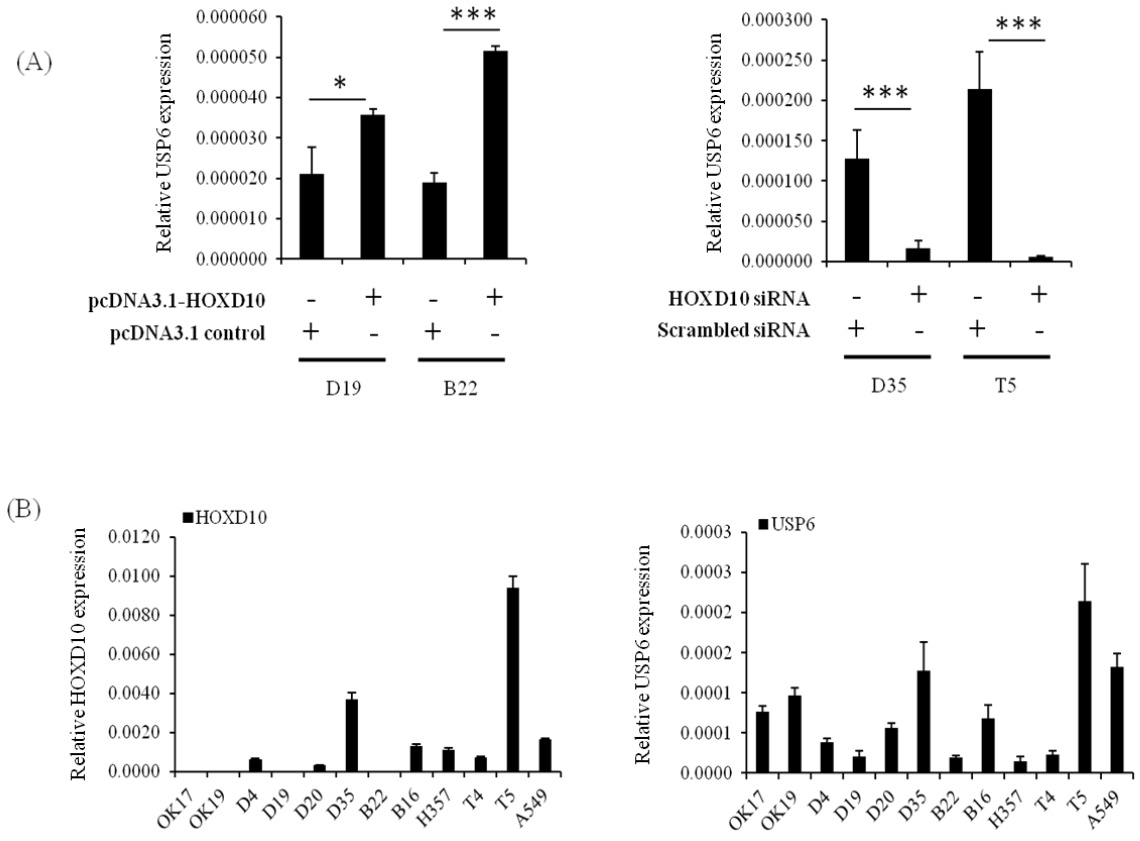


Figure 4.5: HOXD10 basal and manipulated expression level versus USP6 level.

(A) The results of qPCR analysis showed that USP6 level positively correlates with HOXD10 level since it increases with HOXD10 increase, and vice versa. (B) The basal level of USP6 does not completely match the level of HOXD10 in different types of cells (* $p < 0.05$, *** $p < 0.001$ parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

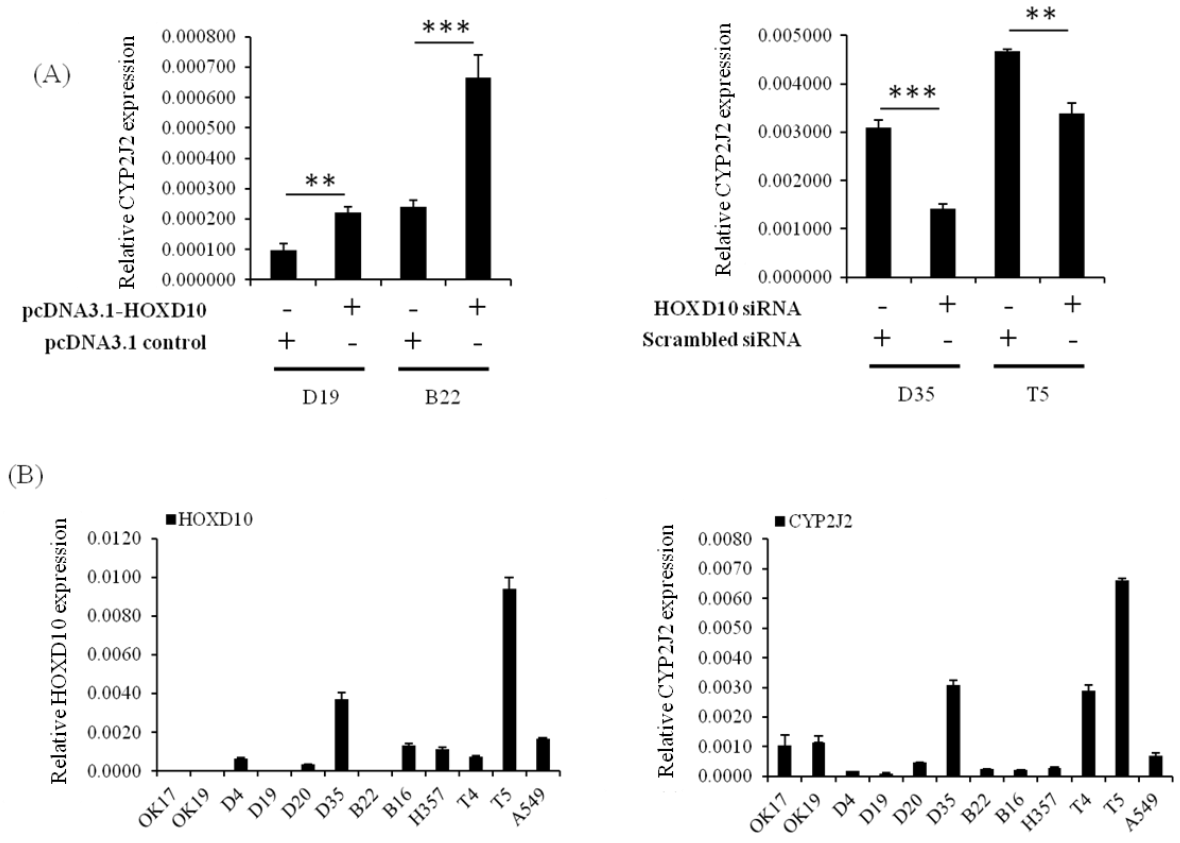


Figure 4.6: HOXD10 basal and manipulated expression level versus CYP2J2 level.

(A) The results of qPCR analysis showed that CYP2J2 level positively correlates with HOXD10 level since it increases with HOXD10 increase, and vice versa. (B) The basal level of CYP2J2 highly matches the level of HOXD10 in different types of cells (** p < 0.01, *** p < 0.001, parametric T-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

4.7 Using ANOVA test to identify HOXD10-associating genes

Moving from the first analysis and the limited number of genes that showed an association with HOXD10 expression, it was necessary to use a more reliable statistical analysis procedure and apply more strict rules to exclude false positive or negative output. Thus, instead of comparing the genes lists as in two groups (Over-expression vs. Silencing) using *T*-test, the more robust ANOVA statistical tool was used, which allows evaluating the significance in the change of the probe's intensity across all the four conditions – Over-exp, Low, siRNA and High – while at the same time retaining the comparison between (Over-exp. and Low) vs. (siRNA and High) parameters. One-way ANOVA with unequal-variance (Welch) test was the method of choice, in which the probe's intensity was considered as a "one variable" during computing the change across all the probe sets. To adjust the *p*-values in order to minimise false positives or negatives, multiple testing correction technique was used. Among the available multiple testing correction options, Benjamini and Hochberg False Discovery Rate (FDR) method was chosen because it is the least stringent and provides a good balance between discovery of statistically significant genes and limitation of false positive occurrences (Loots et al., 2006).

The above analysis method resulted in identifying 9,167 differentially expressed genes out of 42,545 detected genes (Figure 4.7). The listed genes were filtered by:

1) Reciprocal HOXD10-manipulation effect

HOXD10-associated genes level is supposed to have a negative or positive correlation with the manipulation of HOXD10. In other words, if HOXD10 over-expression increased or decreased an associating gene, HOXD10 silencing should cause a decrease or increase, respectively. Therefore, genes that showed no reciprocal effect in a response to HOXD10 manipulation were excluded, which left 3,216 genes differential expressed with HOXD10.

2) Twofold change

The genes list was shortened again by excluding all genes with less than two-fold change in their expression, which brought the number down to 414 genes (Appendix 9).

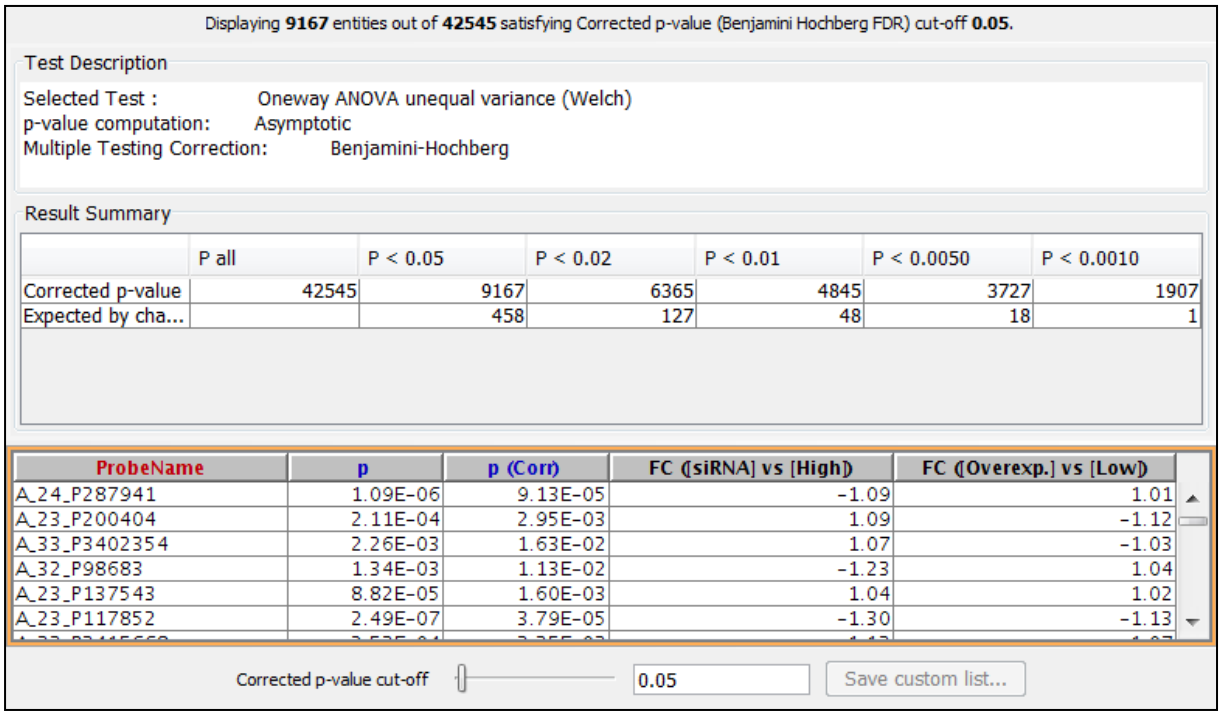


Figure 4.7: A snapshot for the gene expression analysis output using the ANOVA statistical method.

The differences in gene expression between the groups – [(siRNA vs. High) vs. (Over-exp. vs. Low)]* – were identified using the unequal variance ANOVA (Welch) statistical test with a cut-off *p*-value of 0.05. False Discovery Rate (FDR)-corrected *p*-value was computed using the Benjamini-Hochberg procedure for multiple testing correction. In total, 9167 genes were detected to be significantly differentially expressed.

* Key:
siRNA: D35 and T5 cells transiently transfected with HOXD10 siRNA.
High: D35 and T5 cells transiently transfected with scrambled siRNA control.
Over-expression: D19 and B22 cells stably transfected with pcDNA3.1-HOXD10.
Low: D19 and B22 cells stably transfected with pcDNA3.1 control.

4.8 Gene Ontology and pathway analysis

GO analysis is a good method for grouping related genes according to their molecular function, biological process, or cellular component. The 414 HOXD10-associating genes deemed significant using the ANOVA statistical method were uploaded into the GeneRanker tool of the Genomatix software suite for functional clustering with an emphasis on the biological processes. A number of significantly enriched GO categories were identified (complete list is available on: http://www.genomatix.de/cgi-bin//GeneRanker/gene_ranker_web_gui.pl?s=80a3f65115a178c0239f569315c23021;TASK=gene_ranker;show_html_result=1;result_id=882). The main goal of this approach was to explore whether the GO analysis would match to HOXD10-associated phenotypes. In addition to the top 20 GO terms (Table 4.2), the analysis showed a strong association between many of the 414 genes and development, cell proliferation and migration-associated GO terms. This maps well to both the effects demonstrated when HOXD10 level was manipulated and the known role of HOXD10 in development and morphogenesis.

The GePS pathway analysis tool was used to identify the canonical pathways and biological terms of genes of interest. However, due to the large amount of data this tool can generate, an analysis strategy was applied to obtain the most significant results. Genomatix software suite contains a database of gene–disease associations, which is continuously updated through PubMed abstracts and their corresponding medical subject headings (MeSH). Therefore, an investigator can use the "Diseases (MeSH)" annotation to filter the input genes and retain those that are commonly deregulated in pathologies of interest. Sixty five genes out of the 414 genes were found by this method to be shared between NSCLC and HNSCC (Appendix 10). These 65 genes were then re-analysed using the GePS tool but with assigning the "Biological Processes" for pathway mapping. Interestingly, the two most statistically significant pathways identified were associated with cell proliferation (Figure 4.8). Figure 4.9 shows the interacting genes mapped to these two pathways (complete pathways list is available on: http://www.genomatix.de/cgi-bin//GeneRanker/start_pathway_viewer.pl?s=80a3f65115a178c0239f569315c23021;task_name=GePS;result_id=2882;result_name=GePS_9;analysis_date=2013-12-07%2001%3A47%3A28;input_taxon_id=9606;target_taxon_id=;load_result=1).

Some other web-based GO analysis tools, such as DAVID (Dennis et al., 2003) and GOEAST (Zheng and Wang, 2008), have produced highly similar results. It is worth mentioning that 49 genes from the whole gene list have been previously linked to EMT mechanism (Appendix 11) (Jechlinger et al., 2003, Taube et al., 2010). Examples of these genes are JAG2 and OCLN (Figure 4.10).

GO Term	GO-Term id	*p*-value	# input Genes	# Genes (total)
regulation of developmental process	GO:0050793	2.56E-06	55	1237
cell differentiation	GO:0030154	4.11E-06	89	2411
cellular developmental process	GO:0048869	7.92E-06	92	2558
regulation of multicellular organismal development	GO:2000026	1.36E-05	44	959
neuron fate commitment	GO:0048663	1.62E-05	8	49
multicellular organismal development	GO:0007275	2.42E-05	120	3681
anatomical structure morphogenesis	GO:0009653	2.69E-05	70	1850
developmental process	GO:0032502	4.00E-05	130	4109
epithelial cell proliferation	GO:0050673	4.73E-05	16	215
anatomical structure development	GO:0048856	4.79E-05	116	3580
multicellular organismal process	GO:0032501	5.01E-05	157	5207
regulation of cell differentiation	GO:0045595	5.49E-05	39	859
tissue morphogenesis	GO:0048729	6.78E-05	23	399
hemopoiesis	GO:0030097	7.40E-05	25	456
regulation of epithelial cell differentiation	GO:0030856	8.27E-05	8	61
response to steroid hormone stimulus	GO:0048545	8.98E-05	19	301
immune system development	GO:0002520	9.26E-05	27	519
system development	GO:0048731	9.61E-05	102	3100
regulation of epithelial cell proliferation	GO:0050678	1.04E-04	14	183
cell fate specification	GO:0001708	1.04E-04	8	63

Table 4.2: Top 20 GO terms ranked by their *p*-values.

The 414 HOXD10-associating genes deemed significant by ANOVA statistical method were uploaded into the GeneRanker tool of Genomatix software suit for functional clustering with an emphasis on the biological processes. TOP 20 terms by significance are displayed. The analysis showed a strong association between many of the 414 genes and development, cell proliferation and migration-associated GO terms.

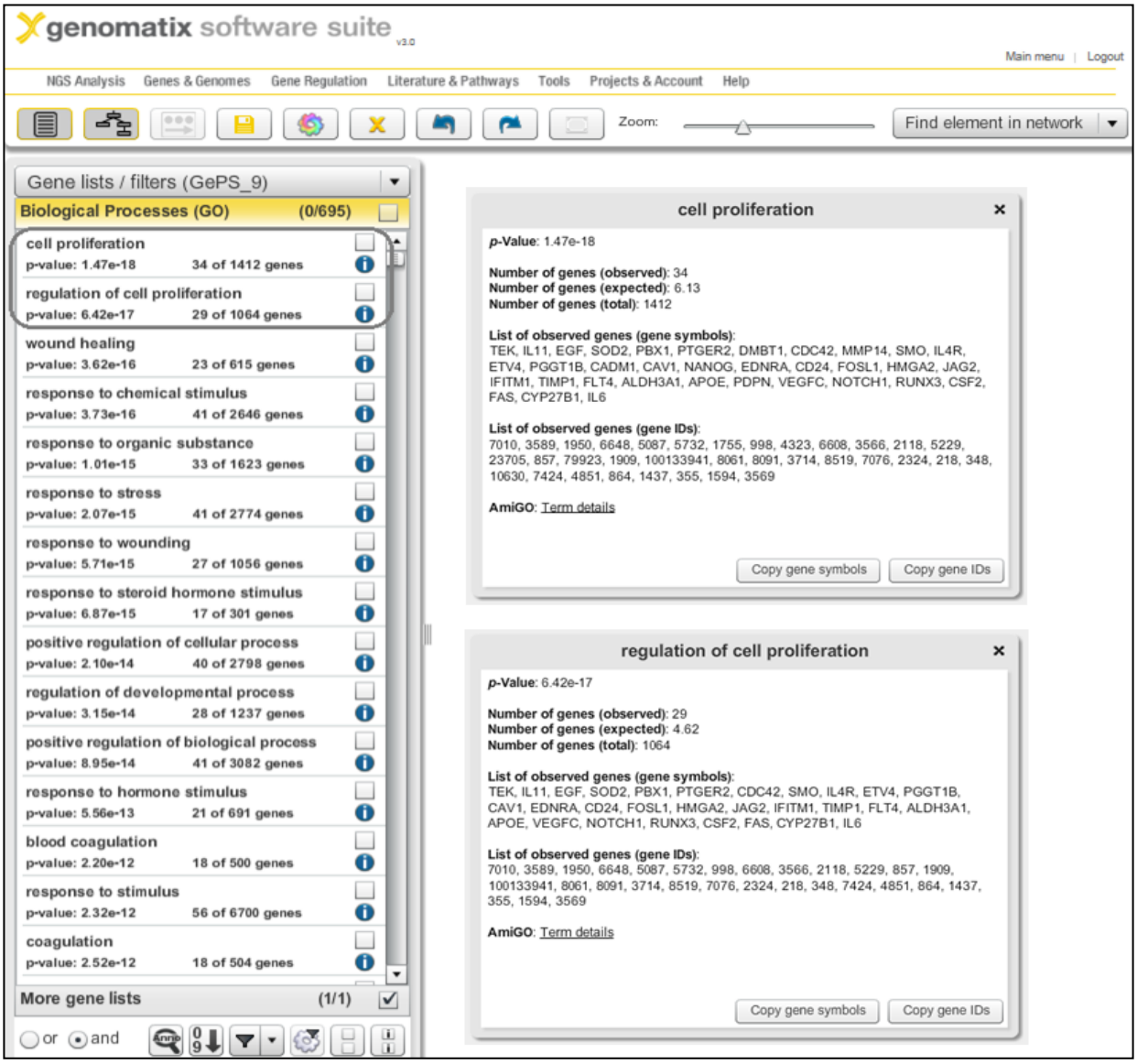


Figure 4.8: Pathway analysis using GePS analysis tool.

GePS pathway analysis tool with "Diseases (MeSH)" annotation was used to identify genes that are commonly deregulated in both NSCLC and HNC. Sixty five genes out of the 414 were analysed using the same tool but with "Biological Processes" for pathway mapping. Thirty four and 29 genes out of the 65 genes (listed above) were mapped to two pathways associated with cell proliferation. Moreover, these two pathways were ranked the two most significant according to their *p*-values.

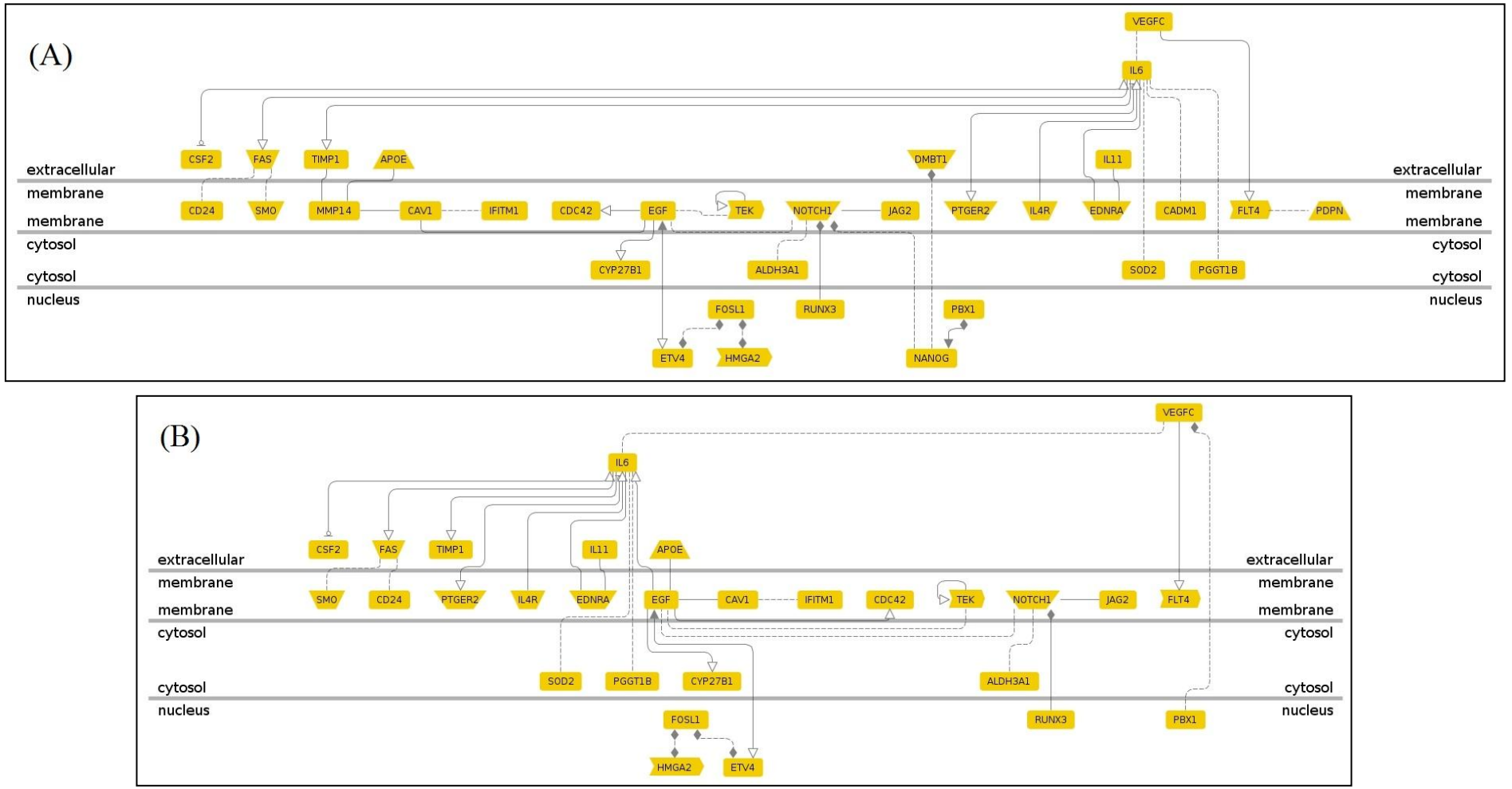


Figure 4.9: Two cell proliferation pathways identified by GO analysis.

GePS analysis tool with "Diseases (MeSH)" annotation was used to identify genes that are commonly dysregulated in both NSCLC and HNC among the genes affected by HOXD10 manipulation. Sixty five genes were found to be shared between these two cancers and analysed using GePS analysis tool but with "Biological Processes" for pathway mapping. Out of the analysed 65 genes, 34 and 29 genes were mapped to two different proliferation-associated pathways (A) and (B), respectively. (*Photos were exported from the Genomatix software suit after completing the analysis*).

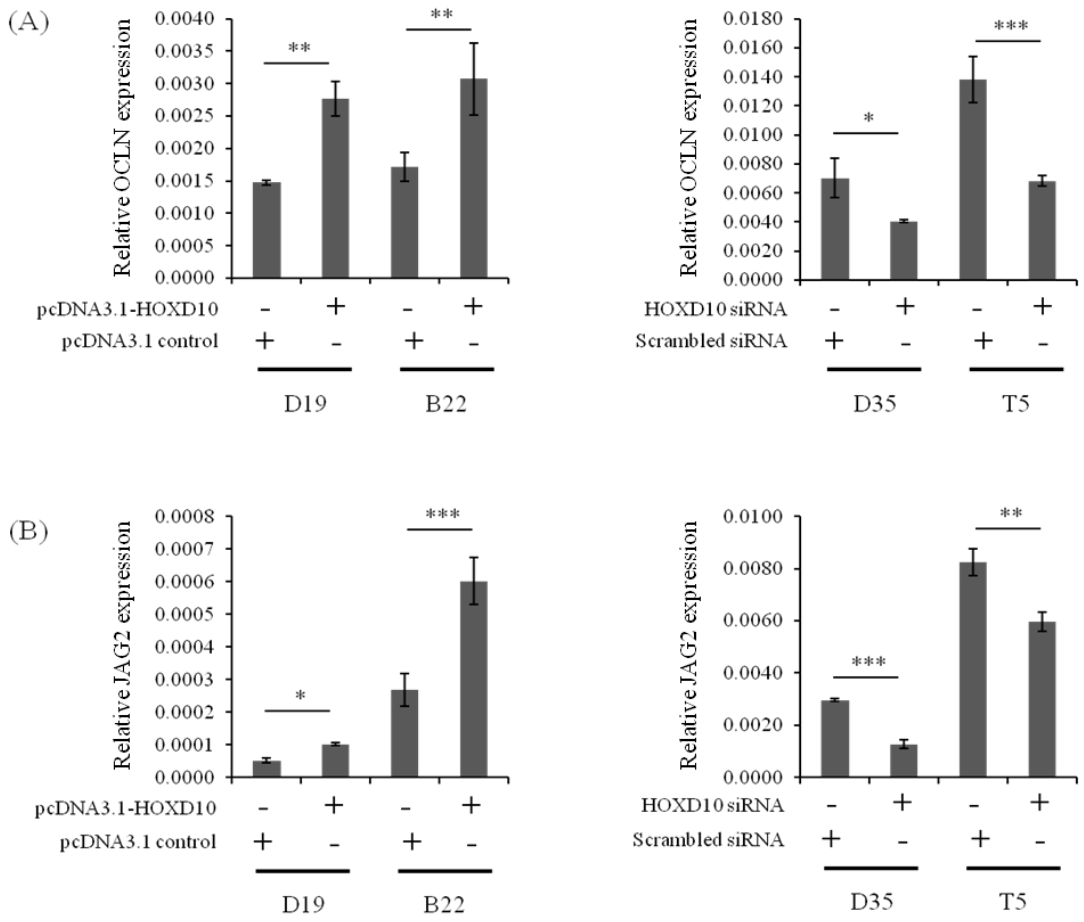


Figure 4.10: OCLN and JAG2 genes associated with EMT are affected by HOXD10 manipulation.

OCLN (A) and JAG2 (B) genes are examples of EMT-related genes that are positively regulated by HOXD10 according to qPCR analysis. Both genes are induced by HOXD10 over-expression and reduced when HOXD10 is knocked down (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

4.9 Validation of HOXD10-associating genes identified by ANOVA test

Although GO analysis could help narrow down the number of HOXD10-associating genes, the frequently identified phenotype by pathway analysis, which is cell proliferation, was the general finding that was relied on to identify the most important HOXD10 associating genes. Through literature mining, the 414 genes were filtered according to their roles in cell proliferation. At the same time, these genes were screened using *in-silico* analysis for HOXD10-binding sites in their promoters. Only 48 genes met these conditions and were considered for further analysis (Table 4.3, Figure 4.3). qPCR analysis was used to validate the differential expression of the 48 genes in the transfected cells and in a panel of HNSCC and NSCLC cells.

Most of the 48 genes (39 out 48) matched the microarray analysis and showed a differential expression caused by HOXD10 manipulation. Genes that show a correlation between their basal expression level and HOXD10 basal expression level in a panel of different types of cells were considered “HOXD10-correlating genes”. Eight genes only were of a high negative or positive correlation (Figure 4.12, Table 4.4). An example of genes that showed weak correlation on the basal level is GBP2 gene (Figure 4.13). On the other hand, the transcript variant two of angiomotin, which encodes an 80 kDa protein (AMOT-p80), showed a very high positive correlation with HOXD10 (Figure 4.14). Strong negative correlation with HOXD10 is exemplified by miR-146a (Figure 4.15). AMOT-p80, miR-146a and CYP2J2 were considered for the DLR assay validation step.

Table 4.3: The 48 genes associated with HOXD10 according to ANOVA test.

Probe ID	Accession number	**Gene symbol**	*p* - value (ANOVA)	Effect of HOXD10 over-expression	Effect of HOXD10 silencing	Fold change
A_33_P3221748	NM_001031680	RUNX3	**8.2E-11**	-	+	2.4
A_23_P23074	NM_006417	IFI44	**1.2E-10**	-	+	4.6
A_33_P3336686	NM_004669	CLIC3	**6.6E-10**	+	-	3
A_23_P85693	NM_004120	GBP2	**4.3E-08**	+	-	2.9
A_33_P3745146	NM_001098517	CADM1	**5.9E-08**	-	+	2.5
A_23_P36611	NM_181861	APAF1	**2.9E-07**	-	+	2.4
A_33_P3290567	NM_003390	WEE1	**3.2E-07**	-	+	2.6
A_23_P300056	NM_044472	CDC42	**4.8E-07**	-	+	2.5
A_24_P344961	NM_133265	AMOT	**9.0E-07**	+	-	3.9
A_32_P164246	NM_033260	FOXQ1	**1.8E-06**	+	-	3.2
A_33_P3718269	NR_029701	MIR146A	**3.0E-06**	-	+	2.8
A_33_P3277198	NM_004854	CHST10	**7.0E-06**	+	-	2.5
A_24_P79403	NM_002619	PF4	**1.3E-05**	-	+	2.7
A_23_P73429	NM_005335	HCLS1	**1.4E-05**	-	+	2.3
A_24_P217572	NM_001957	EDNRA	**2.5E-05**	-	+	3.9
A_23_P401904	NM_001009936	PHF19	**2.8E-05**	-	+	2.5
A_24_P388786	NM_001369	DNAH5	**3.0E-05**	+	-	3.3
A_23_P218770	NM_002872	RAC2	**4.2E-05**	-	+	2.2
A_32_P108156	NR_001458	MIR155HG	**7.2E-05**	-	+	4.3
A_23_P134176	NM_001024465	SOD2	**1.0E-04**	-	+	3.4
A_33_P3318357	NM_138775	ALKBH8	**1.3E-04**	+	-	2.7
A_23_P319859	NM_005244	EYA2	**1.6E-04**	+	-	2.7
A_23_P86599	NM_007329	DMBT1	**5.0E-04**	-	+	2.9
A_24_P146892	NM_032790	ORAI1	**5.2E-04**	-	+	2.7
A_33_P3369844	NM_013230	CD24	**5.5E-04**	+	-	3
A_23_P63896	NM_000043	FAS	**5.5E-04**	-	+	2.7
A_23_P207213	NM_000691	ALDH3A1	**6.2E-04**	+	-	4.1
A_23_P106024	NM_002226	JAG2	**6.4E-04**	+	-	2.8
A_23_P212508	NM_001063	TF	**6.7E-04**	+	-	3.7
A_23_P406424	NM_175744	RHOC	**1.2E-03**	-	+	2.3
A_33_P3627001	NM_144962	PEBP4	**1.3E-03**	+	-	2.3
A_33_P3322085	NM_001102651	ZNF554	**1.4E-03**	+	-	5.8
A_23_P137035	NM_003662	PIR	**1.5E-03**	+	-	2.4
A_24_P383609	NM_199461	NANOS1	**1.7E-03**	+	-	3
A_23_P112482	NM_004925	AQP3	**1.8E-03**	+	-	3
A_33_P3370424	NM_017617	NOTCH1	**1.9E-03**	+	-	2.1
A_23_P204640	NM_024865	NANOG	**4.2E-03**	+	-	2.9
A_23_P18447	NM_013261	PPARGC1A	**5.2E-03**	+	-	3.5
A_23_P60146	NM_006207	PDGFRL	**6.1E-03**	-	+	2.2
A_33_P3211666	NM_003855	IL18R1	**9.9E-03**	-	+	2.8
A_33_P3308744	NM_001105206	LAMA4	**1.0E-02**	+	-	3.1
A_33_P3314902	NM_016252	BIRC6	**1.1E-02**	+	-	2.2
A_23_P134454	NM_001753	CAV1	**1.4E-02**	+	-	2.3
A_24_P365515	NM_021784	FOXA2	**1.9E-02**	+	-	2.8
A_24_P299685	NM_198389	PDPN	**2.9E-02**	-	+	3.4
A_23_P167096	NM_005429	VEGFC	**3.4E-02**	-	+	3.6
A_24_P102053	NM_002538	OCLN	**3.8E-02**	+	-	3.9
A_33_P3235078	NM_001134855	TRIM17	**4.2E-02**	+	-	5.4

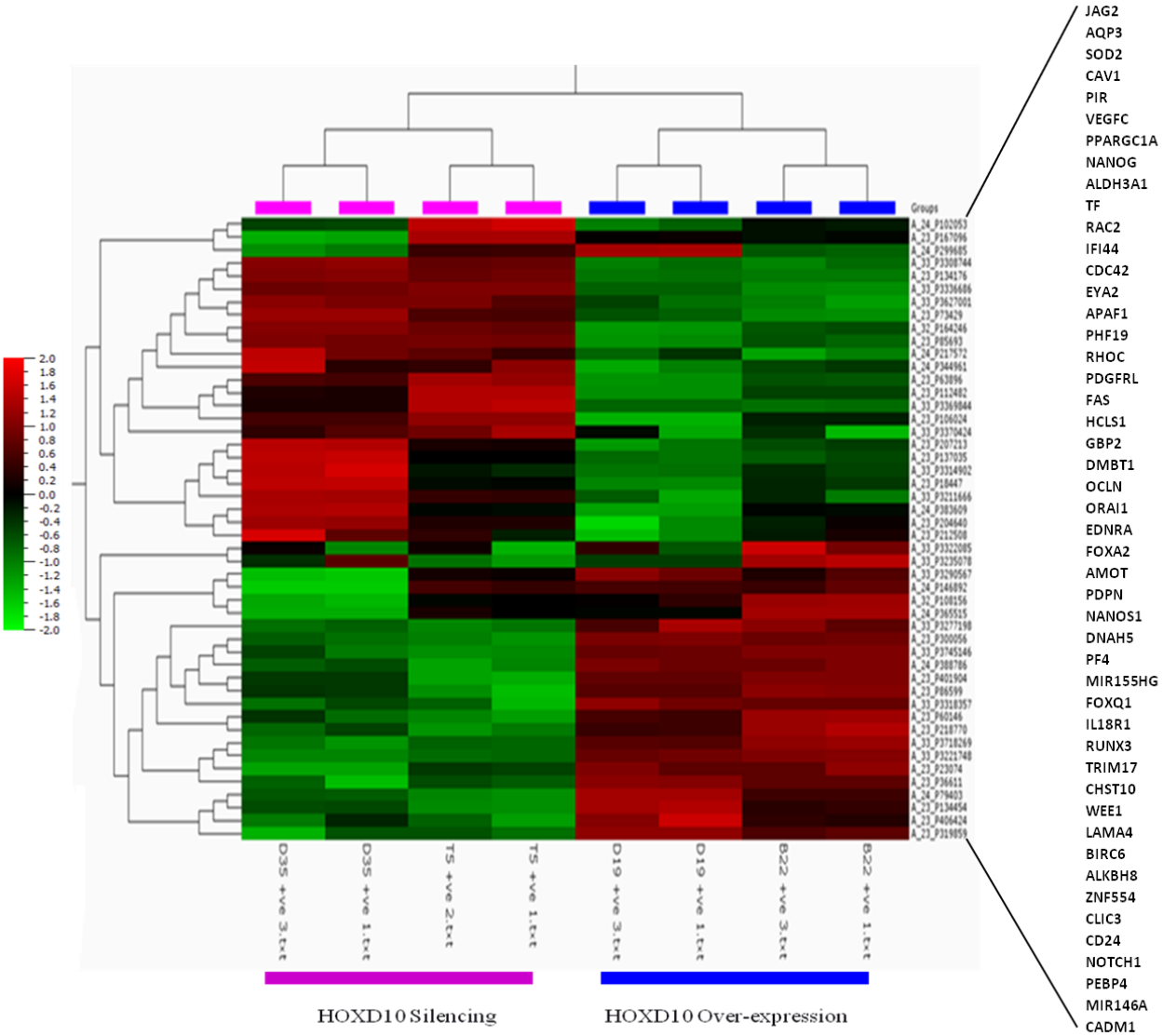


Figure 4.11: A heat-map of the 48 genes associated with HOXD10 according to ANOVA statistical analysis.

The columns represent the assessed cell lines (in duplicates), while the rows represent the detected genes. The (+) mark stands for the HOXD10 manipulation. D35 +ve and T5 +ve are cells subjected to HOXD10 silencing, while D19 +ve and B22 +ve are cells subjected to HOXD10 over-expression.

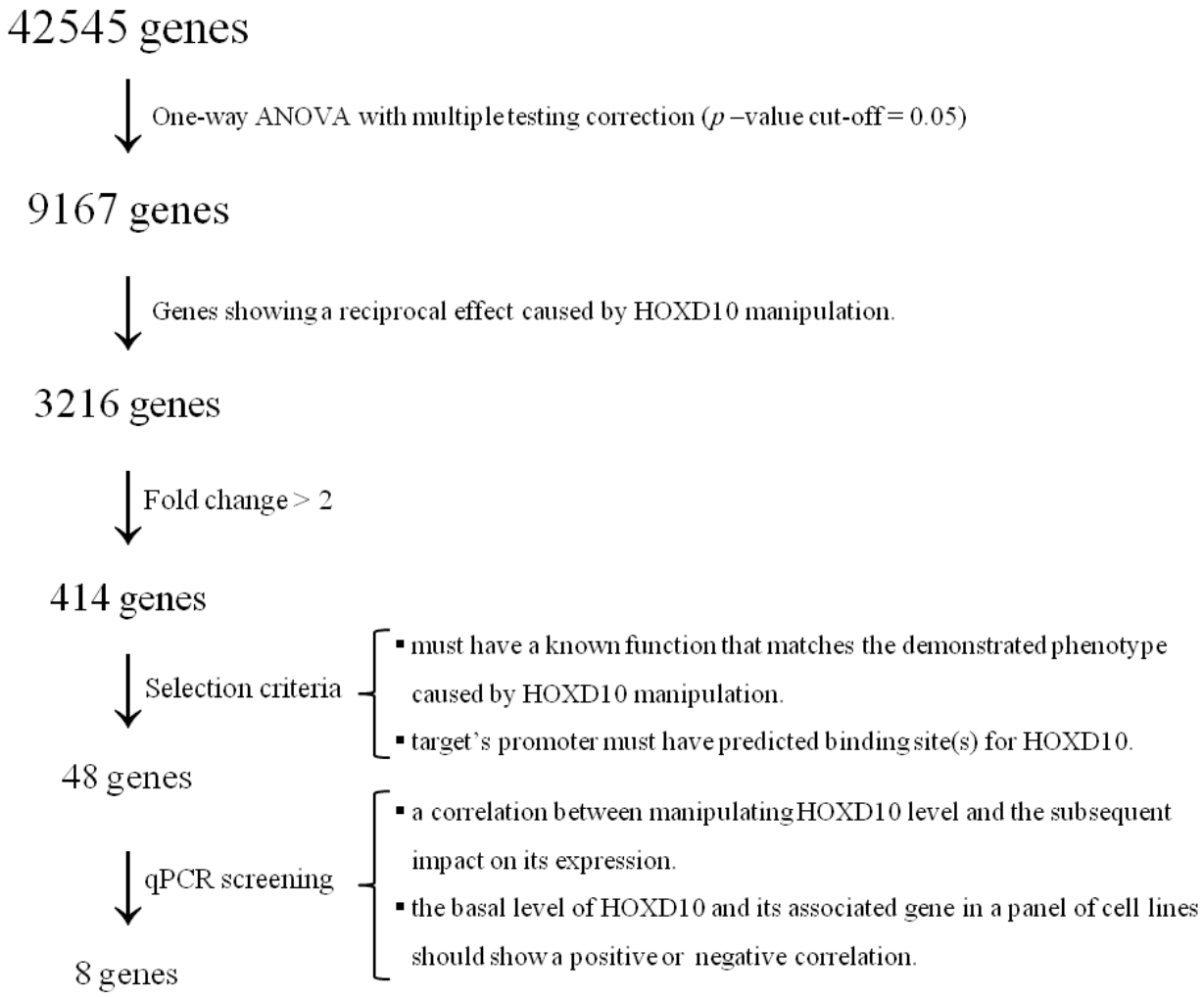


Figure 4.12: A scheme of filtering HOXD10-associated genes identified by ANOVA test.

Table 4.4: The final list of the eight HOXD10-correlated genes identified by ANOVA test.

GENE	Correlation with HOXD10	Examples of reported roles/differential expression
AMOT	+	promotes the proliferation of mammary epithelial cells (Ranahan et al., 2011), regulates endothelial cell migration and is linked to angiogenesis (Ernkvist et al., 2006).
CD24	+	depletion of CD24 reduced cell proliferation, adhesion and enhanced apoptosis (Bretz et al., 2012), while its over-expression contributes to tumour growth in urothelial carcinoma (Overdevest et al., 2012).
CLIC3	+	elevated in breast cancer in comparison with normal breast tissue and drive cell growth and cancer progression (Dozynkiewicz et al., 2012).
AQP3	+	important for growth of oesophageal and oral squamous cell carcinoma (Kusayama et al., 2011).
JAG2	+	highly expressed in tongue carcinoma compared with normal tissues and is a proliferation promoter (Zhang et al., 2011b).
PDGFRL	-	putative tumour suppressor in colorectal cancer cells (Guo et al., 2010).
miR-155	-	suppresses gastric cancer cells growth, migration, and adhesion (Li et al., 2012a).
miR-146a	-	inhibits cell growth, cell migration and induces apoptosis in non-small cell lung cancer cells (Chen et al., 2013a).

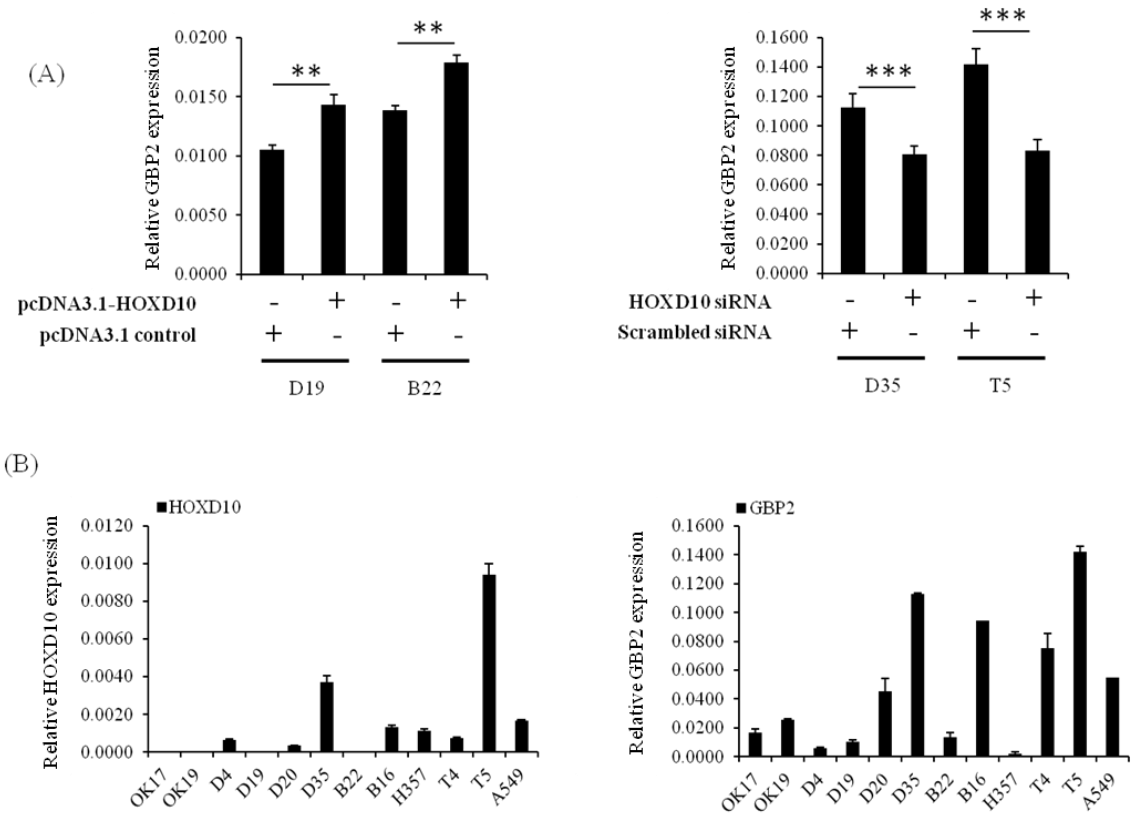


Figure 4.13: HOXD10 basal and manipulated expression level versus GBP2 level.

(A) qPCR analysis showed that GBP2 level positively correlates with HOXD10 level since it increases with HOXD10 increase and vice versa. (B) The basal level of GBP2 does not completely match the level of HOXD10 in different types of cells (** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

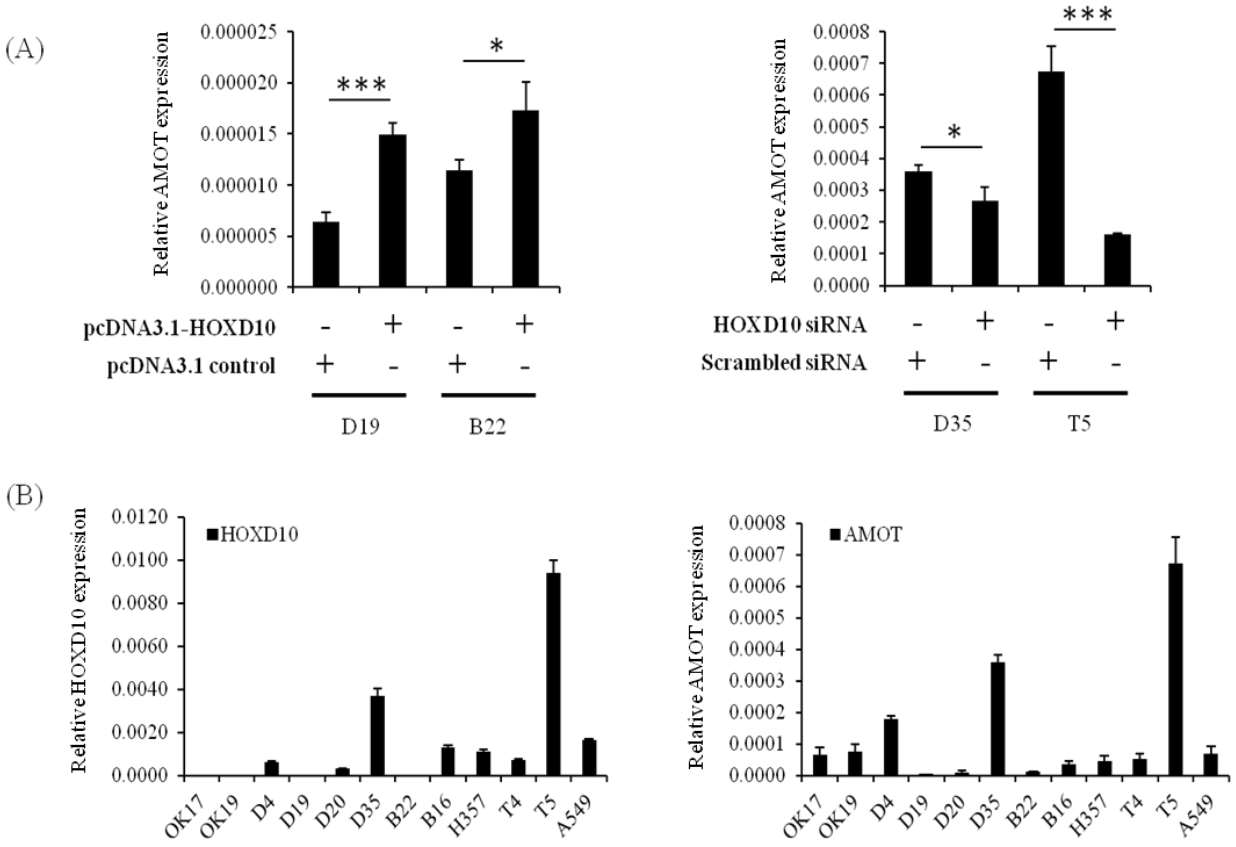


Figure 4.14: HOXD10 basal and manipulated expression level versus AMOT-p80 level.

(A) qPCR analysis shows that AMOT-p80 level positively correlates with HOXD10 level since it increases with HOXD10 increase and vice versa. (B) The basal level of AMOT-p80 matches the level of HOXD10 in different types of cells (* $p < 0.05$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

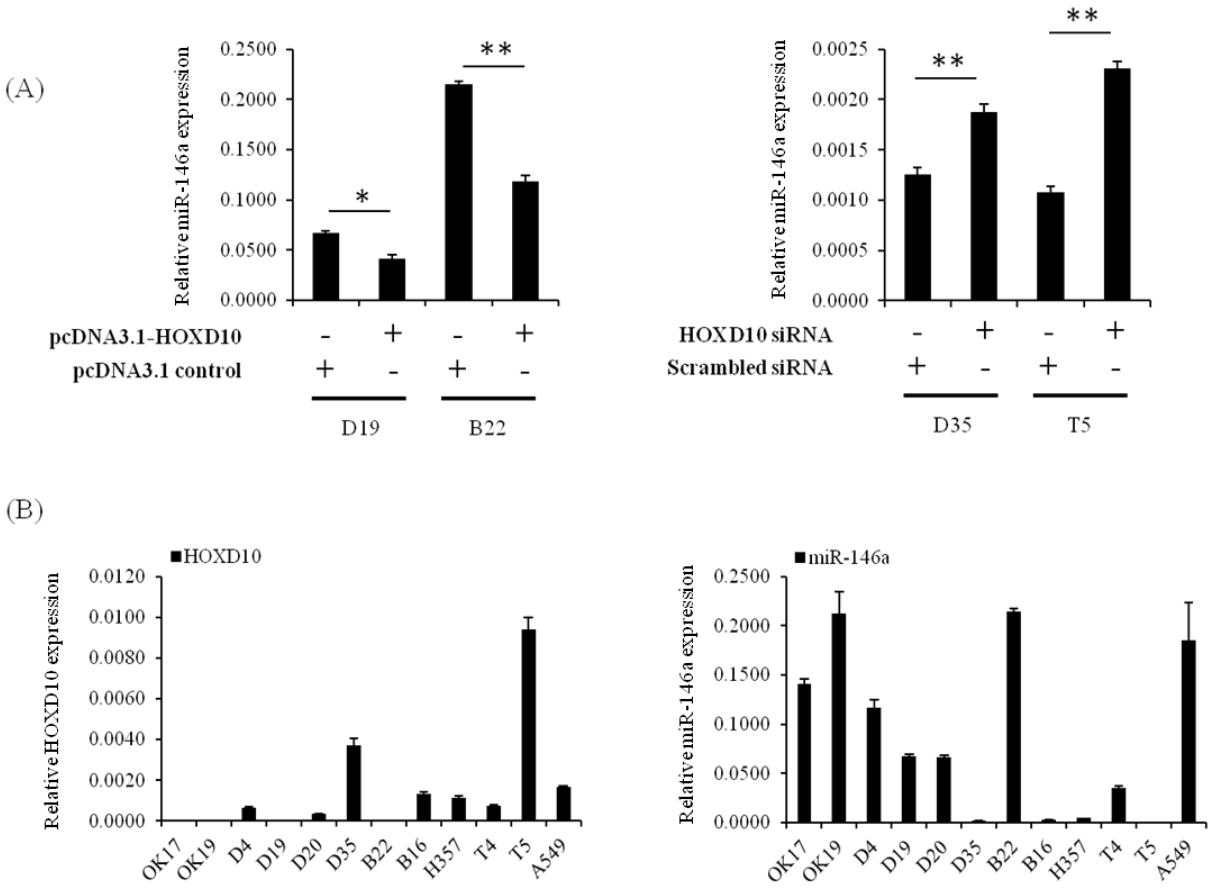


Figure 4.15: HOXD10 basal and manipulated expression level versus miR-146a level.

(A) qPCR analysis shows that miR-146a level negatively correlates with HOXD10 level since it decreases with HOXD10 increase and vice versa. (B) The basal level of miR-146a opposes the level of HOXD10 in different types of cells (* $p < 0.05$, ** $p < 0.01$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

4.10 Discussion

Microarray is a powerful useful tool for a simultaneous analysis of gene expression levels for thousands of genes. However, to avoid misinterpretation, it must be taken into consideration that the output data might be influenced by many factors including variations in sample preparation, labelling and hybridisation steps, and even the method of analysis. Failure to minimise biases before and during analysing the data might reduce the credibility of conclusions. For example, it was demonstrated that having a minimum of a triplicate for each analysed sample made it possible to remove any faulty repeat to avoid its effect on the overall analysis. Following data analysis steps needed continuous evaluation and monitoring to avoid producing skewed results, especially in the case of choosing statistical methods for computing the significance analysis. A number of studies demonstrated these methods might produce overlapping or completely different sets of differentially expressed genes (Kerr et al., 2000, Cui and Churchill, 2003).

Two different statistical methods, unpaired *T*-test and ANOVA (Welch) test, were used for our analysis parameters. *T*-test method was tried first, but the number of genes with differentially expression that satisfied the *T*-test *p*-value and met the filtering criteria was very limited. This was not expected, but not totally surprising since it was reported that *T*-test might not be a proper statistical method for microarray data analysis compared to ANOVA method (Jeanmougin et al., 2010). Moreover, since *T*-tests are mainly used to identify deviation from the mean, large sampling sizes might generate numerous false positives and negatives, which may conclude little if anything about the actual biological change (Nadon and Shoemaker, 2002, Lin et al., 2010). In this particular microarray study, one of the available ways to evaluate the reliability of the used statistical method and the validity of the whole analysis procedure was to detect HOXD10 gene among the output list of genes. *T*-test method failed to detect it while ANOVA did (see Appendix 9). qPCR analysis showed that some genes identified by *T*-test are indeed aligned by HOXD10 expression e.g. CYP2J2. However, their little number mandates re-analysing the raw data but using different statistical method, which is expected to identify different set of genes with differential expression (Jeffery et al., 2006). Also, to perform a GO analysis, a good number of genes were required to develop a complete picture of the biological processes and networks affected by HOXD10 manipulation.

The strategy applied during this study has succeeded in identifying some important HOXD10-interacting molecules in HNSCC and NSCLC. This conclusion was drawn from 1) the previously reported roles of some of these genes in the pathology of HNSCC, NSCLC and many other cancers, and 2) the networks they were mapped to, which explain the phenotypes associated with HOXD10 manipulation. Moreover, according to the GO analysis, it was demonstrated that even in cancer cells, HOXD10 still retains its effect on many genes implicated in development biological processes, which might agree with the "oncology recapitulates ontology" hypothesis. Moreover, some of the HOXD10-associated genes negatively regulated by HOXD10 over-expression are of high importance in inducing cell invasion and metastasis in a

variety of cancers. This includes RhoC (Liu et al., 2012), MMP3 (Zhang et al., 2011a), MMP14 (Li et al., 2012b), INHBA (Kim et al., 2010a), TIMP-1 (Schelter et al., 2011) and TIMP-3 (Zhang et al., 2007). The suppression of some MMPs together with some of their inhibitors, TIMPs, supports the role of HOXD10 in suppressing cell invasion in HNSCC, considering that a previous study showed that their simultaneous up-regulation is needed for cell invasion in this type of cancer (Sutinen et al., 1998). Identifying both genes that match HOXD10-associated phenotypes and genes that link HOXD10 to its role in development strengthens further the validity of our microarray analysis.

Among the genes detected to have an expression that strongly correlate with HOXD10 level in NSCLC and HNSCC cells, CYP2J2 & AMOT-p80, and miR-146a were chosen for further validation as they showed the best positive and negative correlation with HOXD10 expression, respectively, in addition to their important functions reported in cancer studies. CYP2J2 is one of the monooxygenases enzymes family, involved in lipid and steroids metabolism (Berlin et al., 2011), and has an anti-inflammation function (Wray et al., 2009). CYP2J2 was suggested to be a general tumour biomarker due to its frequent over-expression detected in various tumours including oesophageal squamous carcinoma, lung squamous carcinoma and adenocarcinoma, colon carcinoma, hepatocellular carcinoma, breast cancer and gastric cancer (Jiang et al., 2005, Jiang et al., 2009). CYP2J2 is believed to promote tumour cell growth through producing the growth factors regioisomeric epoxyeicosatrienoic acids (EETs) from arachidonic acid (Wang et al., 2005, Chen et al., 2011). Such a role of CYP2J2 in cancer matches HOXD10-associated phenotypes.

AMOT-p80 was firstly identified as a protein target of the angiogenesis inhibitor, angiostatin (Troyanovsky et al., 2001). Then, three more members of the angiomotin family were identified: AMOT-p130 (Ernkvist et al., 2006), AMOT like 1 (AMOTL1) and AMOTL2 (Bratt et al., 2002). AMOT-p130 and AMOT-p80 are splice variants of the angiomotin gene and encode two cell membrane proteins of different sizes, 130 kDa and 80 kDa, respectively. They also have different functions since AMOT-p80 promotes endothelial cell migration and plays a role in angiogenesis, while AMOT-p130 does not, but instead, maintains cell shape through an interaction with actin filaments (Ernkvist et al., 2006). Our microarray study identified, for the first time, AMOT-p80, but not AMOT-p130, as HOXD10-correlating gene, raising the possibility of a variant-specific regulation by HOXD10. Moreover, the up-regulation of the cell motility-inducing variant, AMOT-p80, agrees with HOXD10 positive effect on cell migration demonstrated in our phenotypes studies. Inducing cell proliferation is another important function of AMOT-p80 since it was identified as a potent activator of the extracellular signal-regulated kinase isoforms 1 and 2 (ERK1/2) pathway, which is important for the growth of breast cancer cells (Ranahan et al., 2011), and other types of cancer cells, including oral cancer (Judd et al., 2012) and lung cancer (Chen et al., 2013c). Interestingly, through its N-terminal extension, AMOT-p130, but not AMOT-p80, was found to interact with various molecules to function as a tumour suppressor factor

(Paramasivam et al., 2011, Zhao et al., 2011). However, AMOT-p80 over-expression can antagonise the tumour suppression signalling triggered by AMOT-p130 (Paramasivam et al., 2011).

This study presents the first evidence of an association between miR-146a and HOXD10. miR-146a is one of the abundant microRNAs that have a normal size of 18 to 22 nucleotides and function in modulating gene expression on the post-transcriptional level. MicroRNAs are involved in almost all cellular processes (Bartel, 2004). In cancer, they can function as TSGs or oncogenes (oncomiRs), depending on cancer type and their targets genes, and this applies to miR-146a as well. The expression of miR-146a was reported high in some cancers, such as thyroid carcinoma (He et al., 2005) and cervical cancer (Wang et al., 2008b), while it was suppressed in some other cancers including prostate cancer (Lin et al., 2008), pancreatic cancer (Li et al., 2010c), gastric cancer (Hou et al., 2012a) and breast cancer (Bhaumik et al., 2008).

CHAPTER 5: VALIDATION OF HOXD10 PUTATIVE TARGETS

5.1 Introduction

The mechanism of eukaryotic gene expression is very complex and involves many regulating factors that lead to initiation of or suppression of the transcriptional machinery at the Transcription Starting Site (TSS)-containing core promoter. Enhancer, silencer and insulator elements are examples of DNA elements that mediate turning-on or turning-off the transcriptional machinery (Riethoven, 2010). Typically, the promoter region contains specific binding motifs for the transcription factors (TF) (Lemon and Tjian, 2000). In general, to regulate the expression of a gene, TFs must bind the gene's promoter region, then act as either transcription "activators" through recruiting RNA polymerase II to initiate mRNA synthesis, or as repressors by impeding the transcriptional machinery (Butler and Kadonaga, 2002). In this study, it was demonstrated that HOXD10 induce the expression of *CYP2J2* and *AMOT-p80*, but suppresses the expression of *miR-146a*. To validate the potential interaction between HOXD10 and the promoters of these genes, putative HOXD10-binding sites in a promoter sequence were identified, mutated and then tested for HOXD10 binding compared to wild type sequences. Another approach to confirm that these genes are downstream putative targets of HOXD10 is to manipulate their level to study their effects on cell biological processes, and match them with HOXD10-associated phenotypes.

5.2 Assessment of HOXD10 binding to the promoters of its correlating genes

Promoters of *AMOT-p80*, *CYP2J2* and *miR-146a* (Appendix 12) were identified as previously described in section (2.2.6.2.1). Each promoter sequence was then analysed to determine the possible binding sites for HOXD10 using PROMO online algorithm (Messeguer et al., 2002) (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). A number of studies have shown that the results of this algorithm are reliable (Canbaz et al., 2011, Samarajeewa et al., 2013, Tsai et al., 2013, Wang et al., 2013). The analysis output is exemplified by a schematic diagram for HOXD10 binding sites detected on AMOT-p80 and miR-146a promoters (Figure 5.1). The predicted binding sites with the best score were mutated by mutagenic PCR protocol using mutagenic primers listed in (Table 5.1). Mutagenesis reactions were designed to substitute two to six nucleotides in HOXD10-binding sites with random nucleotides in order to impair HOXD10 binding. The mutated promoter's sequence was then tested again on PROMO online tool to confirm that HOXD10 is not capable of binding at that site. One of the drawbacks of this approach is a possible impairment of other TFs binding or generating a binding site for other TFs. Moreover, HOXD10 might still capable of binding to some other regulatory elements, such as an enhancer sequence which might be far from its predicted binding site on the gene’s promoter. Wild type (WT) and mutated promoters were cloned into pGL3-basic luciferase reporter vector. Cloning and mutagenesis results were confirmed by DNA sequencing as can be seen in the example demonstrated in (Figure 5.2). pGL3-WT and pGL3-mutated promoters, in addition to pRL-TK internal control plasmid, were transfected into D19

cells, which are stably transfected with either pcDNA3.1-HOXD10 or pcDNA3.1 control plasmid. Then, after 48 hours, the change in luciferase expression was detected using the DLR assay. Before computing the statistical analysis, readings of pGL3 luciferase signals were normalised to the readings of pRL-TK luciferase signals to exclude technical errors and any variations in the transfection efficiency. HOXD10 was found able to reduce the expression of luciferase encoded in pGL3 plasmid containing miR-146a WT promoter. However, its effect significantly decreased when its binding sites were mutated (Figure 5.3, A). Also, HOXD10 was demonstrated to drive the expression of luciferase encoded in pGL3 plasmid containing AMOT-p80 WT promoter. However, HOXD10 effect significantly decreased when its binding sites were mutated (Figure 5.3, B). No significant effect occurs on the level of luciferase encoded in pGL3 plasmid containing either CYP2J2 WT promoter or CYP2J2 mutated promoter (Figure 5.3, C).

More validation of HOXD10 binding to AMOT-p80 promoter was carried out by, firstly, mutating another binding site on AMOT-p80 promoter. Then, AMOT-p80 promoter with either one of the mutated binding sites or both were cloned into pGL3 and tested for HOXD10 binding using DLR assay as performed before. Results showed that HOXD10 can indeed bind to AMOT-p80 promoter and its binding can be affected by mutating either of its two binding sites. However, mutating both binding sites in the same promoter did not show a strong further reduction in HOXD10 effect (Figure 5.4).

Table 5.1: The cloning and mutagenic primers used to amplify, clone and mutate AMOT-p80, CYP2J2 and miR-146a promoters.

Promoter	Primer type	Forward (5' → 3')	Reverse (5' → 3')
AMOT-p80	Cloning	TGCTAACTCTCGAGAGTCAACTTC ATATCCACCCCCAAAA	TACGCCAAGCTTACGACCAAGTTCA TGCCACCAT
	Mutagenic (site#1)	CACAATAGCCTCTTGTTTAGT***CC***T ATTAATTTTGAGGGCGGGTGG	CCACCCGCCCTCAAAATTAATA***GG***A CTAAACAAGAGGCTATTGTG
	Mutagenic (site#2)	GGTGGAAGCTGCTTTTCAGCAT***A*** ***A***CTACTATTTGTTTTTCTTCCTCCC	GGGAGGAAGAAAAACAAATAGTAG ***TT***ATGCTGAAAAGCAGCTTCCACC
CYP2J2	Cloning	ATGGATCTCGAGCAGCTGTAGTA AAACAGGAAAATGTGGA	TAACGTCAAGCTTGCTGCAGAACA GAGTTAGGGTCAG
	Mutagenic	CTGTCTCAAAAATAAAGAAATA***CGG*** ***GGGC***ATATTGTAAGCAGAATATGT	ACATATTCTGCTTACAATAT***GCCCC*** ***CG***TATTTCTTTATTTTTGAGACAGA
miR-146a	Cloning	AAAATTCTCGAGTTGAAAAGCCAA CAGGCTCATTGG	CAAATTTAAGCTTCCACTCCAATCG GCCCTGCT
	Mutagenic	AGGGTGTGGAAATGGAATA***TTT***G CATATGCAAATAGGCCTT	AAGGCCTATTTGCATATGC***AAA***TAT TCCATTTCCACACCCT

* Restriction sites (*XhoI* and *HindIII*) are underlined.

** Site of HOXD10 binding is highlighted.

*** Site of mutation is in an italic bold font.

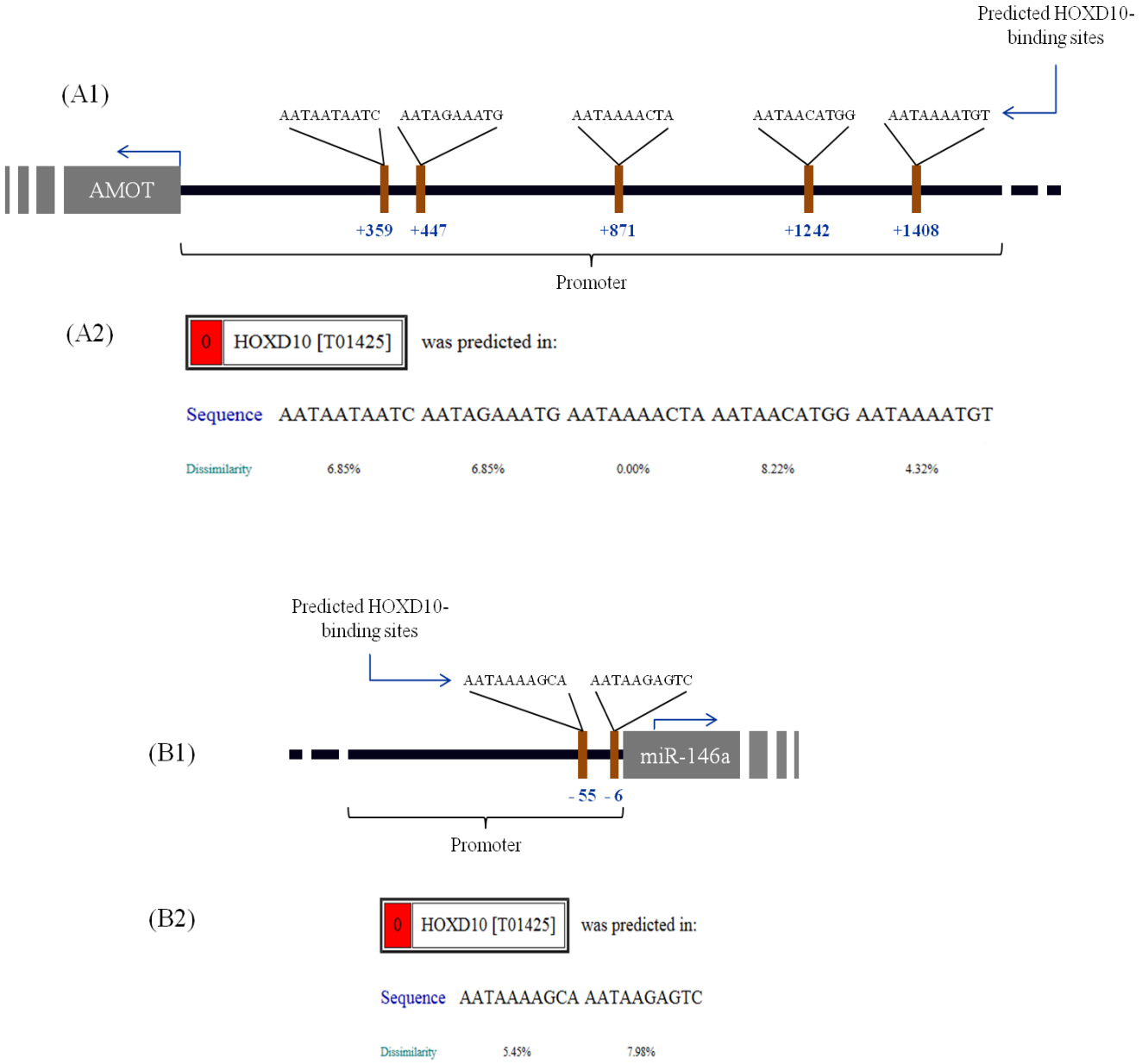


Figure 5.1: A schematic diagram of HOXD10 binding sites in AMOT-p80 and miR-146a promoters.

(A1) *In-silico* analysis using the ALGGEN-PROMO programme identified five different binding sites on AMOT-p80 promoter which differ in their matrix dissimilarity. (A2) Theoretically, the less matrix dissimilarity rate, the highest possibility of a true HOXD10 binding site (maximum matrix dissimilarity rate was set at ten percent). (B1) Two HOXD10-binding sites in the promoter of miR-146a were identified with matrix dissimilarity rates of 5.45% and 7.89% (B2).

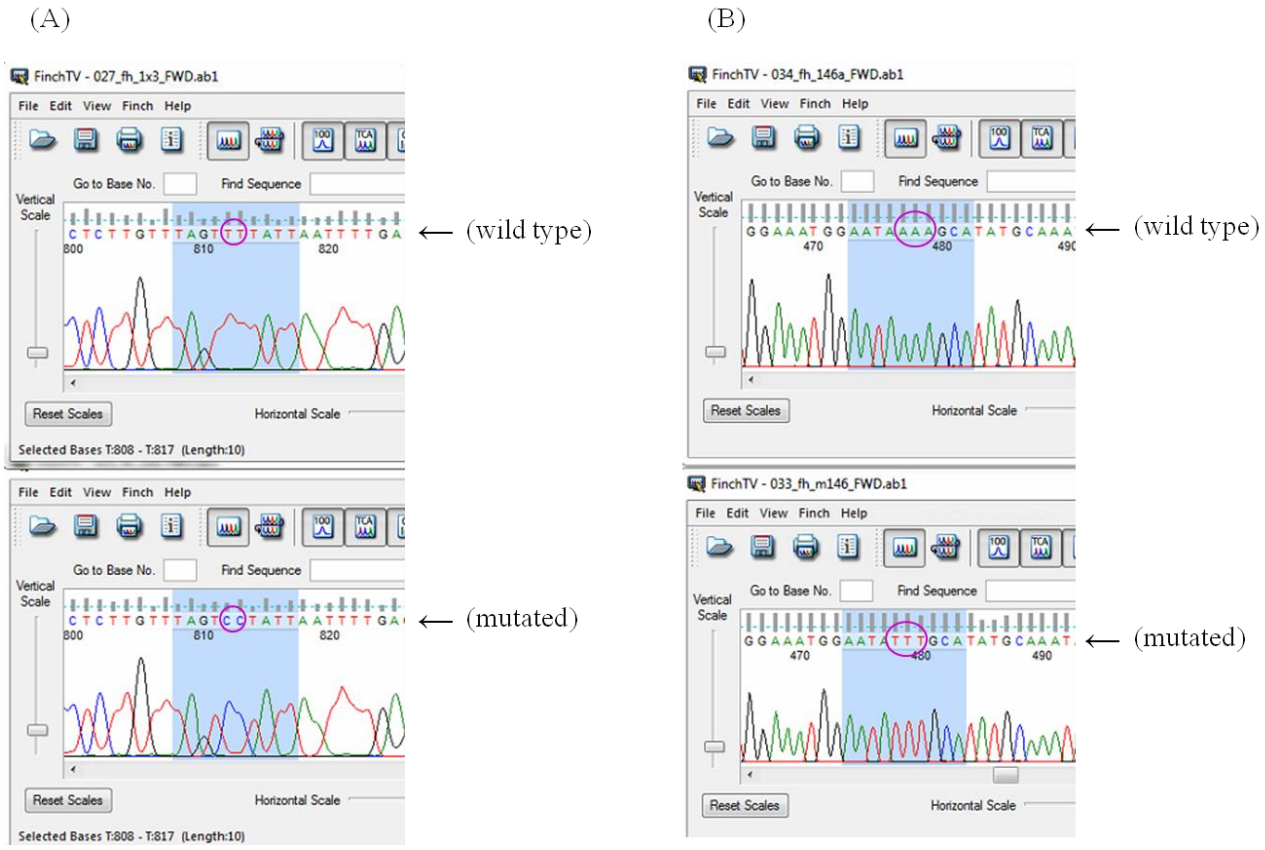


Figure 5.2: Snapshots of the sequencing results demonstrated on Finch TV for HOXD10-binding sites in AMOT-p80 and miR-146a promoters.

(A) Mutagenic PCR was used to mutate one of the HOXD10-binding sites on AMOT-p80 promoter by introducing the sequence TAGTCCTATTA instead of the sequence TAGTTTTATT. (B) The HOXD10-binding site wild type sequence in miR-146a promoter, AATAAAAGCA, was substituted with AATATTTGCA nucleotides in order to impair HOXD10 binding.

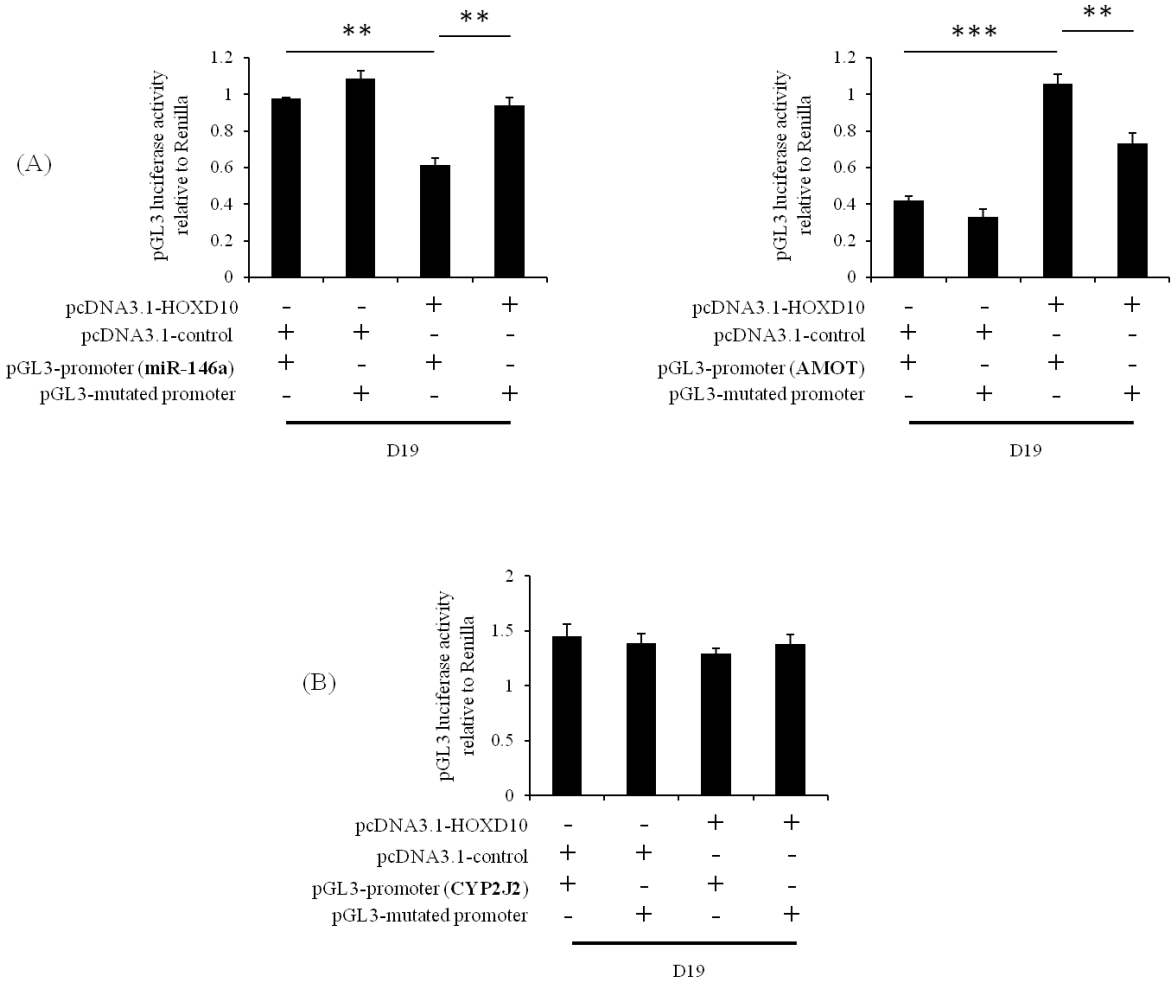


Figure 5.3: DLR analysis of HOXD10 binding to the wild-type and mutated promoters of *AMOT-p80*, *miR-146a* and *CYP2J2* genes.

(A) DLR analysis shows that in the presence of HOXD10, the expression of the luciferase encoded in pGL3 vectors containing AMOT-p80 and miR-146a promoters was affected positively and negatively, respectively. This effect decreased significantly when HOXD10-binding sites were mutated. (B) No effect caused by the presence of HOXD10 was demonstrated in the case of CYP2J2 WT or mutated promoters (** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

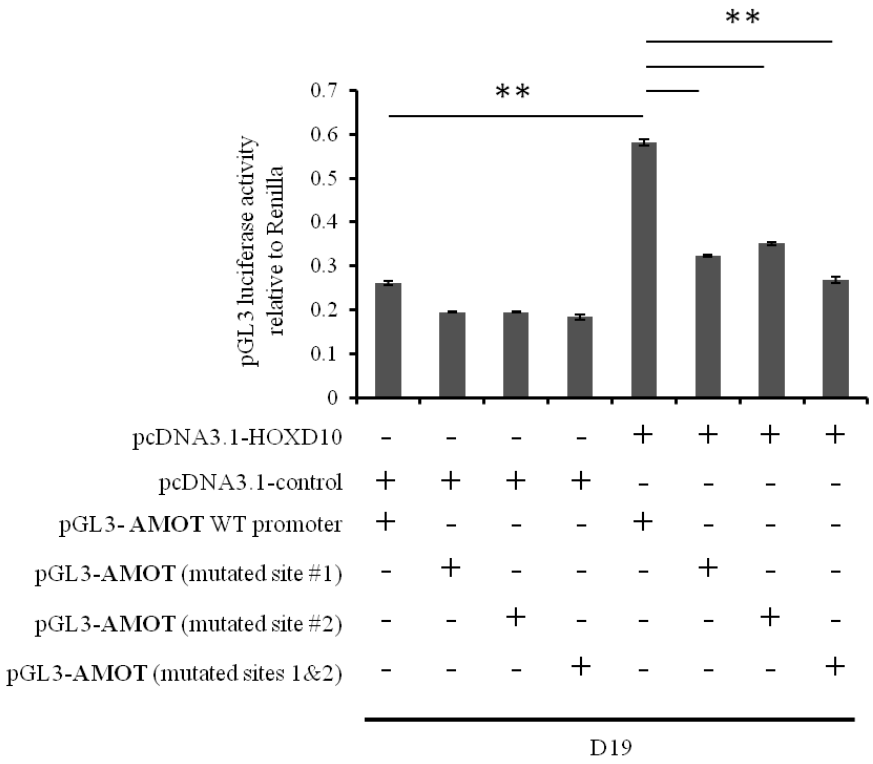


Figure 5.4: DLR analysis of HOXD10 binding to WT and mutated promoters of AMOT-p80.

DLR analysis shows that HOXD10 is capable of driving the luciferase expression when pGL3 vector contains WT promoter. However, luciferase expression decreases when either of the HOXD10 binding sites in AMOT-p80 promoter is mutated. No significant further reduction in luciferase expression was detected when both mutated sites existed in the same pGL3 vector compared to pGL3 vector containing only one mutated binding site (** $p < 0.01$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

5.3 AMOT-p80 protein in HNSCC tissues and cell lines

A small cohort of eight HNSCC and four normal oral mucosa tissue sections were examined by IHC protocol to detect AMOT-p80 protein using a commercial specific Ab. High level of AMOT-p80 was detected in HNSCC samples compared to no detection or very low level detected in normal tissues (Figure 5.5). Moreover, in a small panel of cell lines, AMOT-p80 protein level, which was detected by WB, matches its mRNA level, which was detected by qPCR analysis (Figure 5.6).

5.4 AMOT-p80 protein expression in cells transfected with cloned HOXD10 or AMOT-p80

Using WB protocol, AMOT-p80 was detected in lysates of D19 and B22 cells stably transfected with pcDNA-3.1-HOXD10 compared to no detection detected in control cells (Figure 5.7, A). To validate these bands, and for future usage, AMOT-p80 coding sequence was cloned into the pcDNA3.1 mammalian expression vector (pcDNA3.1-AMOT-p80) using the following primers: forward primer, 5'-CACCATGCCTCGGGCTCAGCCATCCTC-3' and reverse primer, 5'-TTAGATGAGATATTCCACCATCTCTGCATCAGGCTCTTGTC-3'. pcDNA3.1-AMOT-p80 was transfected into a new batch of low-AMOT-p80 expressing D19 and B22 cells. qPCR analysis showed a marked increase in AMOT-p80, which validates the successful cloning of AMOT-p80, while the detected bands by WB were of the same size detected previously in cells transfected with HOXD10 (Figure 5.7, B).

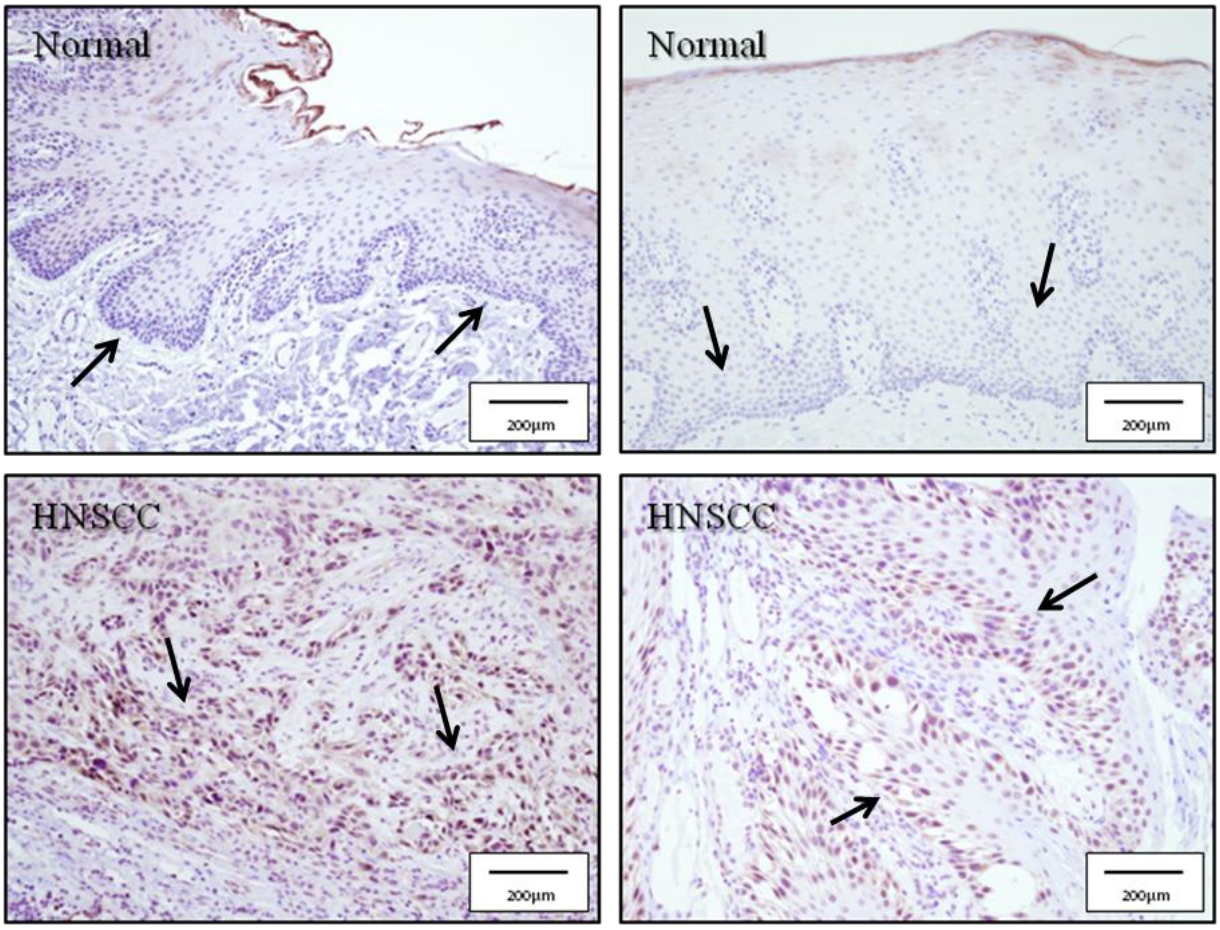


Figure 5.5: Detection of AMOT-p80 protein in normal and tumour tissue sections.

A panel of four normal and eight HNSCC tumour tissue sections were screened for AMOT-p80 protein expression level. AMOT-p80 protein is highly expressed in tumour tissues compared to very low expression demonstrated in normal tissues.

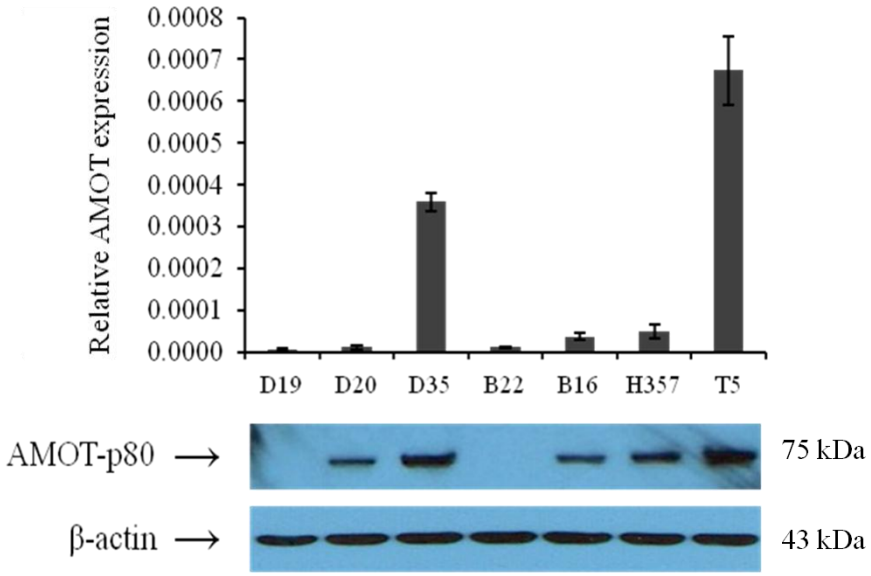


Figure 5.6: AMOT-p80 protein and mRNA levels in a panel of cell lines.

A bar graph of AMOT-p80 mRNA level in a panel of HNSCC cells as determined by qPCR (upper panel). Western blot analysis of AMOT-p80 protein in the same panel of HNSCC cells (lower panel). Beta actin protein expression is shown as a control for loading (qPCR data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

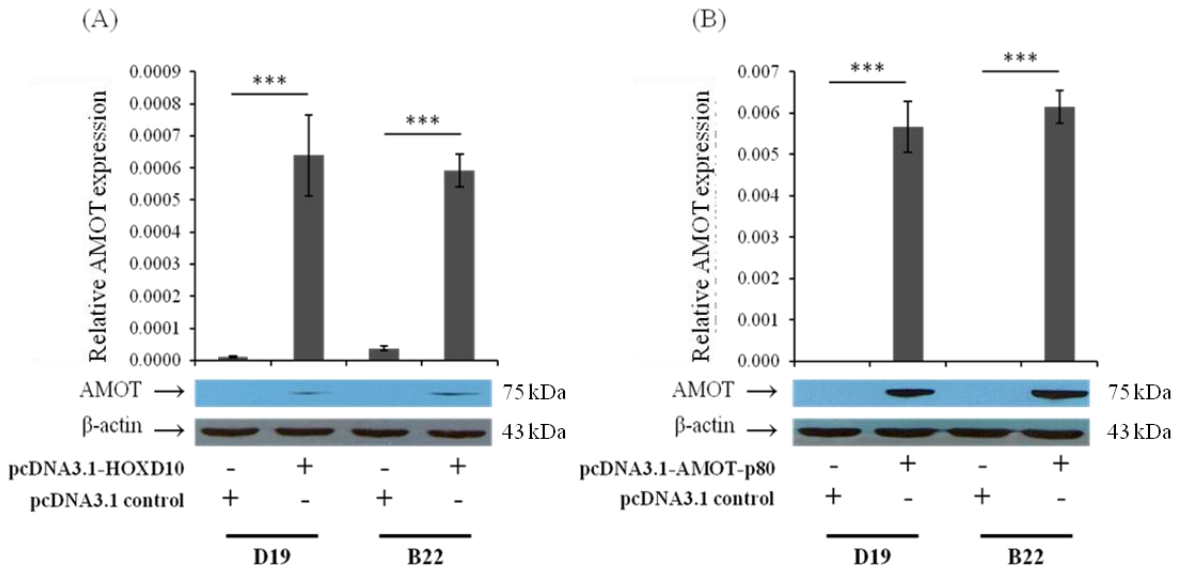


Figure 5.7: qPCR and WB analysis of AMOT-p80 expression driven by either HOXD10 or AMOT-p80 transfection.

(A) A bar graph of AMOT-p80 mRNA level in D19 and B22 HNSCC cells stably transfected with HOXD10 or a control plasmid as determined by qPCR (upper panel). Western blot analysis of AMOT-p80 protein in the same HNSCC cells (lower panel). (B) A bar graph of AMOT-p80 mRNA level in D19 and B22 HNSCC cells transiently transfected with AMOT-p80 coding plasmid or a control plasmid as determined by qPCR (upper panel). Western blot analysis of AMOT-p80 protein in the same HNSCC cells (lower panel). Beta actin protein expression is shown as a control for loading. (*** $p < 0.001$, parametric *T*-test, qPCR data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

5.5 The function of AMOT-p80 in HNSCC cells

Transient transfection of AMOT-p80 into low-HOXD10 and AMOT-p80 expressing HNSCC cells (B22 and D19) resulted in phenotypic changes similar to transfection with HOXD10 as in both cell lines there was an increase in cell proliferation and migration (Figure 5.8, A & B). There was no significant effect on cell adhesion caused by high AMOT-p80 level (Figure 5.8, C). This might indicate that other genes are involved in this phenotype demonstrated before when HOXD10 expression level was manipulated.

High-HOXD10 expressing T5 cells were stably transfected with HOXD10 shRNA, which resulted in a reduction of HOXD10 level and HOXD10-putative target, AMOT-p80, by more than sixty percent according to qPCR analysis (Figure 5.9, A). To study the effect of retrieving AMOT-p80 level in cells which underwent stable HOXD10 knock down, the amount of pcDNA3.1-AMOT-p80 transfected DNA was firstly optimised using a range of DNA amount (10 ng to 200 ng) to identify the proper amount capable of introducing AMOT-p80 level in a similar value to its basal level in T5 cells not subjected to HOXD10 silencing. As can be seen in (Figure 5.9, B), 50 ng was the suitable amount to rescue the reduction in AMOT-p80 level in HOXD10 shRNA-stably transfected T5 cells on both mRNA and protein levels. Functional assays were performed for cells stably transfected with HOXD10 shRNA and transiently transfected with 50 ng of pcDNA3.1-AMOT-p80 DNA. It was demonstrated that re-expressing AMOT-p80 was able to rescue the negative effect of HOXD10 silencing on cell phenotype to some extent, with partial reversal of the reduction in proliferation (Figure 5.10, A) and reversion of migration to a level higher than original control levels (Figure 5.10, B).

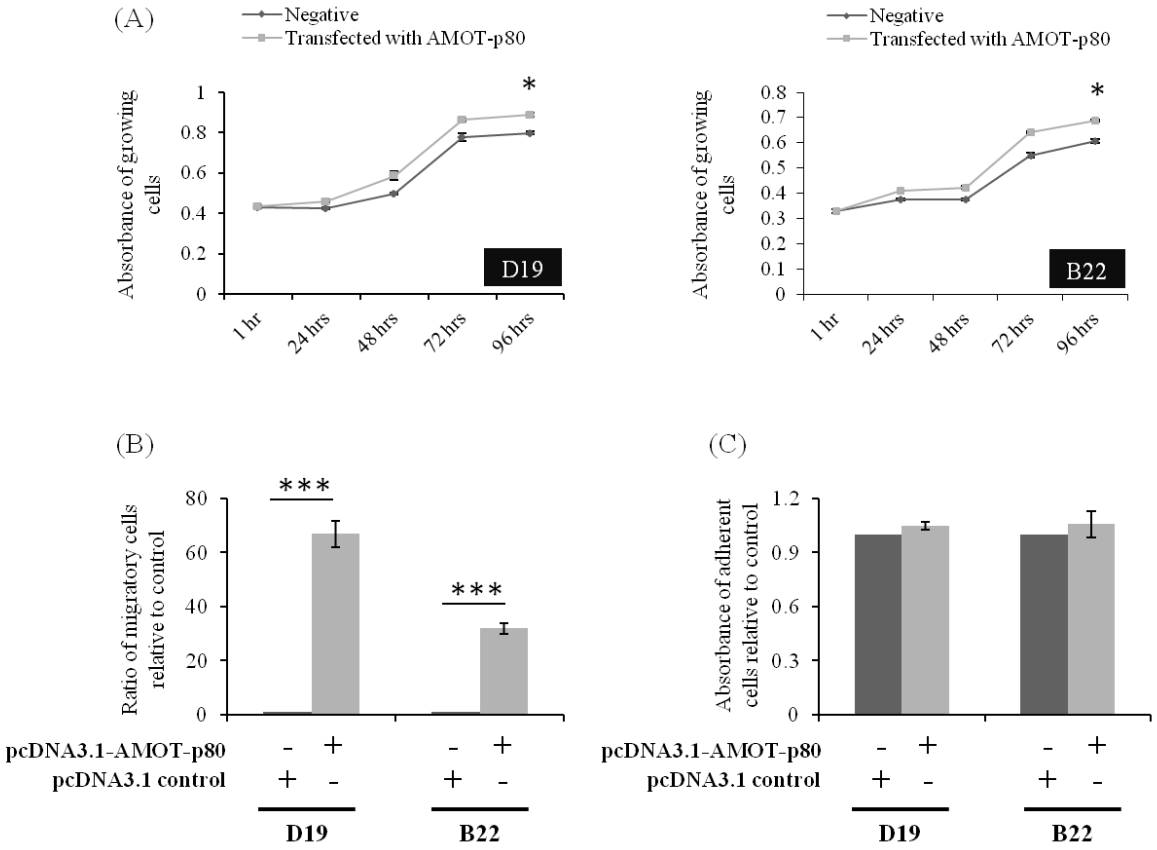


Figure 5.8: Effect of AMOT-p80 over-expression on cell proliferation, migration and adhesion.

(A) Introducing AMOT-p80 into low-AMOT-p80 cells (D19 and B22) resulted in an increase in their proliferation. (B) Cells transfected with AMOT-p80 showed very significant increase in their ability to migrate through the porous membrane of the Transwell migration system compared to cells transfected with a control plasmid. (C) Assessment of cell adhesion has shown that introducing high levels of AMOT-p80 by transfection had no significant effect (* $p < 0.05$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

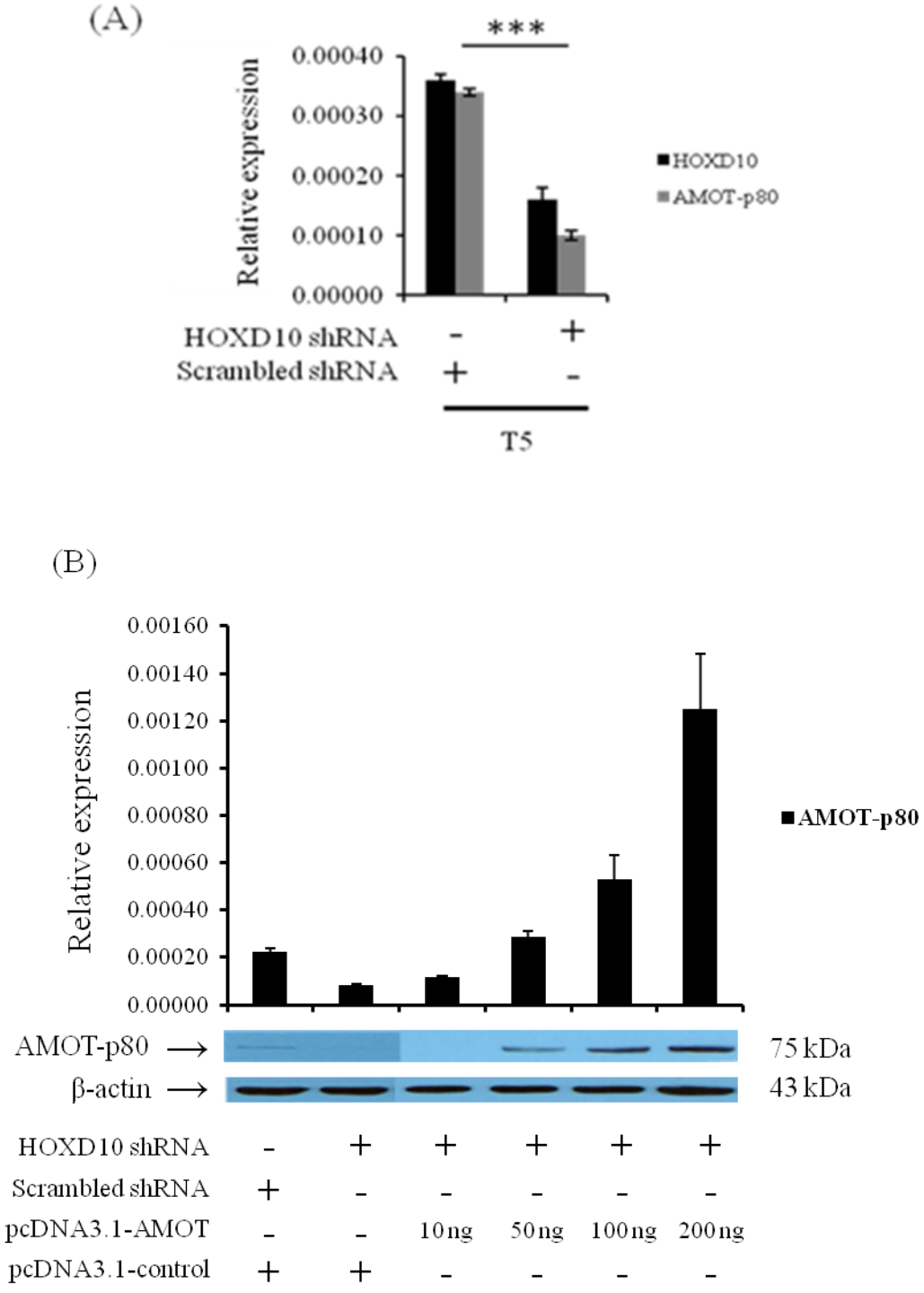


Figure 5.9: Rescuing AMOT-p80 expression following HOXD10 knocks down.

(A) The direct silencing effect of HOXD10 shRNA on HOXD10 expression level and its indirect effect on AMOT-p80 expression level as determined by qPCR. (B). qPCR and WB analysis of AMOT-p80 mRNA and protein expression levels, respectively, in HOXD10-silenced cells following introducing different amounts of AMOT-p80 encoding vector (*** $p < 0.001$, parametric *T*-test, qPCR data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

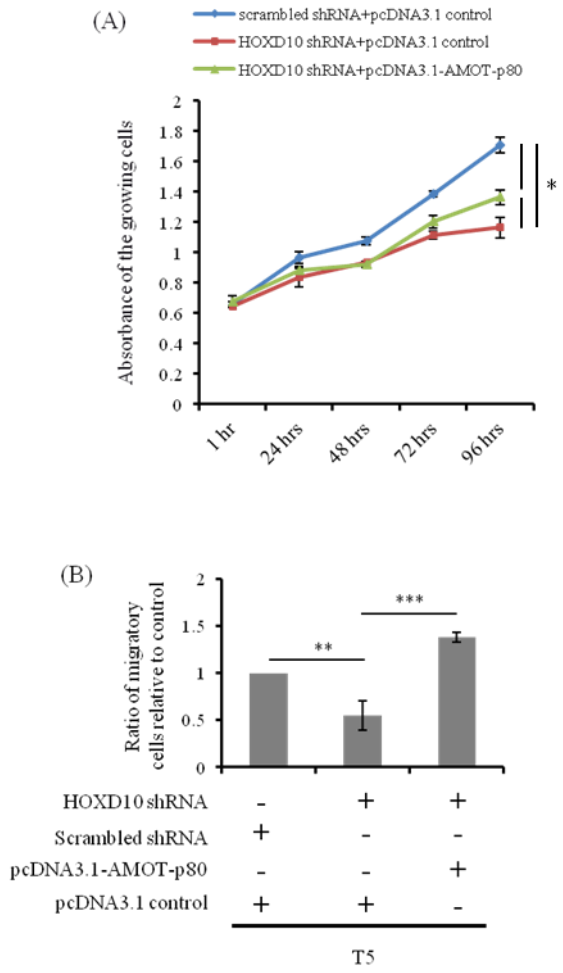


Figure 5.10: The impact of rescuing AMOT-p80 expression on the migration and proliferation of HOXD10-silenced cells.

(A) The MTS assay demonstrated that the re-increase in AMOT-p80 level in HOXD10-knocked down cells led to re-promoting cells proliferation after 96 hours compared to HOXD10- knocked down cells transfected with a control pcDNA3.1 plasmid. (B) The decrease in T5 cells migration caused by HOXD10 silencing was significantly rescued when AMOT-p80 level was recovered (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

5.6 The effect of miR-146a on its downstream targets in HNSCC and NSCLC

Although it is a future plan to investigate further the role of miR-146a in NSCLC and HNSCC, its role in these pathologies was briefly studied at the molecular interaction level. Firstly, a potential regulatory loop between miR-146a and HOXD10 was investigated. It was shown previously that HOXD10 over-expression down-regulates miR-146a level (Figure 4.15), however, introducing high levels of miR-146a molecules into normally high-HOXD10 expressing cells, such as D35 and T5 cells, did not affect the level of HOXD10 (Figure 5.11). Therefore, it is unlikely that miR-146a targets HOXD10 mRNA.

The importance of miR-146a in cancer is linked to its role in affecting the level of some cancer-associated genes, such as EGFR (Bhaumik et al., 2008, Xu et al., 2012), BRCA1 (Garcia et al., 2011), IRAK1 (Hung et al., 2013), CXCR4 (Labbaye et al., 2008) and L1CAM (Hou et al., 2012b). Most of these molecules, if not all, promote tumour cell progression in different types of cancer. In addition to these miR-146a targets, a target that also exists in the 414 genes list is NOTCH1. NOTCH1 is over-expressed in HNSCC (Yoshida et al., 2013) and NSCLC (Li et al., 2010b). In glioblastoma, miR-146a was found to inhibit cell proliferation and migration through targeting NOTCH1 (Mei et al., 2011). This molecular interaction might occur in HNSCC and NSCLC as well since HOXD10 negatively and positively regulates the expression of *miR-146a* and *NOTCH1*, respectively (Figure 5.12, A and B). As miR-146a was confirmed as a highly putative target of HOXD10 while no putative HOXD10-binding sites in NOTCH1 promoter sequence were detected, it is more likely that the up-regulation in NOTCH1 is a result of suppressing its up-stream regulator, miR-146a, by HOXD10. This is also supported by the negative correlation between miR-146a and NOTCH1 expression in a panel of cell lines (Figure 5.12, C).

To explore further the indirect effect of HOXD10 on the downstream targets of miR-146a, two more known miR-146a targets were chosen for qPCR analysis, L1CAM (Hou et al., 2012b) and CXCR4 (Labbaye et al., 2008). Interestingly, L1CAM is over-expressed in both HNSCC (Hung et al., 2010) and NSCLC (Katayama et al., 1997). CXCR4 is also up-regulated in HNSCC (Delilbasi et al., 2004) and NSCLC (Na et al., 2008). Indeed, HOXD10 over-expression causes a reduction in miR-146a and an increase, mostly indirectly, in the level of miR-146a downstream targets, CXCR4 and L1CAM (Figure 5.13, A). An opposite effect was seen in the case of HOXD10 silencing (Figure 5.13, B). However, it cannot be confirmed at this point that the previously reported over-expression of L1CAM and CXCR4 in HNSCC and NSCLC is always a result of HOXD10 over-expression.

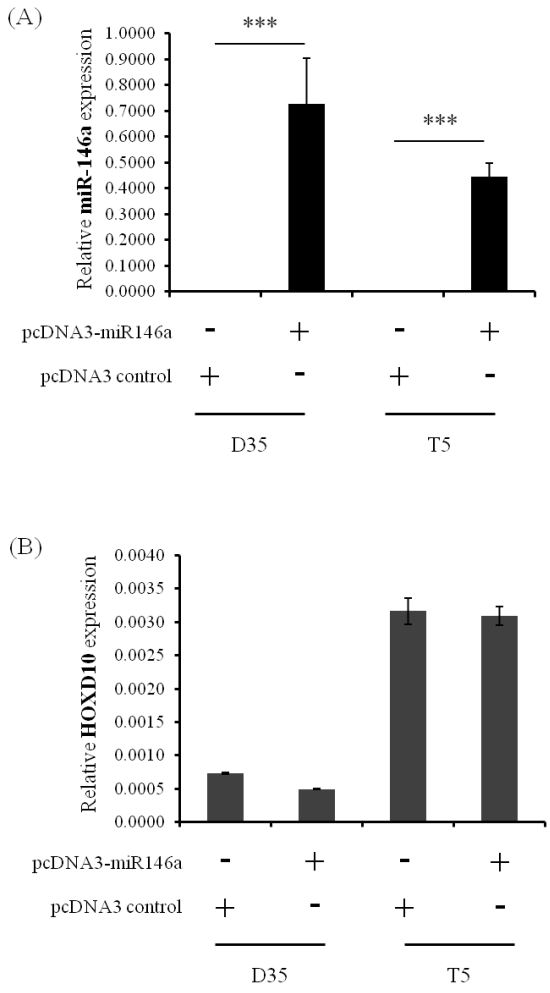


Figure 5.11: The effect of miR-146a over-expression on HOXD10 level.

(A) Pre-miR-146a transfection into D35 and T5 cells resulted in a marked increase in miR-146a level. (B) HOXD10 level was not affected by the increase in miR-146a expression level in both normally high-HOXD10 expressing cells, D35 and T5 cells (*** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

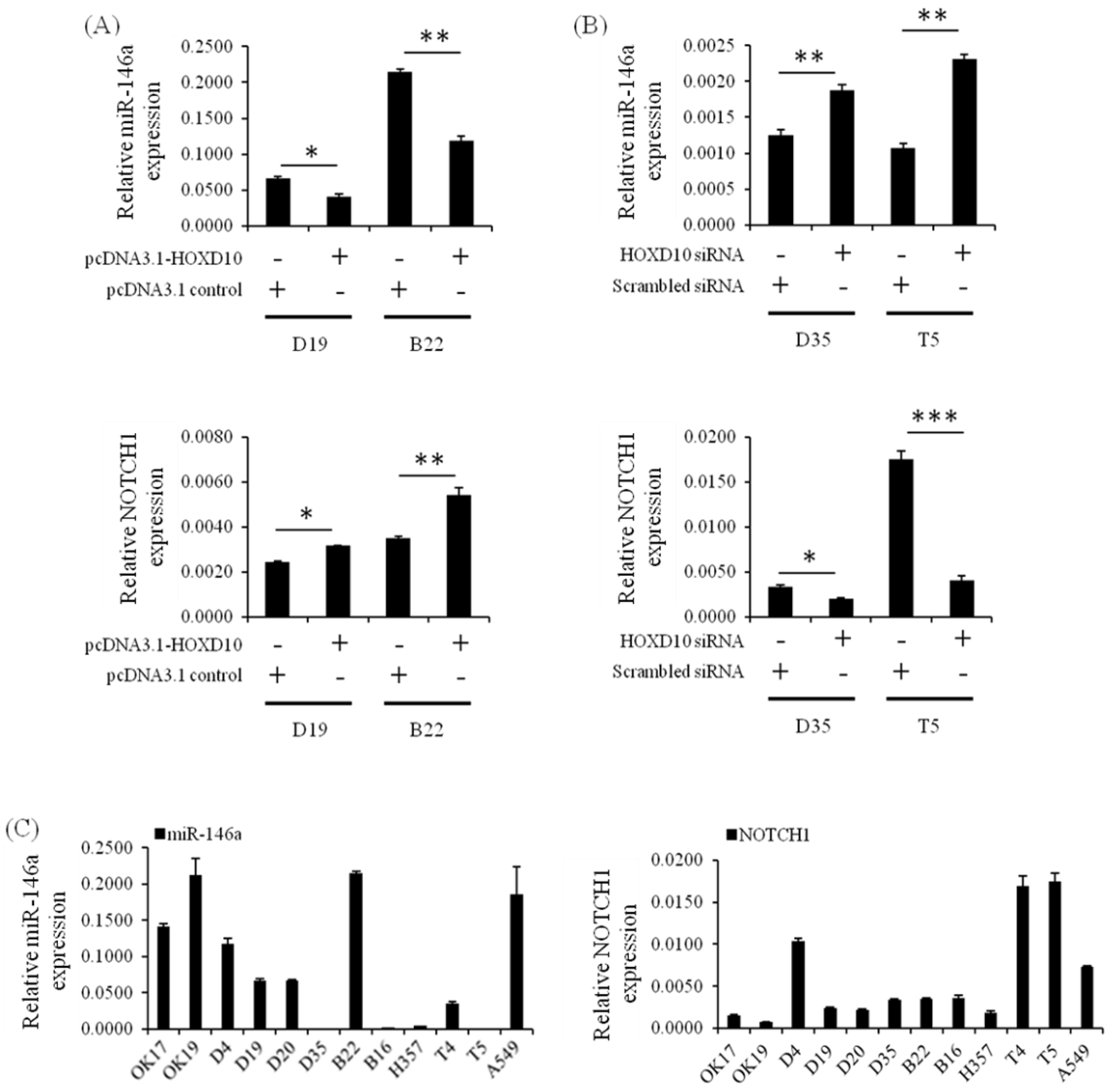


Figure 5.12: miR-146a and NOTCH1 levels in HOXD10-manipulated cells and in a panel of cell lines.

(A) HOXD10 over-expression caused a reduction in miR-146a and an increase in the level of its downstream targets, NOTCH1. An opposite effect was seen in the case of HOXD10 silencing (B). (C) Comparing the basal level of miR-146a with NOTCH1 emphasises their negative correlation except for D4 cells. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

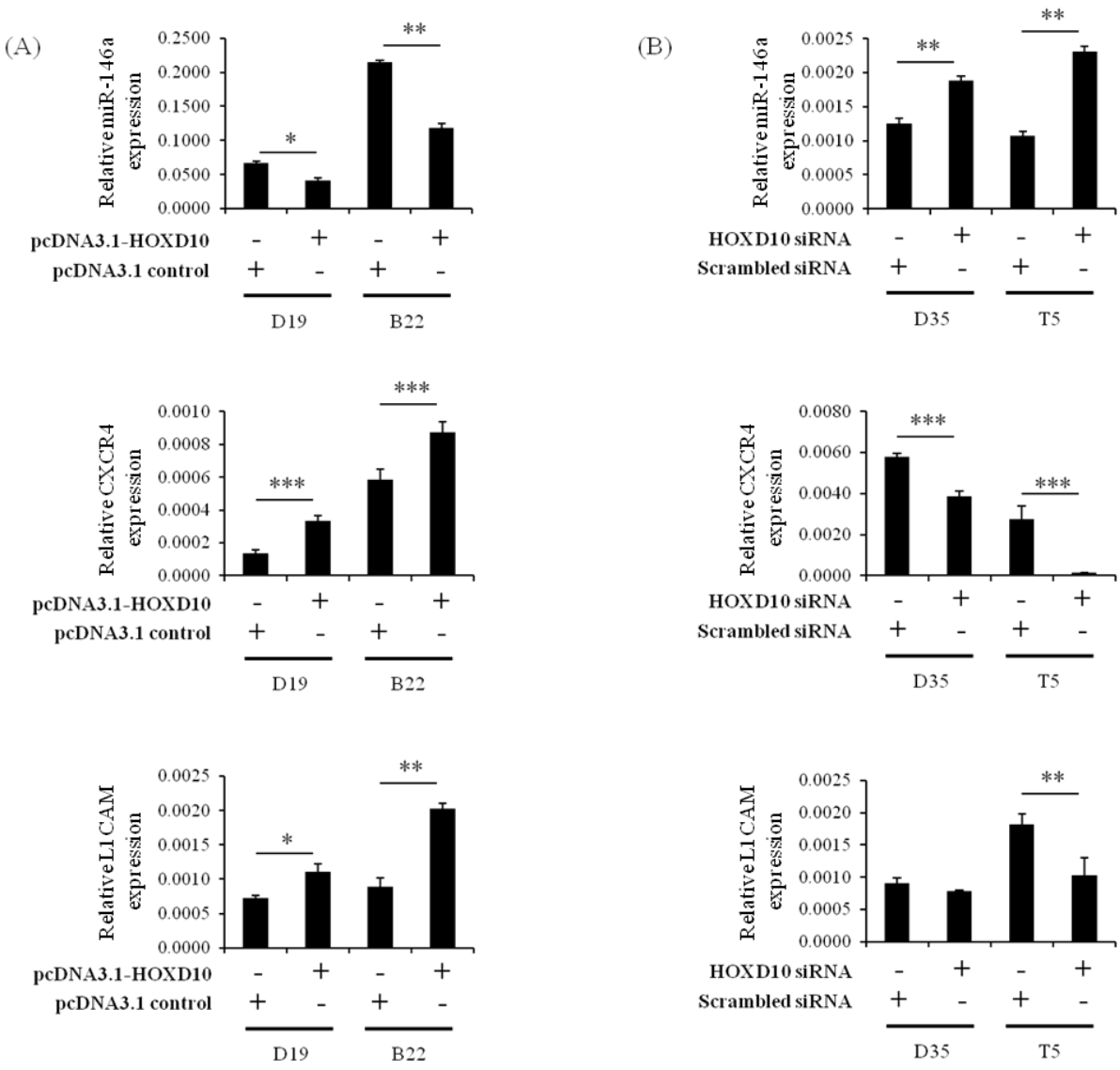


Figure 5.13: Effect of HOXD10 manipulation on miR-146a and its downstream targets, CXCR4 and L1CAM.

(A) HOXD10 over-expression causes a reduction in miR-146a and an increase in the level of its downstream targets, CXCR4 and L1CAM. An opposite effect was seen in the case of HOXD10 silencing (B) (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, parametric *T*-test, data are represented as mean +/- S.E.M taken over a minimum of three independent experiments).

5.7 Discussion

This microarray study identified a large number of HOXD10-correlated genes, some of which may be targets of HOXD10. Since it is a transcription factor, HOXD10 exerts its effects on the expression of its targets through their promoters. Our strategy for HOXD10 candidate identification nominated three genes, CYP2J2, AMOT-p80 and miR-146a. Their previously reported functions in cancer agreed with HOXD10 demonstrated effects in HNSCC and NSCLC, and showed a correlation with HOXD10 expression that indicates behaviour of a direct target. Multiple studies have used methods like Chromatin Immunoprecipitation (ChIP) to study an interaction between a promoter sequence and some transcription factors. However, the low sensitivity of the only Ab found specific for HOXD10 protein in cell lysate in this study abrogated the possibility of using this technique. Moreover, the strategy applied in this study to identify an interaction between a TF and a target's promoter, which involves predicting TF binding sites by an *in-silico* analysis, PCR mutagenesis and DLR assay, has been successfully applied before in several studies (Canbaz et al., 2011, Jiang et al., 2011, Snead et al., 2011, Hou et al., 2013, Thompson et al., 2013). In our study, this strategy produced high quality results and identified two novel targets of HOXD10, AMOT-p80 and miR-146a. It was decided to study the link between AMOT-p80 functions and phenotypes demonstrated by HOXD10 manipulation, while the miR-146a-associated phenotypes were left for future work.

The protein expression level of AMOT-p80 matched the previous qPCR analysis. Assessment of expression of AMOT-p80 in a panel of normal oral mucosa and HNSCC tissues showed that the expression of AMOT-p80 was abundant in the HNSCC samples compared to very low expression in normal oral mucosa. Matching levels of HOXD10 protein detected previously in the same panel. Moreover, the differential expression of AMOT-p80 in a panel of cell lines was validated and showed that its mRNA level reflects its protein level. The link between HOXD10 and the up-regulation of AMOT-p80 was further confirmed by detecting an increase in AMOT-p80 protein caused by HOXD10 over-expression. Also, targeted knockdown of HOXD10 level by shRNA in some cell lines has reduced AMOT-p80 protein level as well. Collectively, these findings support AMOT-p80 as a direct target of HOXD10, and that HOXD10 is responsible for increasing AMOT-p80 level in the studied cancer tissues and cell lines.

To explore the phenotypes that would be affected by AMOT-p80 manipulation, AMOT-p80 was successfully cloned and transfected into some low-AMOT-p80 expression cells, as confirmed by qPCR and WB. Introducing high levels of AMOT-p80 caused a significant increase in cell migration and proliferation, which agrees with HOXD10 effects and suggests that AMOT-p80 is one of the downstream targets of HOXD10 that mediates its role in cell migration and proliferation. This was further confirmed by designing a rescue experiment in which, HOXD10 was knocked down to cause a decrease in AMOT-p80 level and to negatively affect cell migration and proliferation. However, when AMOT-p80 level was restored without restoring HOXD10 level, the affected cell proliferation and migration were rescued.

The role of the other HOXD10 novel target, miR-146a, in HNSCC and NSCLC was studied briefly. An agreement between the suppression of miR-146a by HOXD10 and expected up-regulation of its downstream targets was demonstrated. Many of these targets are known to be over-expressed in HNSCC and NSCLC and important for their progression, which supports our findings that their regulator, miR-146a, is down-regulated in these two cancers. Also, HOXD10 level was demonstrated to be unaffected by high miR-146a level, which exclude the possibility of a regulation loop between HOXD10 and miR-146a.

CHAPTER 6: GENERAL DISCUSSION AND FUTURE WORK

6.1 General discussion

The complexity of HNSCC and NSCLC is a result of the various types of genetic alterations that occur at different stages of these diseases. The effect of these alterations ranges between promoting the tumour development, inducing the spread of cancer cells into adjacent and distant body parts and resisting therapies. Moreover, the effect of a particular gene alteration may be greater if it has the ability to up-regulate oncogenes or down-regulate tumour suppressor genes. This might be one of the main functions of cancer-associated transcription factors, such as *HOX* genes. Although they are known as the master regulators of body formation during embryogenesis, many of them have been described as playing a role in the development of a wide range of cancers. Their roles range between tumour suppression and promoting progression, depending on cancer type and stage. Many genetic alterations and abnormal genes expression are shared between HNSCC and NSCLC including *HOXD10* and several other *HOX* genes, which are deregulated in cancer tissues compared to normal tissues.

The number of cancer studies reporting differential expression of HOXD10 and its interacting molecules is small; studies in breast and gastric tissue report association with the expression of *miR-10b* & *miR-7* and *IGFBP3*, respectively. This is the first study to explore the functional roles and molecular effect of HOXD10 in HNSCC and NSCLC progression. The up-regulation or down-regulation of *HOXD10* in different types of cancers highlights the context and cell-specific roles of this transcription factor, particularly as its expression may be regulated by a number of different mechanisms. The cancer-type dependent role of HOXD10 might be supported by the lack of association between its level in HNSCC and NSLCL and some previously reported interacting molecules, such as *miR-7* and miR-10b. Such an interaction was identified to mediate the anti-metastasis role of HOXD10 in breast and ovarian cancer. The same role was demonstrated when HOXD10-associated phenotypes were assessed, but its anti-metastasis function in HNSCC and NSCLC might be mediated through different pathways, such as modulating EMT-related genes.

It has been shown that high expression of HOXD10 in cells from primary HNSCCs promotes their proliferation, migration and adhesion. These effects may promote the development of the primary tumour as it becomes established at the primary site. Moreover, our microarray study identified a wide spectrum of genes associated with HOXD10 expression; many are implicated in HNSCC and NSCLC development and progression according to our GO analysis. Thus, it is likely that increasing HOXD10 expression during HNSCC development supports the emergence of cells which populate the primary tumour. It is also demonstrated that HOXD10 expression is reduced in HNSCC metastases relative to their paired primary tumours. The observations that HOXD10 suppresses invasion into Matrigel, that expression of HOXD10 is low in cells derived from metastases, that the tumour loses expression of HOXD10 in its metastases and that HOXD10 modulates EMT by decreasing and increasing the level of some EMT-related genes suggest a change in the role HOXD10 is playing as cancer develops and

spreads to sites distant from the primary tumour. The pattern of high HOXD10 expression in the primary tumour, with loss of expression in metastases which was observed in the studies of cancers, has been reported in other cancers, such as bladder cancer (Nakayama et al., 2013).

The differential pattern of expression observed here poses interesting questions regarding the possible uses of HOXD10 both as a biomarker and potential therapeutic target. Of the tumours which metastasised, all demonstrated the pattern of high expression in the primary tumour, low expression in metastases. No significant association of loss of HOXD10 expression in the primary tumour with the presence of metastasis was identified. Thus, it is unlikely to be useful as a prognostic biomarker. As high and low HOXD10 expression can support tumour development, then metastasis, respectively, considering HOXD10 a potential therapeutic target might need more investigation accompanied by an *in-vivo* experiment. It was demonstrated that sustained decrease of HOXD10 by around sixty percent of its normal level endows cancer cells with an invasive ability *in-vitro*. However, knocking down HOXD10 by ninety percent or more of its normal level using a high amount of siRNA molecules kills the transfected cells. This might be a potential therapeutic strategy considering that using siRNA to target cancer cells is a very promising field in cancer therapy; a number of clinical trials are based on this approach (Lee et al., 2013). However, since stable and effective delivery of siRNA *in-vivo* is still challenging (Seth et al., 2012), considering anti-HOXD10 siRNA as a valid approach to tackle HNSCC and NSCLC cells might depend on identifying and interrupting the main downstream targets of HOXD10 responsible for inducing cell invasion when HOXD10 level is not completely lost. Therefore, blocking HOXD10 might inhibit tumour development as long as the invasion-inducing elements are also blocked to inhibit the metastatic transition. On the other hand, HOXD10 possible utility as a marker of progression of OPLs to HNSCC warrants further investigation, given the high expression in some OPL cells and tissues which may be related to risk of progression to an invasive disease. Also, identification of the targets of HOXD10 at the particular stage of disease development indicated may identify useful biomarkers or novel therapeutic targets.

The expression microarray analysis to identify possible targets of HOXD10 was used. This is subject to a number of limitations, including the large number of indirect effects which will be seen on manipulation of HOXD10 expression. However, the large number of genes affected, directly or indirectly, by HOXD10 manipulation affirms the multi-roles of HOXD10 in HNSCC and NSCLC. The algorithm applied for target identification was successful in the identification of a number of novel candidate HOXD10 targets, including AMOT-p80 and miR-146a. The role of miR-146a in NSCLC was not studied extensively. However, one study has identified a down-regulation of miR-146a in this cancer, and when its expression is restored, it impairs cell proliferation and migration (Chen et al., 2013a). In the case of HNSCC, miR-146a expression level in primary tumour cells has been a controversial subject. According to some studies, up-regulation of miR-146a was detected in HNSCC tissues and cell lines (Cervigne et al., 2009, Lajer et al., 2011, Hung et al., 2013). However, no alteration in miR-146a expression was

demonstrated in a number of other HNSCC studies (Chang et al., 2008, Hui et al., 2010, Kimura et al., 2010, Gombos et al., 2013). The reason for this discrepancy is not clear. However, it was reported in different types of cancer, including oral cancer, that the single nucleotide polymorphism (G>C; rs2910164) in the pre-miR-146a sequence can affect the mature miR-146a expression level, with an association between variant C allele and the high miR-146a level (Shen et al., 2008, Jazdzewski et al., 2009, Xu et al., 2010, Hung et al., 2012). The screened primary HNSCC and NSCLC cell lines showed a down-regulation of this miR compared to normal cells. In addition, an up-regulation in the level of some of its previously reported downstream targets was identified, which agrees with the loss of miR-146a in HNSCC and NSCLC. Moreover, most of these downstream targets are known to promote HNSCC and NSCLC progression, raising the possibility that HOXD10 induces cancer development by up-regulating these genes indirectly through suppressing miR-146a. Moreover, miR-146a has been reported as a tumour suppressor in a number of studies: inhibiting TNF receptor-associated factor 6 (TRAF6)/nuclear factor κB (NFκB) pathway in lymphoma (Paik et al., 2011); impairing cell proliferation in myeloid leukaemia (Boldin et al., 2011); reducing cell proliferation and migration in hepatocellular carcinoma through targeting Rho-associated protein kinase (ROCK1) (Rong et al., 2013), suppressing tumour growth and progression by targeting EGFR pathway in prostate cancer (Xu et al., 2012) and NSCLC (Chen et al., 2013a). Thus, the finding that HOXD10 decreases miR-146a expression is in keeping with the observed phenotype and the role of HOXD10 in promoting tumour development.

Although HOXD10 was reported to inhibit angiogenesis in breast cancer (Myers et al., 2002), the differences between HOXD10-interacting molecules in breast cancer and what has been identified here suggests a different relationship between HOXD10 and angiogenesis in HNSCC and NSCLC. The known function of AMOT-p80 in angiogenesis might link HOXD10, indirectly, to blood vessel formation in HNSCC and NSCLC primary tumour sites. Again, this makes HOXD10 an attractive target for cancer therapeutics as if inhibited; there may be effects on both the primary tumour and the tumour vasculature. Moreover, AMOT-p80 has been reported to promote tumour growth in a number of different contexts: signalling via RAS-MAPK to promote both proliferation of embyronal kidney cell lines and enhanced development of xenograft tumours in mice using NF2 null Schwann cells (Yi et al., 2011); signalling via ERK1/2 in breast cancer cells, resulting in increased proliferation and dysregulated cell polarity which resulted in a more neoplastic growth pattern (Ranahan et al., 2011). Our data in HNSCC support the role of AMOT-p80 in the promotion of tumour growth, although it is not clear whether HOXD10 plays a role in the control of AMOT-p80 expression in these other tumour types. Nevertheless, AMOT-p80 is a potential target for novel therapeutics which may be useful in a number of cancers.

The mechanism of HOXD10 down-regulation in HNSCC cells as they metastasise is not clear. In contrast with that seen in breast and ovarian cancer, there is no direct relationship between the expression of HOXD10 and its negative regulator, miR-10b, in HNSCC and NSCLC

according to our study. This finding is supported by a recent HNSCC study (Severino et al., 2013). Thus, there is scope for further investigation of other potential mechanisms responsible for HOXD10 expression modulation, both in terms of the initiating promotion of expression and its subsequent loss in lymph node metastases.

Although identifying the regulatory mechanism behind HOXD10 over-expression in primary tumour site and its down-regulation in metastasising cells is beyond the scope of this project, some possible ways are suggested within these studies. Demonstrating a dominant HOXD10 over-expression in primary cancer cells was demonstrated, with occasional low expression and a variable expression in OPL cells, suggests that the regulatory factor that initiates HOXD10 over-expression in primary tumour sites might be triggered by changes in cancer cell density. This is supported by the finding that cell proliferation, which is probably the main function of HOXD10 in HNSCC and NSCLC according to our study, might be induced by irregular cell-to-cell communication, or quorum sensing mechanism disruption (Hickson et al., 2009, Agur et al., 2010). Moreover, the potential link between HOXD10 and blood vessel formation in the tumour microenvironment might suggest the presence of a regulatory mechanism triggered by the need for growth factors and nutrients. Studying the HOXD10 promoter to detect potential regulatory elements that are already involved in cell-to-cell communication, cell growth, or angiogenesis, might identify the factor behind HOXD10 over-expression. The significant reduction of HOXD10 in metastasising cells is mostly due to its anti-invasion function. Therefore, it is possible to combine literature mining with *in-silico* analysis to identify some invasion-promoting molecules predicted to target HOXD10 mRNA, such as microRNAs, or *HOXD10* promoter, such as a transcription factor.

In conclusion, HOXD10 is not expressed in normal OKs, but is detectable in some OPL cells and is highly expressed in primary tumour cells then significantly reduced in metastases. Such stage-specific expression suggests that HOXD10 supports the development of the tumour at different stages by differential modulation of the phenotype, including the development of EMT which is so closely linked to metastasis. Also, it has been demonstrated that HOXD10 might exert some of its roles in HNSCC and NSCLC through affecting the level of many important molecules including AMOT-p80 and miR-146a. This also has an impact on their previously reported interacting molecules. Such direct or indirect effect of HOXD10 might be behind some of its effect on cell proliferation, migration adhesion and invasion (Figure 6.1). HOXD10 and its novel associating molecules may represent therapeutic targets which will be able to specifically interfere with these particular tumour promoting stages.

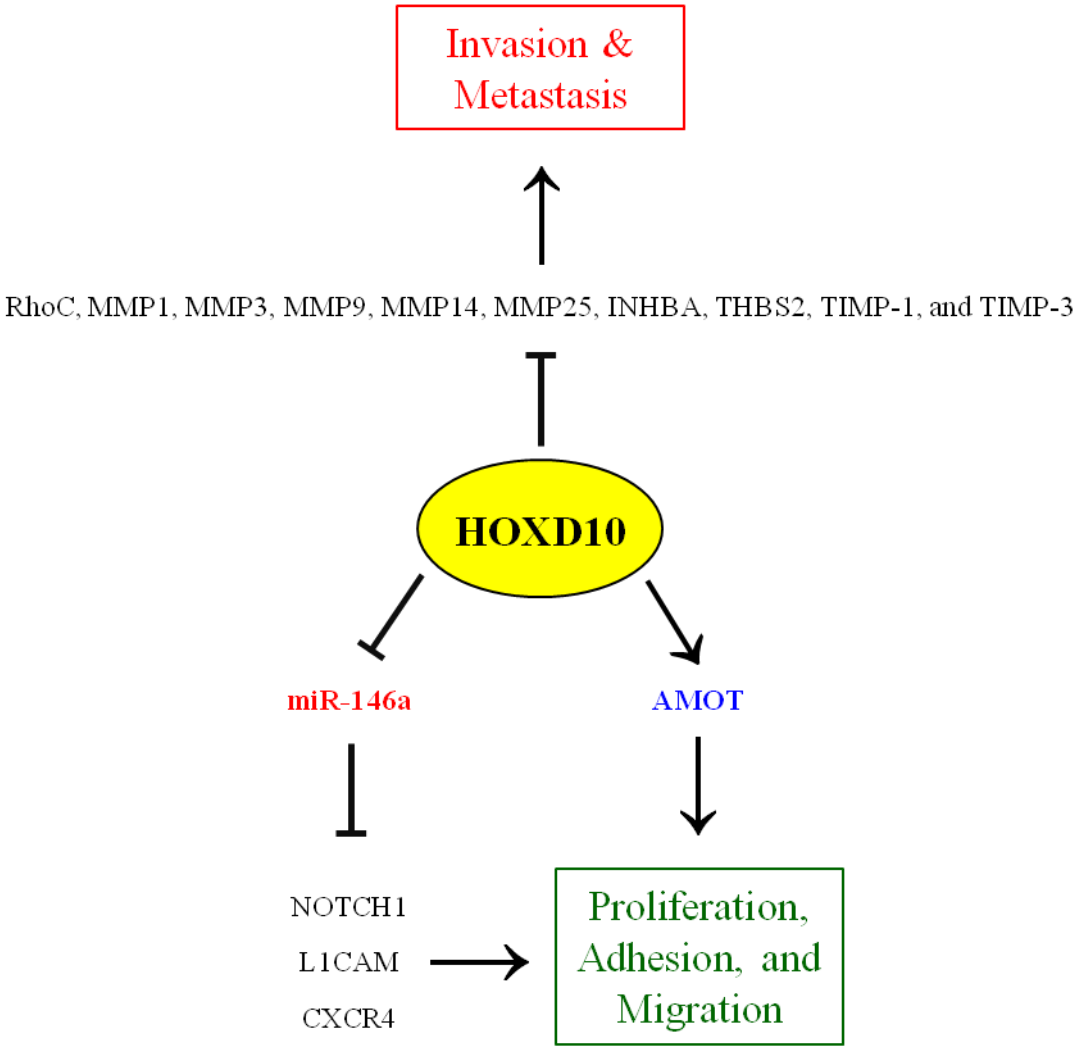


Figure 6.1: The identified HOXD10-associating molecules and their effect on cellular functions.

6.2 Future work

In a similar way as HNSCC, studying the expression of HOXD10 in NSCLC tissue samples might further confirm the high and low expression level in primary tumours and lymph node metastases, respectively.

Targeting HOX proteins complexes using HXR9 has been established in different types of cancers. However, the effect of this peptide on HOX proteins in HNSCC and NSCLC is still not studied. Therefore, the effect of HXR9 on HOXD10 complexes in these two cancers might be explored *in-vitro* and *in-vivo*. The *in-vivo* study might also include:

- introducing high level of HOXD10 siRNA to tackle NSCLC and HNSCC cells using one of the available *in-vivo* siRNA delivery systems including liposomal siRNA delivery system (Sakurai et al., 2013) or polysaccharide nanoparticles (Yang et al., 2013).
- investigating the potential role of HOXD10 in promoting HNSCC and NSCLC-associated angiogenesis mediated by its direct target AMOT-p80. Moreover, since AMOT-p80 is involved in modulating ERK1/2 signalling pathway, a potential activating role of HOXD10 might be explored.

To verify the miR-146a findings, the level of miR-146a in patients' tissue samples could be assessed using qPCR analysis and laser capture microdissection (LCM). In addition, the phenotypes associated with miR-146a over-expression or silencing can be investigated and compared to HOXD10 roles. Moreover, since the significance of *NOTCH1* gene in HNSCC is growing, and in our study a possible modulation of NOTCH1 level by HOXD10 was demonstrated, an interaction that involves HOXD10-miR-146a-NOTCH1 might be worth more exploring.

Finally, further investigation and DLR analysis might reveal more HOXD10-associating molecules among the remaining six significant genes. Studying their functions in depth should develop a more complete picture of the roles of HOXD10 in HNSCC and NSCLC. There is also room to re-evaluate the large list of HOXD10 putative associating molecules to identify the key invasion-associated genes and their associated pathways. This might help in designing a strategy for blocking both HOXD10, in order to inhibit tumour development caused by its over-expression, and key genes responsible for cell invasion, which function when HOXD10 is reduced but not completely lost. A good start in this direction might be studying the role of RhoC, INHBA, MMP3, MMP14, TIMP-1 and TIMP-3 proteins in inducing cell invasion following HOXD10 silencing.

CHAPTER 7: REFERENCES

ABATE-SHEN, C. 2002. Deregulated homeobox gene expression in cancer: cause or consequence? *Nat Rev Cancer,* 2**,** 777-85.

ABE, M., HAMADA, J., TAKAHASHI, O., TAKAHASHI, Y., TADA, M., MIYAMOTO, M., MORIKAWA, T., KONDO, S. & MORIUCHI, T. 2006. Disordered expression of HOX genes in human non-small cell lung cancer. *Oncol Rep,* 15**,** 797-802.

ABERLE, D. R., ADAMS, A. M., BERG, C. D., BLACK, W. C., CLAPP, J. D., FAGERSTROM, R. M., GAREEN, I. F., GATSONIS, C., MARCUS, P. M. & SICKS, J. D. 2011. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med,* 365**,** 395-409.

ABUZEID, W. M., LI, D. & O'MALLEY, B. W., JR. 2011. Gene therapy for head and neck cancer. *Adv Otorhinolaryngol,* 70**,** 141-51.

ADJAYE, J. & MONK, M. 2000. Transcription of homeobox-containing genes detected in cDNA libraries derived from human unfertilized oocytes and preimplantation embryos. *Mol Hum Reprod,* 6**,** 707-11.

AFFOLTER, M., PERCIVAL-SMITH, A., MÜLLER, M., BILLETER, M., QIAN, Y., OTTING, G., WÜTHRICH, K. & GEHRING, W. 1991. Similarities between the homeodomain and the Hin recombinase DNA-binding domain. *Cell,* 64**,** 879-880.

AGRAWAL, N., FREDERICK, M. J., PICKERING, C. R., BETTEGOWDA, C., CHANG, K., LI, R. J., FAKHRY, C., XIE, T. X., ZHANG, J., WANG, J., ZHANG, N., EL-NAGGAR, A. K., JASSER, S. A., WEINSTEIN, J. N., TREVINO, L., DRUMMOND, J. A., MUZNY, D. M., WU, Y., WOOD, L. D., HRUBAN, R. H., WESTRA, W. H., KOCH, W. M., CALIFANO, J. A., GIBBS, R. A., SIDRANSKY, D., VOGELSTEIN, B., VELCULESCU, V. E., PAPADOPOULOS, N., WHEELER, D. A., KINZLER, K. W. & MYERS, J. N. 2011. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science,* 333**,** 1154-7.

AGUR, Z., KOGAN, Y., LEVI, L., HARRISON, H., LAMB, R., KIRNASOVSKY, O. U. & CLARKE, R. B. 2010. Disruption of a Quorum Sensing mechanism triggers tumorigenesis: a simple discrete model corroborated by experiments in mammary cancer stem cells. *Biol Direct,* 5**,** 20.

AI, L., STEPHENSON, K. K., LING, W., ZUO, C., MUKUNYADZI, P., SUEN, J. Y., HANNA, E. & FAN, C. Y. 2003. The p16 (CDKN2a/INK4a) tumor-suppressor gene in head and neck squamous cell carcinoma: a promoter methylation and protein expression study in 100 cases. *Mod Pathol,* 16**,** 944-50.

AKAM, M. 1989. Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell,* 57**,** 347-9.

AKARSU, A. N., STOILOV, I., YILMAZ, E., SAYLI, B. S. & SARFARAZI, M. 1996. Genomic structure of HOXD13 gene: a nine polyalanine duplication causes synpolydactyly in two unrelated families. *Hum Mol Genet,* 5**,** 945-52.

AKIN, Z. & NAZARALI, A. 2005. Hox genes and their candidate downstream targets in the developing central nervous system. *Cellular and molecular neurobiology,* 25**,** 697-741.

ALASTI, F., SADEGHI, A., SANATI, M. H., FARHADI, M., STOLLAR, E., SOMERS, T. & VAN CAMP, G. 2008. A Mutation in< i> HOXA2</i> Is Responsible for Autosomal-Recessive Microtia in an Iranian Family. *The American Journal of Human Genetics,* 82**,** 982-991.

ALBERG, A. J., FORD, J. G. & SAMET, J. M. 2007. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest,* 132**,** 29S-55S.

ALHARBI, R. A., PETTENGELL, R., PANDHA, H. S. & MORGAN, R. 2013. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia,* 27**,** 1000-8.

ANBAZHAGAN, R. & RAMAN, V. 1997. Homeobox genes: molecular link between congenital anomalies and cancer. *European Journal of Cancer,* 33**,** 635-637.

ANDL, T., KAHN, T., PFUHL, A., NICOLA, T., ERBER, R., CONRADT, C., KLEIN, W., HELBIG, M., DIETZ, A., WEIDAUER, H. & BOSCH, F. X. 1998. Etiological

involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinomas lacking retinoblastoma cell cycle control. *Cancer Res,* 58**,** 5-13.

ANG, K. K., BERKEY, B. A., TU, X., ZHANG, H. Z., KATZ, R., HAMMOND, E. H., FU, K. K. & MILAS, L. 2002. Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma. *Cancer Res,* 62**,** 7350-6.

ANGULO, B., CONDE, E., SUAREZ-GAUTHIER, A., PLAZA, C., MARTINEZ, R., REDONDO, P., IZQUIERDO, E., RUBIO-VIQUEIRA, B., PAZ-ARES, L., HIDALGO, M. & LOPEZ-RIOS, F. 2012. A comparison of EGFR mutation testing methods in lung carcinoma: direct sequencing, real-time PCR and immunohistochemistry. *PLoS One,* 7**,** e43842.

ANSARY-MOGHADDAM, A., HUXLEY, R. R., LAM, T. H. & WOODWARD, M. 2009. The risk of upper aero digestive tract cancer associated with smoking, with and without concurrent alcohol consumption. *Mt Sinai J Med,* 76**,** 392-403.

ANTCZAK, A., MIGDALSKA-SĘK, M., PASTUSZAK-LEWANDOSKA, D., CZARNECKA, K., NAWROT, E., DOMAŃSKA, D., KORDIAK, J., GÓRSKI, P. & BRZEZIAŃSKA, E. 2013. Significant frequency of allelic imbalance in 3p region covering RARβ and MLH1 loci seems to be essential in molecular non-small cell lung cancer diagnosis. *Medical Oncology,* 30**,** 1-10.

APIOU, F., FLAGIELLO, D., CILLO, C., MALFOY, B., POUPON, M. F. & DUTRILLAUX, B. 1996. Fine mapping of human HOX gene clusters. *Cytogenet Cell Genet,* 73**,** 114-5.

ARORA, S., AGGARWAL, P., PATHAK, A., BHANDARI, R., DUFFOO, F. & GULATI, S. C. 2012. Molecular genetics of head and neck cancer (Review). *Mol Med Rep,* 6**,** 19-22.

ATSDR. 2010. *What is Radon?* [Online]. Atlanta: Agency for Toxic Substances and Disease Registry. Available: http://www.atsdr.cdc.gov/csem/radon/what.html [Accessed 24 April 2011].

AWGULEWITSCH, A. 2003. Hox in hair growth and development. *Naturwissenschaften,* 90**,** 193-211.

BACH, C., BUHL, S., MUELLER, D., GARCIA-CUELLAR, M. P., MAETHNER, E. & SLANY, R. K. 2010. Leukemogenic transformation by HOXA cluster genes. *Blood,* 115**,** 2910-8.

BACH, P. B. 2009. Smoking as a factor in causing lung cancer. *JAMA: The Journal of the American Medical Association,* 301**,** 539-541.

BADILLO-ALMARAZ, I., ZAPATA-BENAVIDES, P., SAAVEDRA-ALONSO, S., ZAMORA-AVILA, D., RESENDEZ-PEREZ, D., TAMEZ-GUERRA, R., HERRERA-ESPARZA, R. & RODRIGUEZ-PADILLA, C. 2013. Human papillomavirus 16/18 infections in lung cancer patients in Mexico. *Intervirology,* 56**,** 310-5.

BAEZ, A. 2008. Genetic and environmental factors in head and neck cancer genesis. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev,* 26**,** 174-200.

BAGNARDI, V., ROTA, M., BOTTERI, E., SCOTTI, L., JENAB, M., BELLOCCO, R., TRAMACERE, I., PELUCCHI, C., NEGRI, E., LA VECCHIA, C., CORRAO, G. & BOFFETTA, P. 2011. Alcohol consumption and lung cancer risk in never smokers: a meta-analysis. *Ann Oncol,* 22**,** 2631-9.

BAILEY-WILSON, J., AMOS, C., PINNEY, S., PETERSEN, G., DE ANDRADE, M., WIEST, J., FAIN, P., SCHWARTZ, A., YOU, M. & FRANKLIN, W. 2004. A major lung cancer susceptibility locus maps to chromosome 6q23-25. *The American Journal of Human Genetics,* 75**,** 460-474.

BALDERAS-LOAEZA, A., ANAYA-SAAVEDRA, G., RAMIREZ-AMADOR, V. A., GUIDO-JIMENEZ, M. C., KALANTARI, M., CALLEJA-MACIAS, I. E., BERNARD, H. U. & GARCIA-CARRANCA, A. 2007. Human papillomavirus-16 DNA methylation patterns support a causal association of the virus with oral squamous cell carcinomas. *Int J Cancer,* 120**,** 2165-9.

BALLAZ, S. & MULSHINE, J. L. 2003. The potential contributions of chronic inflammation to lung carcinogenesis. *Clin Lung Cancer,* 5**,** 46-62.

BANERJEE-BASU, S. & BAXEVANIS, A. D. 2001. Molecular evolution of the homeodomain family of transcription factors. *Nucleic Acids Res,* 29**,** 3258-69.

BANSAL, D., SCHOLL, C., FROHLING, S., MCDOWELL, E., LEE, B. H., DOHNER, K., ERNST, P., DAVIDSON, A. J., DALEY, G. Q., ZON, L. I., GILLILAND, D. G. & HUNTLY, B. J. 2006. Cdx4 dysregulates Hox gene expression and generates acute myeloid leukemia alone and in cooperation with Meis1a in a murine model. *Proc Natl Acad Sci U S A,* 103**,** 16924-9.

BARLOW, J. H., FARYABI, R. B., CALLÉN, E., WONG, N., MALHOWSKI, A., CHEN, H. T., GUTIERREZ-CRUZ, G., SUN, H.-W., MCKINNON, P. & WRIGHT, G. 2013. Identification of early replicating fragile sites that contribute to genome instability. *Cell*.

BARNES, L. 2005. *Pathology & genetics: head and neck tumours*, World Health Organization.

BARTEL, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell,* 116**,** 281-97.

BASKA, T., WARREN, C. W., BASKOVA, M. & JONES, N. R. 2009. Prevalence of youth cigarette smoking and selected social factors in 25 European countries: findings from the Global Youth Tobacco Survey. *Int J Public Health,* 54**,** 439-45.

BATESON, W. 1894. Materials for the study of variation, treated with especial regard to discontinuity in the origin of species. *London, Macmillan.*

BAUER, V. L., BRASELMANN, H., HENKE, M., MATTERN, D., WALCH, A., UNGER, K., BAUDIS, M., LASSMANN, S., HUBER, R. & WIENBERG, J. 2008. Chromosomal changes characterize head and neck cancer with poor prognosis. *Journal of molecular medicine,* 86**,** 1353-1365.

BECHTOLD, V., BEARD, P. & RAJ, K. 2003. Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA. *J Virol,* 77**,** 2021-8.

BECK, F. 2002. Homeobox genes in gut development. *Gut,* 51**,** 450-4.

BÉRGAMO, N. A., DA SILVA VEIGA, L. C., DOS REIS, P. P., NISHIMOTO, I. N., MAGRIN, J., KOWALSKI, L. P., SQUIRE, J. A. & ROGATTO, S. R. 2005. Classic and molecular cytogenetic analyses reveal chromosomal gains and losses correlated with survival in head and neck cancer patients. *Clinical cancer research,* 11**,** 621-631.

BERLIN, D. S., SANGKUHL, K., KLEIN, T. E. & ALTMAN, R. B. 2011. PharmGKB summary: cytochrome P450, family 2, subfamily J, polypeptide 2: CYP2J2. *Pharmacogenet Genomics,* 21**,** 308-11.

BHATTACHARJEE, A., RICHARDS, W. G., STAUNTON, J., LI, C., MONTI, S., VASA, P., LADD, C., BEHESHTI, J., BUENO, R., GILLETTE, M., LODA, M., WEBER, G., MARK, E. J., LANDER, E. S., WONG, W., JOHNSON, B. E., GOLUB, T. R., SUGARBAKER, D. J. & MEYERSON, M. 2001. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A,* 98**,** 13790-5.

BHAUMIK, D., SCOTT, G. K., SCHOKRPUR, S., PATIL, C. K., CAMPISI, J. & BENZ, C. C. 2008. Expression of microRNA-146 suppresses NF-kappaB activity with reduction of metastatic potential in breast cancer cells. *Oncogene,* 27**,** 5643-7.

BITU, C. C., DESTRO, M. F., CARRERA, M., DA SILVA, S. D., GRANER, E., KOWALSKI, L. P., SOARES, F. A. & COLETTA, R. D. 2012. HOXA1 is overexpressed in oral squamous cell carcinomas and its expression is correlated with poor prognosis. *BMC cancer,* 12**,** 146.

BOCKMÜHL, U., SCHLÜNS, K., KÜCHLER, I., PETERSEN, S. & PETERSEN, I. 2000. Genetic imbalances with impact on survival in head and neck cancer patients. *The American journal of pathology,* 157**,** 369-375.

BOEING, H., DIETRICH, T., HOFFMANN, K., PISCHON, T., FERRARI, P., LAHMANN, P. H., BOUTRON-RUAULT, M. C., CLAVEL-CHAPELON, F., ALLEN, N., KEY, T., SKEIE, G., LUND, E., OLSEN, A., TJONNELAND, A., OVERVAD, K., JENSEN, M. K., ROHRMANN, S., LINSEISEN, J., TRICHOPOULOU, A., BAMIA, C., PSALTOPOULOU, T., WEINEHALL, L., JOHANSSON, I., SANCHEZ, M. J., JAKSZYN, P., ARDANAZ, E., AMIANO, P., CHIRLAQUE, M. D., QUIROS, J. R., WIRFALT, E., BERGLUND, G., PEETERS, P. H., VAN GILS, C. H., BUENO-DE-MESQUITA, H. B., BUCHNER, F. L., BERRINO, F., PALLI, D., SACERDOTE, C., TUMINO, R., PANICO, S., BINGHAM, S., KHAW, K. T., SLIMANI, N., NORAT, T., JENAB, M. & RIBOLI, E. 2006. Intake of fruits and vegetables and risk of cancer of the upper aero-digestive tract: the prospective EPIC-study. *Cancer Causes Control,* 17**,** 957-69.

BOLDIN, M. P., TAGANOV, K. D., RAO, D. S., YANG, L., ZHAO, J. L., KALWANI, M., GARCIA-FLORES, Y., LUONG, M., DEVREKANLI, A. & XU, J. 2011. miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *The Journal of experimental medicine,* 208**,** 1189-1201.

BONCINELLI, E., SIMEONE, A., ACAMPORA, D. & MAVILIO, F. 1991. HOX gene activation by retinoic acid. *Trends Genet,* 7**,** 329-34.

BONNER, J. A., HARARI, P. M., GIRALT, J., AZARNIA, N., SHIN, D. M., COHEN, R. B., JONES, C. U., SUR, R., RABEN, D., JASSEM, J., OVE, R., KIES, M. S., BASELGA, J., YOUSSOUFIAN, H., AMELLAL, N., ROWINSKY, E. K. & ANG, K. K. 2006. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N Engl J Med,* 354**,** 567-78.

BOSLEY, T. M., ALORAINY, I. A., SALIH, M. A., ALDHALAAN, H. M., ABU-AMERO, K. K., OYSTRECK, D. T., TISCHFIELD, M. A., ENGLE, E. C. & ERICKSON, R. P. 2008. The clinical spectrum of homozygous HOXA1 mutations. *Am J Med Genet A,* 146A**,** 1235-40.

BOUDREAU, M. & COUSSENS, L. unpublished work. *Macrophage-Mediated Delivery of the Breast Tumor Suppressor HoxD10 via Autologous Transfer to Breast Tumors* [Online]. Available: http://cancer.ucsf.edu/research/breast-spore/breast-cancer-spore-developmental-research-project-3 [Accessed 29/11/2013.

BOURGUIGNON, L. Y., WONG, G., EARLE, C., KRUEGER, K. & SPEVAK, C. C. 2010. Hyaluronan-CD44 interaction promotes c-Src-mediated twist signaling, microRNA-10b expression, and RhoA/RhoC up-regulation, leading to Rho-kinase-associated cytoskeleton activation and breast tumor cell invasion. *J Biol Chem,* 285**,** 36721-35.

BOVA, R. J., QUINN, D. I., NANKERVIS, J. S., COLE, I. E., SHERIDAN, B. F., JENSEN, M. J., MORGAN, G. J., HUGHES, C. J. & SUTHERLAND, R. L. 1999. Cyclin D1 and p16INK4A expression predict reduced survival in carcinoma of the anterior tongue. *Clin Cancer Res,* 5**,** 2810-9.

BOY, S., VAN RENSBURG, E. J., ENGELBRECHT, S., DREYER, L., VAN HEERDEN, M. & VAN HEERDEN, W. 2006. HPV detection in primary intra-oral squamous cell carcinomas--commensal, aetiological agent or contamination? *J Oral Pathol Med,* 35**,** 86-90.

BOYER, S. N., WAZER, D. E. & BAND, V. 1996. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res,* 56**,** 4620-4.

BOYLE, J. O., HAKIM, J., KOCH, W., VAN DER RIET, P., HRUBAN, R. H., ROA, R. A., CORREO, R., EBY, Y. J., RUPPERT, J. M. & SIDRANSKY, D. 1993. The incidence of p53 mutations increases with progression of head and neck cancer. *Cancer Res,* 53**,** 4477-80.

BRATT, A., WILSON, W. J., TROYANOVSKY, B., AASE, K., KESSLER, R., VAN MEIR, E. G. & HOLMGREN, L. 2002. Angiomotin belongs to a novel protein family with conserved coiled-coil and PDZ binding domains. *Gene,* 298**,** 69-77.

BRAZIL, D. P., YANG, Z.-Z. & HEMMINGS, B. A. 2004. Advances in protein kinase B signalling:< i> AKT</i> ion on multiple fronts. *Trends in biochemical sciences,* 29**,** 233-242.

BRENNAN, J. A. & SIDRANSKY, D. 1996. Molecular staging of head and neck squamous carcinoma. *Cancer Metastasis Rev,* 15**,** 3-10.

BRENNER, D. R., BRENNAN, P., BOFFETTA, P., AMOS, C. I., SPITZ, M. R., CHEN, C., GOODMAN, G., HEINRICH, J., BICKEBOLLER, H., ROSENBERGER, A., RISCH, A., MULEY, T., MCLAUGHLIN, J. R., BENHAMOU, S., BOUCHARDY, C., LEWINGER, J. P., WITTE, J. S., CHEN, G., BULL, S. & HUNG, R. J. 2013. Hierarchical modeling identifies novel lung cancer susceptibility variants in inflammation pathways among 10,140 cases and 11,012 controls. *Hum Genet,* 132**,** 579-89.

BRETZ, N. P., SALNIKOV, A. V., PERNE, C., KELLER, S., WANG, X., MIERKE, C. T., FOGEL, M., ERBE-HOFMANN, N., SCHLANGE, T., MOLDENHAUER, G. & ALTEVOGT, P. 2012. CD24 controls Src/STAT3 activity in human tumors. *Cell Mol Life Sci,* 69**,** 3863-79.

BRIDGES, C. B. & MORGAN, T. H. 1923. *The third-chromosome group of mutant characters of Drosophila melanogaster*, Carnegie Institution of Washington.

BROCK, H. W., HODGSON, J. W., PETRUK, S. & MAZO, A. 2009. Regulatory noncoding RNAs at Hox loci. *Biochemistry and Cell Biology,* 87**,** 27-34.

BRZEZIANSKA, E., DUTKOWSKA, A. & ANTCZAK, A. 2013. The significance of epigenetic alterations in lung carcinogenesis. *Mol Biol Rep,* 40**,** 309-25.

BRZEZNIAK, C., CARTER, C. A. & GIACCONE, G. 2013. Dacomitinib, a new therapy for the treatment of non-small cell lung cancer. *Expert Opin Pharmacother,* 14**,** 247-53.

BUTASH, K. A., NATARAJAN, P., YOUNG, A. & FOX, D. K. 2000. Reexamination of the effect of endotoxin on cell proliferation and transfection efficiency. *Biotechniques,* 29**,** 610-4, 616, 618-9.

BUTLER, J. E. & KADONAGA, J. T. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & development,* 16**,** 2583-2592.

BYKOV, V. J., ISSAEVA, N., SHILOV, A., HULTCRANTZ, M., PUGACHEVA, E., CHUMAKOV, P., BERGMAN, J., WIMAN, K. G. & SELIVANOVA, G. 2002. Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nat Med,* 8**,** 282-8.

CALIFANO, J., VAN DER RIET, P., WESTRA, W., NAWROZ, H., CLAYMAN, G., PIANTADOSI, S., CORIO, R., LEE, D., GREENBERG, B., KOCH, W. & SIDRANSKY, D. 1996. Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res,* 56**,** 2488-92.

CALLAGHAN, R. C., ALLEBECK, P. & SIDORCHUK, A. 2013. Marijuana use and risk of lung cancer: a 40-year cohort study. *Cancer Causes & Control***,** 1-10.

CALVO, R., WEST, J., FRANKLIN, W., ERICKSON, P., BEMIS, L., LI, E., HELFRICH, B., BUNN, P., ROCHE, J., BRAMBILLA, E., ROSELL, R., GEMMILL, R. M. & DRABKIN, H. A. 2000. Altered HOX and WNT7A expression in human lung cancer. *Proc Natl Acad Sci U S A,* 97**,** 12776-81.

CAMPISI, G., PANZARELLA, V., GIULIANI, M., LAJOLO, C., DI FEDE, O., FALASCHINI, S., DI LIBERTO, C., SCULLY, C. & LO MUZIO, L. 2007. Human papillomavirus: its identity and controversial role in oral oncogenesis, premalignant and malignant lesions (review). *Int J Oncol,* 30**,** 813-23.

CANBAZ, D., KIRIMTAY, K., KARACA, E. & KARABAY, A. 2011. SPG4 gene promoter regulation via Elk1 transcription factor. *J Neurochem,* 117**,** 724-34.

CANCERRESEARCHUK. 2010. *Cancer Research UK: the UK's leading cancer charity* [Online]. Available: http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/commoncancers/ [Accessed 12 June 2013].

CAPPUZZO, F., HIRSCH, F. R., ROSSI, E., BARTOLINI, S., CERESOLI, G. L., BEMIS, L., HANEY, J., WITTA, S., DANENBERG, K., DOMENICHINI, I., LUDOVINI, V., MAGRINI, E., GREGORC, V., DOGLIONI, C., SIDONI, A., TONATO, M., FRANKLIN, W. A., CRINO, L., BUNN, P. A., JR. & VARELLA-GARCIA, M. 2005. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst,* 97**,** 643-55.

CARE, A., SILVANI, A., MECCIA, E., MATTIA, G., PESCHLE, C. & COLOMBO, M. P. 1998. Transduction of the SkBr3 breast carcinoma cell line with the HOXB7 gene induces bFGF expression, increases cell proliferation and reduces growth factor dependence. *Oncogene,* 16**,** 3285-9.

CARNEY, D. N., GAZDAR, A. F., BEPLER, G., GUCCION, J. G., MARANGOS, P. J., MOODY, T. W., ZWEIG, M. H. & MINNA, J. D. 1985. Establishment and identification of small cell lung cancer cell lines having classic and variant features. *Cancer Res,* 45**,** 2913-23.

CARPENTER, E. M., GODDARD, J. M., DAVIS, A. P., NGUYEN, T. P. & CAPECCHI, M. R. 1997. Targeted disruption of Hoxd-10 affects mouse hindlimb development. *Development,* 124**,** 4505-14.

CARRIO, M., ARDERIU, G., MYERS, C. & BOUDREAU, N. J. 2005. Homeobox D10 induces phenotypic reversion of breast tumor cells in a three-dimensional culture model. *Cancer Res,* 65**,** 7177-85.

CARTHARIUS, K., FRECH, K., GROTE, K., KLOCKE, B., HALTMEIER, M., KLINGENHOFF, A., FRISCH, M., BAYERLEIN, M. & WERNER, T. 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics,* 21**,** 2933-42.

CARVALHO, A. L., CHUANG, A., JIANG, W.-W., LEE, J., BEGUM, S., POETA, L., ZHAO, M., JERÓNIMO, C., HENRIQUE, R. & NAYAK, C. S. 2006. Deleted in colorectal cancer is a putative conditional tumor-suppressor gene inactivated by promoter hypermethylation in head and neck squamous cell carcinoma. *Cancer research,* 66**,** 9401-9407.

CERVIGNE, N. K., REIS, P. P., MACHADO, J., SADIKOVIC, B., BRADLEY, G., GALLONI, N. N., PINTILIE, M., JURISICA, I., PEREZ-ORDONEZ, B., GILBERT, R., GULLANE, P., IRISH, J. & KAMEL-REID, S. 2009. Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma. *Hum Mol Genet,* 18**,** 4818-29.

CHALBOS, D., VIGNON, F., KEYDAR, I. & ROCHEFORT, H. 1982. Estrogens stimulate cell proliferation and induce secretory proteins in a human breast cancer cell line (T47D). *J Clin Endocrinol Metab,* 55**,** 276-83.

CHANG, A. 2011. Chemotherapy, chemoresistance and the changing treatment landscape for NSCLC. *Lung Cancer,* 71**,** 3-10.

CHANG, C. P., SHEN, W. F., ROZENFELD, S., LAWRENCE, H. J., LARGMAN, C. & CLEARY, M. L. 1995. Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev,* 9**,** 663-74.

CHANG, S. S., JIANG, W. W., SMITH, I., POETA, L. M., BEGUM, S., GLAZER, C., SHAN, S., WESTRA, W., SIDRANSKY, D. & CALIFANO, J. A. 2008. MicroRNA alterations in head and neck squamous cell carcinoma. *Int J Cancer,* 123**,** 2791-7.

CHAPMAN, C. J., MURRAY, A., MCELVEEN, J. E., SAHIN, U., LUXEMBURGER, U., TÜRECI, Ö., WIEWRODT, R., BARNES, A. C. & ROBERTSON, J. 2008. Autoantibodies in lung cancer: possibilities for early detection and subsequent cure. *Thorax,* 63**,** 228-233.

CHARIOT, A. & CASTRONOVO, V. 1996. Detection of HOXA1 expression in human breast cancer. *Biochem Biophys Res Commun,* 222**,** 292-7.

CHARIOT, A., MOREAU, L., SENTERRE, G., SOBEL, M. E. & CASTRONOVO, V. 1995. Retinoic acid induces three newly cloned HOXA1 transcripts in MCF7 breast cancer cells. *Biochem Biophys Res Commun,* 215**,** 713-20.

CHATURVEDI, A. K., ENGELS, E. A., ANDERSON, W. F. & GILLISON, M. L. 2008. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol,* 26**,** 612-9.

CHAUDHARY, A. K., SINGH, M., SUNDARAM, S. & MEHROTRA, R. 2009. Role of human papillomavirus and its detection in potentially malignant and malignant head and neck lesions: updated review. *Head Neck Oncol,* 1**,** 22.

CHEN, C., WEI, X., RAO, X., WU, J., YANG, S., CHEN, F., MA, D., ZHOU, J., DACKOR, R. T. & ZELDIN, D. C. 2011. Cytochrome P450 2J2 is highly expressed in hematologic malignant diseases and promotes tumor cell growth. *Journal of Pharmacology and Experimental Therapeutics,* 336**,** 344-355.

CHEN, F. & CAPECCHI, M. R. 1999. Paralogous mouse Hox genes, Hoxa9, Hoxb9, and Hoxd9, function together to control development of the mammary gland in response to pregnancy. *Proceedings of the National Academy of Sciences,* 96**,** 541-546.

CHEN, G., UMELO, I. A., LV, S., TEUGELS, E., FOSTIER, K., KRONENBERGER, P., DEWAELE, A., SADONES, J., GEERS, C. & DE GREVE, J. 2013a. miR-146a inhibits cell growth, cell migration and induces apoptosis in non-small cell lung cancer cells. *PLoS One,* 8**,** e60317.

CHEN, J., BI, H., HOU, J., ZHANG, X., ZHANG, C., YUE, L., WEN, X., LIU, D., SHI, H. & YUAN, J. 2013b. Atorvastatin overcomes gefitinib resistance in KRAS mutant human non-small cell lung carcinoma cells. *Cell death & disease,* 4**,** e814.

CHEN, J., WU, X., LIN, J. & LEVINE, A. J. 1996. mdm-2 inhibits the G1 arrest and apoptosis functions of the p53 tumor suppressor protein. *Mol Cell Biol,* 16**,** 2445-2452.

CHEN, T. J., GAO, F., YANG, T., THAKUR, A., REN, H., LI, Y., ZHANG, S., WANG, T. & CHEN, M. W. 2013c. CDK-associated Cullin 1 promotes cell proliferation with activation of ERK1/2 in human lung cancer A549 cells. *Biochem Biophys Res Commun,* 437**,** 108-13.

CHENG, Y.-W., CHIOU, H.-L., SHEU, G.-T., HSIEH, L.-L., CHEN, J.-T., CHEN, C.-Y., SU, J.-M. & LEE, H. 2001. The association of human papillomavirus 16/18 infection with lung cancer among nonsmoking Taiwanese women. *Cancer research,* 61**,** 2799-2803.

CHISAKA, O., MUSCI, T. S. & CAPECCHI, M. R. 1992. Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox geneHox-# 150; 1.6. *Nature,* 355**,** 516-520.

CHOU, Y.-T., LIN, H.-H., LIEN, Y.-C., WANG, Y.-H., HONG, C.-F., KAO, Y.-R., LIN, S.-C., CHANG, Y.-C., LIN, S.-Y. & CHEN, S.-J. 2010. EGFR promotes lung tumorigenesis by activating miR-7 through a Ras/ERK/Myc pathway that targets the Ets2 transcriptional repressor ERF. *Cancer research,* 70**,** 8822-8831.

CHRISTIANI, D. C. 2009. Lung cancer genetics: a family affair? *Clinical cancer research,* 15**,** 2581-2582.

CHRISTY, A. W. & BOJAN, A. 2013. Targeted therapy: A novel approach in head and neck cancer. *Indian J Dent Res,* 24**,** 261-6.

CILLO, C., CANTILE, M., FAIELLA, A. & BONCINELLI, E. 2001. Homeobox genes in normal and malignant cells. *J Cell Physiol,* 188**,** 161-9.

CIMMINO, A., CALIN, G. A., FABBRI, M., IORIO, M. V., FERRACIN, M., SHIMIZU, M., WOJCIK, S. E., AQEILAN, R. I., ZUPO, S. & DONO, M. 2005. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 13944-13949.

CONWAY, D. I., HASHIBE, M., BOFFETTA, P., WUNSCH-FILHO, V., MUSCAT, J., LA VECCHIA, C. & WINN, D. M. 2009. Enhancing epidemiologic research on head and neck cancer: INHANCE - The international head and neck cancer epidemiology consortium. *Oral Oncol,* 45**,** 743-6.

COOPER, J. S., PORTER, K., MALLIN, K., HOFFMAN, H. T., WEBER, R. S., ANG, K. K., GAY, E. G. & LANGER, C. J. 2009. National Cancer Database report on cancer of the head and neck: 10-year update. *Head Neck,* 31**,** 748-58.

COOPER, W. A., LAM, D. C., O'TOOLE, S. A. & MINNA, J. D. 2013. Molecular biology of lung cancer. *J Thorac Dis,* 5**,** S479-S490.

COSTA, D. B., HALMOS, B., KUMAR, A., SCHUMER, S. T., HUBERMAN, M. S., BOGGON, T. J., TENEN, D. G. & KOBAYASHI, S. 2007. BIM mediates EGFR tyrosine kinase inhibitor-induced apoptosis in lung cancers with oncogenic EGFR mutations. *PLoS Med,* 4**,** 1669-79; discussion 1680.

CUGELL, D. W. & KAMP, D. W. 2004. Asbestos and the pleura: a review. *Chest,* 125**,** 1103-17.

CUI, X. & CHURCHILL, G. A. 2003. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol,* 4**,** 210.

D'ANGELO, S. P., PIETANZA, M. C., JOHNSON, M. L., RIELY, G. J., MILLER, V. A., SIMA, C. S., ZAKOWSKI, M. F., RUSCH, V. W., LADANYI, M. & KRIS, M. G. 2011. Incidence of EGFR exon 19 deletions and L858R in tumor specimens from men and cigarette smokers with lung adenocarcinomas. *J Clin Oncol,* 29**,** 2066-70.

DAFTARY, G. S. & TAYLOR, H. S. 2006. Endocrine regulation of HOX genes. *Endocr Rev,* 27**,** 331-55.

DANAHEY, D. G., TOBIN, E. J., SCHULLER, D. E., BIER-LANING, C. M., WEGHORST, C. M. & LANG, J. C. 1999. p16 mutation frequency and clinical correlation in head and neck cancer. *Acta Otolaryngol,* 119**,** 285-8.

DANIELS, T., NEACATO, I., RODRIGUEZ, J., PANDHA, H., MORGAN, R. & PENICHET, M. 2010. Disruption of HOX activity leads to cell death that can be enhanced by the interference of iron uptake in malignant B cells. *Leukemia,* 24**,** 1555-1565.

DAPSON, R. W. 2007. Macromolecular changes caused by formalin fixation and antigen retrieval. *Biotech Histochem,* 82**,** 133-40.

DARBY, S., HILL, D., AUVINEN, A., BARROS-DIOS, J. M., BAYSSON, H., BOCHICCHIO, F., DEO, H., FALK, R., FORASTIERE, F., HAKAMA, M., HEID, I., KREIENBROCK, L., KREUZER, M., LAGARDE, F., MAKELAINEN, I., MUIRHEAD, C., OBERAIGNER, W., PERSHAGEN, G., RUANO-RAVINA, A., RUOSTEENOJA, E., ROSARIO, A. S., TIRMARCHE, M., TOMASEK, L., WHITLEY, E., WICHMANN, H. E. & DOLL, R. 2005. Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. *BMJ,* 330**,** 223.

DARNTON, A. J., MCELVENNY, D. M. & HODGSON, J. T. 2006. Estimating the number of asbestos-related lung cancer deaths in Great Britain from 1980 to 2000. *Ann Occup Hyg,* 50**,** 29-38.

DE LUCA, A. & NORMANNO, N. 2010. Predictive biomarkers to tyrosine kinase inhibitors for the epidermal growth factor receptor in non-small-cell lung cancer. *Curr Drug Targets,* 11**,** 851-64.

DE SOUZA SETUBAL DESTRO, M. F., BITU, C. C., ZECCHIN, K. G., GRANER, E., LOPES, M. A., KOWALSKI, L. P. & COLETTA, R. D. 2010. Overexpression of HOXB7 homeobox gene in oral cancer induces cellular proliferation and is associated with poor prognosis. *Int J Oncol,* 36**,** 141-9.

DE VILLIERS, E. M., WEIDAUER, H., OTTO, H. & ZUR HAUSEN, H. 1985. Papillomavirus DNA in human tongue carcinomas. *International journal of cancer,* 36**,** 575-578.

DEARDEN, S., STEVENS, J., WU, Y. L. & BLOWERS, D. 2013. Mutation incidence and coincidence in non small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap). *Ann Oncol,* 24**,** 2371-6.

DELILBASI, C. B., OKURA, M., IIDA, S. & KOGO, M. 2004. Investigation of CXCR4 in squamous cell carcinoma of the tongue. *Oral Oncol,* 40**,** 154-7.

DEMIRHAN, O., TASTEMIR, D., HASTURK, S., KULECI, S. & HANTA, I. 2010. Alterations in p16 and p53 genes and chromosomal findings in patients with lung cancer: fluorescence in situ hybridization and cytogenetic studies. *Cancer Epidemiol,* 34**,** 472-7.

DEMOKAN, S., CHUANG, A., SUOGLU, Y., ULUSAN, M., YALNIZ, Z., CALIFANO, J. A. & DALAY, N. 2012. Promoter methylation and loss of p16(INK4a) gene expression in head and neck cancer. *Head Neck,* 34**,** 1470-5.

DENNIS, G., JR., SHERMAN, B. T., HOSACK, D. A., YANG, J., GAO, W., LANE, H. C. & LEMPICKI, R. A. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol,* 4**,** P3.

DENOIX, P. 1944. Tumor, node and metastasis (TNM). *Bull. Inst. Nat. Hyg.(Paris),* 1**,** 1-69.

DEUTSCH, J. S. 2010. Hox genes: studies from the 20th to the 21st century. Preface. *Adv Exp Med Biol,* 689**,** ix-xi.

DICKSON, M. A., HAHN, W. C., INO, Y., RONFARD, V., WU, J. Y., WEINBERG, R. A., LOUIS, D. N., LI, F. P. & RHEINWALD, J. G. 2000. Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span become immortal yet retain normal growth and differentiation characteristics. *Mol Cell Biol,* 20**,** 1436-47.

DOBBS, M. B., GURNETT, C. A., PIERCE, B., EXNER, G. U., ROBARGE, J., MORCUENDE, J. A., COLE, W. G., TEMPLETON, P. A., FOSTER, B. & BOWCOCK, A. M. 2006. HOXD10 M319K mutation in a family with isolated congenital vertical talus. *J Orthop Res,* 24**,** 448-53.

DOLL, R., PETO, R., BOREHAM, J. & SUTHERLAND, I. 2005. Mortality from cancer in relation to smoking: 50 years observations on British doctors. *Br J Cancer,* 92**,** 426-9.

DOLLE, P., IZPISUA-BELMONTE, J. C., BROWN, J. M., TICKLE, C. & DUBOULE, D. 1991. HOX-4 genes and the morphogenesis of mammalian genitalia. *Genes Dev,* 5**,** 1767-7.

DORN, A., AFFOLTER, M., GEHRING, W. J. & LEUPIN, W. 1994. Homeodomain proteins in development and therapy. *Pharmacol Ther,* 61**,** 155-84.

DOWNWARD, J. 2003. Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer,* 3**,** 11-22.

DOZYNKIEWICZ, M. A., JAMIESON, N. B., MACPHERSON, I., GRINDLAY, J., VAN DEN BERGHE, P. V., VON THUN, A., MORTON, J. P., GOURLEY, C., TIMPSON, P. & NIXON, C. 2012. Rab25 and CLIC3 collaborate to promote integrin recycling from late endosomes/lysosomes and drive cancer progression. *Developmental cell,* 22**,** 131-145.

DU, H., DAFTARY, G. S., LALWANI, S. I. & TAYLOR, H. S. 2005. Direct regulation of HOXA10 by 1, 25-(OH) 2D3 in human myelomonocytic cells and human endometrial stromal cells. *Molecular Endocrinology,* 19**,** 2222-2233.

DU, H. & TAYLOR, H. S. 2004. Molecular regulation of mullerian development by Hox genes. *Ann N Y Acad Sci,* 1034**,** 152-65.

DUBOULE, D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Dev Suppl***,** 135-42.

DUBOULE, D. 1995a. Guidebook to the homeobox genes.

DUBOULE, D. 1995b. Vertebrate Hox genes and proliferation: an alternative pathway to homeosis? *Curr Opin Genet Dev,* 5**,** 525-8.

DUBOULE, D. & MORATA, G. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends in Genetics,* 10**,** 358-364.

DUESTER, G. 2008. Retinoic acid synthesis and signaling during early organogenesis. *Cell,* 134**,** 921-31.

DURISETI, S., WINNARD, P. T., JR., MIRONCHIK, Y., VESUNA, F., RAMAN, A. & RAMAN, V. 2006. HOXA5 regulates hMLH1 expression in breast cancer cells. *Neoplasia,* 8**,** 250-8.

ECONOMIDES, K. D. & CAPECCHI, M. R. 2003. Hoxb13 is required for normal differentiation and secretory function of the ventral prostate. *Development,* 130**,** 2061-2069.

EDGE, S. B. & COMPTON, C. C. 2010. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol,* 17**,** 1471-4.

EDINGTON, K. G., LOUGHRAN, O. P., BERRY, I. J. & PARKINSON, E. K. 1995. Cellular immortality: a late event in the progression of human squamous cell carcinoma of the head and neck associated with p53 alteration and a high frequency of allele loss. *Mol Carcinog,* 13**,** 254-65.

EGLOFF, A. M. & GRANDIS, J. R. 2008. Targeting epidermal growth factor receptor and SRC pathways in head and neck cancer. *Semin Oncol,* 35**,** 286-97.

EKLUND, E. A. 2007. The role of HOX genes in malignant myeloid disease. *Curr Opin Hematol,* 14**,** 85-9.

ENGELS, E. A. 2008. Inflammation in the development of lung cancer: epidemiological evidence. *Expert Rev Anticancer Ther,* 8**,** 605-15.

ERNKVIST, M., AASE, K., UKOMADU, C., WOHLSCHLEGEL, J., BLACKMAN, R., VEITONMAKI, N., BRATT, A., DUTTA, A. & HOLMGREN, L. 2006. p130-angiomotin associates to actin and controls endothelial cell shape. *FEBS J,* 273**,** 2000-11.

FARHAT, F. S. & HOUHOU, W. 2013. Targeted therapies in non-small cell lung carcinoma: what have we achieved so far? *Therapeutic Advances in Medical Oncology*.

FARLEY, B. M. & RYDER, S. P. 2008. Regulation of maternal mRNAs in early development. *Crit Rev Biochem Mol Biol,* 43**,** 135-62.

FARR, A. & ROMAN, A. 1992. A pitfall of using a second plasmid to determine transfection efficiency. *Nucleic acids research,* 20**,** 920.

FARRELL, R. E. 1998. *RNA methodologies*, Wiley Online Library.

FARROW, D. C., VAUGHAN, T. L., BERWICK, M., LYNCH, C. F., SWANSON, G. M. & LYON, J. L. 1998. Diet and nasopharyngeal cancer in a low-risk population. *Int J Cancer,* 78**,** 675-9.

FERLAY, J., SHIN, H. R., BRAY, F., FORMAN, D., MATHERS, C. & PARKIN, D. M. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer,* 127**,** 2893-917.

FIELD, J. K., OUDKERK, M., PEDERSEN, J. H. & DUFFY, S. W. 2013. Prospects for population screening and diagnosis of lung cancer. *Lancet,* 382**,** 732-41.

FIELD, R. W. & WITHERS, B. L. 2012. Occupational and environmental causes of lung cancer. *Clin Chest Med,* 33**,** 681-703.

FLICEK, P., AHMED, I., AMODE, M. R., BARRELL, D., BEAL, K., BRENT, S., CARVALHO-SILVA, D., CLAPHAM, P., COATES, G., FAIRLEY, S., FITZGERALD, S., GIL, L., GARCÍA-GIRÓN, C., GORDON, L., HOURLIER, T., HUNT, S., JUETTEMANN, T., KÄHÄRI, A. K., KEENAN, S., KOMOROWSKA, M., KULESHA, E., LONGDEN, I., MAUREL, T., MCLAREN, W. M., MUFFATO, M., NAG, R., OVERDUIN, B., PIGNATELLI, M., PRITCHARD, B., PRITCHARD, E., RIAT, H. S., RITCHIE, G. R. S., RUFFIER, M., SCHUSTER, M., SHEPPARD, D., SOBRAL, D., TAYLOR, K., THORMANN, A., TREVANION, S., WHITE, S., WILDER, S. P., AKEN, B. L., BIRNEY, E., CUNNINGHAM, F., DUNHAM, I., HARROW, J., HERRERO, J., HUBBARD, T. J. P., JOHNSON, N., KINSELLA, R., PARKER, A., SPUDICH, G., YATES, A., ZADISSA, A. & SEARLE, S. M. J. 2013. Ensembl 2013. *Nucleic Acids Research,* 41**,** D48-D55.

FORASTIERE, A., KOCH, W., TROTTI, A. & SIDRANSKY, D. 2001. Head and neck cancer. *N Engl J Med,* 345**,** 1890-900.

FORLANI, S., LAWSON, K. A. & DESCHAMPS, J. 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development,* 130**,** 3807-19.

FOWLER, A. M., SOLODIN, N., PREISLER-MASHEK, M. T., ZHANG, P., LEE, A. V. & ALARID, E. T. 2004. Increases in estrogen receptor-α concentration in breast cancer cells promote serine 118/104/106-independent AF-1 transactivation and growth in the absence of estrogen. *The FASEB journal,* 18**,** 81-93.

FOWLER, C. B., EVERS, D. L., O'LEARY, T. J. & MASON, J. T. 2011. Antigen retrieval causes protein unfolding: evidence for a linear epitope model of recovered immunoreactivity. *J Histochem Cytochem,* 59**,** 366-81.

FOX, E. M., ANDRADE, J. & SHUPNIK, M. A. 2009. Novel actions of estrogen to promote proliferation: integration of cytoplasmic and nuclear pathways. *Steroids,* 74**,** 622-7.

FREEDMAN, N. D., PARK, Y., SUBAR, A. F., HOLLENBECK, A. R., LEITZMANN, M. F., SCHATZKIN, A. & ABNET, C. C. 2008. Fruit and vegetable intake and head and neck cancer risk in a large United States prospective cohort study. *Int J Cancer,* 122**,** 2330-6.

FRIEDL, P. & WOLF, K. 2003. Proteolytic and non-proteolytic migration of tumour cells and leucocytes. *Biochem Soc Symp***,** 277-85.

FROMENTAL-RAMAIN, C., WAROT, X., LAKKARAJU, S., FAVIER, B., HAACK, H., BIRLING, C., DIERICH, A., DOLL E, P. & CHAMBON, P. 1996. Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. *Development,* 122**,** 461-72.

FU, Y., LI, F., ZHAO, D. Y., ZHANG, J.-S., LV, Y. & LI-LING, J. 2012. Interaction between Tbx1 and HoxD10 and connection with TGFβ-BMP signal pathway during kidney development. *Gene*.

FUKUI, T., OHE, Y., TSUTA, K., FURUTA, K., SAKAMOTO, H., TAKANO, T., NOKIHARA, H., YAMAMOTO, N., SEKINE, I. & KUNITOH, H. 2008. Prospective Study of the Accuracy of EGFR Mutational Analysis by High-Resolution Melting Analysis in Small Samples Obtained from Patients with Non–Small Cell Lung Cancer. *Clinical cancer research,* 14**,** 4751-4757.

FUNG, C. & GRANDIS, J. R. 2010. Emerging drugs to treat squamous cell carcinomas of the head and neck. *Expert opinion on emerging drugs,* 15**,** 355-373.

FUSCO, A. & FEDELE, M. 2007. Roles of HMGA proteins in cancer. *Nat Rev Cancer,* 7**,** 899-910.

GARBER, M. E., TROYANSKAYA, O. G., SCHLUENS, K., PETERSEN, S., THAESLER, Z., PACYNA-GENGELBACH, M., VAN DE RIJN, M., ROSEN, G. D., PEROU, C. M., WHYTE, R. I., ALTMAN, R. B., BROWN, P. O., BOTSTEIN, D. & PETERSEN, I. 2001. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A,* 98**,** 13784-9.

GARCIA, A. I., BUISSON, M., BERTRAND, P., RIMOKH, R., ROULEAU, E., LOPEZ, B. S., LIDEREAU, R., MIKAELIAN, I. & MAZOYER, S. 2011. Down-regulation of BRCA1 expression by miR-146a and miR-146b-5p in triple negative sporadic breast cancers. *EMBO Mol Med,* 3**,** 279-90.

GAUFO, G. O., THOMAS, K. R. & CAPECCHI, M. R. 2003. Hox3 genes coordinate mechanisms of genetic suppression and activation in the generation of branchial and somatic motoneurons. *Development,* 130**,** 5191-5201.

GEHRING, W. J. 1987. Homeo boxes in the study of development. *Science,* 236**,** 1245-52.

GEHRING, W. J., QIAN, Y. Q., BILLETER, M., FURUKUBO-TOKUNAGA, K., SCHIER, A. F., RESENDEZ-PEREZ, D., AFFOLTER, M., OTTING, G. & WUTHRICH, K. 1994. Homeodomain-DNA recognition. *Cell,* 78**,** 211-23.

GEORGIEVA, S., IORDANOV, V. & SERGIEVA, S. 2009. Nature of cervical cancer and other HPV - associated cancers. *J BUON,* 14**,** 391-8.

GIARD, D. J., AARONSON, S. A., TODARO, G. J., ARNSTEIN, P., KERSEY, J. H., DOSIK, H. & PARKS, W. P. 1973. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J Natl Cancer Inst,* 51**,** 1417-23.

GILBERT, P. M., MOUW, J. K., UNGER, M. A., LAKINS, J. N., GBEGNON, M. K., CLEMMER, V. B., BENEZRA, M., LICHT, J. D., BOUDREAU, N. J., TSAI, K. K., WELM, A. L., FELDMAN, M. D., WEBER, B. L. & WEAVER, V. M. 2010. HOXA9 regulates BRCA1 expression to modulate human breast tumor phenotype. *J Clin Invest,* 120**,** 1535-50.

GILLISON, M. L. & SHAH, K. V. 2001. Human papillomavirus-associated head and neck squamous cell carcinoma: mounting evidence for an etiologic role for human papillomavirus in a subset of head and neck cancers. *Curr Opin Oncol,* 13**,** 183-8.

GIOVANNELLI, L., CAMPISI, G., LAMA, A., GIAMBALVO, O., OSBORN, J., MARGIOTTA, V. & AMMATUNA, P. 2002. Human papillomavirus DNA in oral mucosal lesions. *J Infect Dis,* 185**,** 833-6.

GOLPON, H. A., GERACI, M. W., MOORE, M. D., MILLER, H. L., MILLER, G. J., TUDER, R. M. & VOELKEL, N. F. 2001. HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. *Am J Pathol,* 158**,** 955-66.

GOLUB, T. R., SLONIM, D. K., TAMAYO, P., HUARD, C., GAASENBEEK, M., MESIROV, J. P., COLLER, H., LOH, M. L., DOWNING, J. R., CALIGIURI, M. A., BLOOMFIELD, C. D. & LANDER, E. S. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science,* 286**,** 531-7.

GOMBOS, K., HORVATH, R., SZELE, E., JUHASZ, K., GOCZE, K., SOMLAI, K., PAJKOS, G., EMBER, I. & OLASZ, L. 2013. miRNA expression profiles of oral squamous cell carcinomas. *Anticancer Res,* 33**,** 1511-7.

GONZALEZ-MOLES, M. A., SCULLY, C., RUIZ-AVILA, I. & PLAZA-CAMPILLO, J. J. 2013. The cancer stem cell hypothesis applied to oral carcinoma. *Oral Oncol,* 49**,** 738-46.

GONZALEZ, M. V., PELLO, M. F., LÓPEZ-LARREA, C., SUÁREZ, C., MENÉNDEZ, M. J. & COTO, E. 1995. Loss of heterozygosity and mutation analysis of the p16 (9p21) and p53 (17p13) genes in squamous cell carcinoma of the head and neck. *Clinical cancer research,* 1**,** 1043-1049.

GORE, A. I. & SPENCER, J. P. 2004. The newborn foot. *Am Fam Physician,* 69**,** 865-72.

GORSKI, D. H. & WALSH, K. 2003. Control of vascular cell differentiation by homeobox transcription factors. *Trends Cardiovasc Med,* 13**,** 213-20.

GOULD, M. K., SILVESTRI, G. A. & DETTERBECK, F. 2004. Multidisciplinary management of lung cancer. *N Engl J Med,* 350**,** 2008-10; author reply 2008-10.

GOY, J., HALL, S. F., FELDMAN-STEWART, D. & GROOME, P. A. 2009. Diagnostic delay and disease stage in head and neck cancer: a systematic review. *Laryngoscope,* 119**,** 889-98.

GRANDIS, J. R. & TWEARDY, D. J. 1993. Elevated levels of transforming growth factor alpha and epidermal growth factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer. *Cancer Res,* 53**,** 3579-84.

GRAY, S., PANDHA, H. S., MICHAEL, A., MIDDLETON, G. & MORGAN, R. 2011. HOX genes in pancreatic development and cancer. *JOP,* 12**,** 216-9.

GREGOIRE, V., LEFEBVRE, J.-L., LICITRA, L. & FELIP, E. 2010. Squamous cell carcinoma of the head and neck: EHNS–ESMO–ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology,* 21**,** v184-v186.

GUAN, J. L. 2005. *Cell migration: developmental methods and protocols*, Humana Pr Inc.

GUAN, Y.-S., LIU, Y., ZOU, Q., HE, Q., LA, Z., YANG, L. & HU, Y. 2009. Adenovirus-mediated wild-type p53 gene transfer in combination with bronchial arterial infusion for treatment of advanced non-small-cell lung cancer, one year follow-up. *Journal of Zhejiang University SCIENCE B,* 10**,** 331-340.

GUI, X.-H., QIU, L.-X., ZHANG, H.-F., ZHANG, D.-P., ZHONG, W.-Z., LI, J. & XIAO, Y.-L. 2009. < i> MDM2</i> 309 T/G polymorphism is associated with lung cancer risk among Asians. *European journal of cancer,* 45**,** 2023-2026.

GUO, F. J., ZHANG, W. J., LI, Y. L., LIU, Y., LI, Y. H., HUANG, J., WANG, J. J., XIE, P. L. & LI, G. C. 2010. Expression and functional characterization of platelet-derived growth factor receptor-like gene. *World J Gastroenterol,* 16**,** 1465-72.

HÅKANSSON, N., FLODERUS, B., GUSTAVSSON, P., FEYCHTING, M. & HALLIN, N. 2001. Occupational sunlight exposure and cancer incidence among Swedish construction workers. *Epidemiology,* 12**,** 552-557.

HALEES, A. S., LEYFER, D. & WENG, Z. 2003. PromoSer: A large-scale mammalian promoter and transcription start site identification service. *Nucleic Acids Res,* 31**,** 3554-9.

HAMMERSCHMIDT, S. & WIRTZ, H. 2009. Lung cancer: current diagnosis and treatment. *Dtsch Arztebl Int,* 106**,** 809-18; quiz 819-20.

HANAHAN, D. 1983. Studies on transformation of Escherichia coli with plasmids. *J Mol Biol,* 166**,** 557-80.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-674.

HANNAN, K. M., KENNEDY, B. K., CAVANAUGH, A. H., HANNAN, R. D., HIRSCHLER-LASZKIEWICZ, I., JEFFERSON, L. S. & ROTHBLUM, L. I. 2000. RNA polymerase I transcription in confluent cells: Rb downregulates rDNA transcription during confluence-induced cell cycle arrest. *Oncogene,* 19**,** 3487-97.

HARADA, H., NAKAGAWA, K., IWATA, S., SAITO, M., KUMON, Y., SAKAKI, S., SATO, K. & HAMADA, K. 1999. Restoration of wild-type p16 down-regulates vascular endothelial growth factor expression and inhibits angiogenesis in human gliomas. *Cancer research,* 59**,** 3783-3789.

HARRIS, T. J. & MCCORMICK, F. 2010. The molecular pathology of cancer. *Nat Rev Clin Oncol,* 7**,** 251-65.

HARRISON, L. B., SESSIONS, R. B. & HONG, W. K. 2009. *Head and neck cancer : a multidisciplinary approach,* Philadelphia, Pa. ; London, Lippincott Williams & Wilkins.

HARTWIG, S., SYRJÄNEN, S., DOMINIAK-FELDEN, G., BROTONS, M. & CASTELLSAGUÉ, X. 2012. Estimation of the epidemiological burden of human papillomavirus-related cancers and non-malignant diseases in men in Europe: a review. *BMC Cancer,* 12**,** 30.

HASHIBE, M., BRENNAN, P., CHUANG, S., BOCCIA, S., CASTELLSAGUE, X., CHEN, C., CURADO, M. P., DAL MASO, L., DAUDT, A. W. & FABIANOVA, E. 2009. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiology Biomarkers & Prevention,* 18**,** 541.

HASSAN, N. M., HAMADA, J., MURAI, T., SEINO, A., TAKAHASHI, Y., TADA, M., ZHANG, X., KASHIWAZAKI, H., YAMAZAKI, Y., INOUE, N. & MORIUCHI, T. 2006. Aberrant expression of HOX genes in oral dysplasia and squamous cell carcinoma tissues. *Oncol Res,* 16**,** 217-24.

HE, H., JAZDZEWSKI, K., LI, W., LIYANARACHCHI, S., NAGY, R., VOLINIA, S., CALIN, G. A., LIU, C. G., FRANSSILA, K., SUSTER, S., KLOOS, R. T., CROCE, C. M. & DE LA CHAPELLE, A. 2005. The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci U S A,* 102**,** 19075-80.

HEIDENREICH, W. F., TOMASEK, L., GROSCHE, B., LEURAUD, K. & LAURIER, D. 2012. Lung cancer mortality in the European uranium miners cohorts analyzed

with a biologically based model taking into account radon measurement error. *Radiat Environ Biophys,* 51**,** 263-75.
HENDERSON, G. S., VAN DIEST, P. J., BURGER, H., RUSSO, J. & RAMAN, V. 2006. Expression pattern of a homeotic gene, HOXA5, in normal breast and in breast tumors. *Cell Oncol,* 28**,** 305-13.
HESS, J. L. 2004. MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol Med,* 10**,** 500-7.
HICKSON, J., DIANE YAMADA, S., BERGER, J., ALVERDY, J., O'KEEFE, J., BASSLER, B. & RINKER-SCHAEFFER, C. 2009. Societal interactions in ovarian cancer metastasis: a quorum-sensing hypothesis. *Clin Exp Metastasis,* 26**,** 67-76.
HIGUCHI, E., ORIDATE, N., HOMMA, A., SUZUKI, F., ATAGO, Y., NAGAHASHI, T., FURUTA, Y. & FUKUDA, S. 2007. Prognostic significance of cyclin D1 and p16 in patients with intermediate-risk head and neck squamous cell carcinoma treated with docetaxel and concurrent radiotherapy. *Head Neck,* 29**,** 940-7.
HILL, J., CLARKE, J. D., VARGESSON, N., JOWETT, T. & HOLDER, N. 1995. Exogenous retinoic acid causes specific alterations in the development of the midbrain and hindbrain of the zebrafish embryo including positional respecification of the Mauthner neuron. *Mech Dev,* 50**,** 3-16.
HIRSCH, F. R., VARELLA-GARCIA, M., BUNN, P. A., DI MARIA, M. V., VEVE, R., BREMNES, R. M., BARÓN, A. E., ZENG, C. & FRANKLIN, W. A. 2003. Epidermal growth factor receptor in non–small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *Journal of clinical oncology,* 21**,** 3798-3807.
HOEY, T., DOYLE, H. J., HARDING, K., WEDEEN, C. & LEVINE, M. 1986. Homeo box gene expression in anterior and posterior regions of the Drosophila embryo. *Proc Natl Acad Sci U S A,* 83**,** 4809-13.
HOLLAND, P. W. H., BOOTH, H. A. F. & BRUFORD, E. A. 2007. Classification and nomenclature of all human homeobox genes. *BMC biology,* 5**,** 47.
HOLVE, S., FRIEDMAN, B., HOYME, H. E., TARBY, T. J., JOHNSTONE, S. J., ERICKSON, R. P., CLERICUZIO, C. L. & CUNNIFF, C. 2003. Athabascan brainstem dysgenesis syndrome. *Am J Med Genet A,* 120A**,** 169-73.
HOMBRÍA, J. C.-G. & LOVEGROVE, B. 2003. Beyond homeosis—HOX function in morphogenesis and organogenesis. *Differentiation,* 71**,** 461-476.
HONG, X., JIANG, F., KALKANIS, S. N., ZHANG, Z. G., ZHANG, X., ZHENG, X., JIANG, H. & CHOPP, M. 2009. Intracellular free calcium mediates glioma cell detachment and cytotoxicity after photodynamic therapy. *Lasers Med Sci,* 24**,** 777-86.
HORAN, G., RAMÍREZ-SOLIS, R., FEATHERSTONE, M. S., WOLGEMUTH, D. J., BRADLEY, A. & BEHRINGER, R. R. 1995. Compound mutants for the paralogous hoxa-4, hoxb-4, and hoxd-4 genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. *Genes & development,* 9**,** 1667-1677.
HÖRMANN, H. 1982. Fibronectin—mediator between cells and connective tissue. *Klinische Wochenschrift,* 60**,** 1265-1277.
HOU, J., XU, J., JIANG, R., WANG, Y., CHEN, C., DENG, L., HUANG, X., WANG, X. & SUN, B. 2013. Estrogen-sensitive PTPRO expression represses hepatocellular carcinoma progression by control of STAT3. *Hepatology,* 57**,** 678-88.
HOU, Z., XIE, L., YU, L., QIAN, X. & LIU, B. 2012a. MicroRNA-146a is down-regulated in gastric cancer and regulates cell proliferation and apoptosis. *Medical Oncology,* 29**,** 886-892.
HOU, Z., YIN, H., CHEN, C., DAI, X., LI, X., LIU, B. & FANG, X. 2012b. microRNA-146a targets the L1 cell adhesion molecule and suppresses the metastatic potential of gastric cancer. *Mol Med Rep,* 6**,** 501-6.
HRUZ, T., WYSS, M., DOCQUIER, M., PFAFFL, M. W., MASANETZ, S., BORGHI, L., VERBRUGGHE, P., KALAYDJIEVA, L., BLEULER, S., LAULE, O., DESCOMBES,

P., GRUISSEM, W. & ZIMMERMANN, P. 2011. RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalization. *BMC Genomics,* 12**,** 156.

HU, L., CROWE, D. L., RHEINWALD, J. G., CHAMBON, P. & GUDAS, L. J. 1991. Abnormal expression of retinoic acid receptors and keratin 19 by human oral and epidermal squamous cell carcinoma cell lines. *Cancer Res,* 51**,** 3972-81.

HU, X., CHEN, D., CUI, Y., LI, Z. & HUANG, J. 2013. Targeting microRNA-23a to inhibit glioma cell invasion via HOXD10. *Sci Rep,* 3**,** 3423.

HUANG, S., BENAVENTE, S., ARMSTRONG, E. A., LI, C., WHEELER, D. L. & HARARI, P. M. 2011. p53 modulates acquired resistance to EGFR inhibitors and radiation. *Cancer Res,* 71**,** 7071-9.

HUGHES, F. J. & ROMANOS, M. A. 1993. E1 protein of human papillomavirus is a DNA helicase/ATPase. *Nucleic Acids Res,* 21**,** 5817-23.

HUI, A. B., LENARDUZZI, M., KRUSHEL, T., WALDRON, L., PINTILIE, M., SHI, W., PEREZ-ORDONEZ, B., JURISICA, I., O'SULLIVAN, B., WALDRON, J., GULLANE, P., CUMMINGS, B. & LIU, F. F. 2010. Comprehensive MicroRNA profiling for head and neck squamous cell carcinomas. *Clin Cancer Res,* 16**,** 1129-39.

HUNG, P.-S., CHANG, K.-W., KAO, S.-Y., CHU, T.-H., LIU, C.-J. & LIN, S.-C. 2012. Association between the rs2910164 polymorphism in pre< i>-mir-146a</i> and oral carcinoma progression. *Oral oncology,* 48**,** 404-408.

HUNG, P. S., LIU, C. J., CHOU, C. S., KAO, S. Y., YANG, C. C., CHANG, K. W., CHIU, T. H. & LIN, S. C. 2013. miR-146a Enhances the Oncogenicity of Oral Carcinoma by Concomitant Targeting of the IRAK1, TRAF6 and NUMB Genes. *PLoS One,* 8**,** e79926.

HUNG, S.-C., WU, I.-H., HSUE, S.-S., LIAO, C.-H., WANG, H.-C., CHUANG, P.-H., SUNG, S.-Y. & HSIEH, C.-L. 2010. Targeting l1 cell adhesion molecule using lentivirus-mediated short hairpin RNA interference reverses aggressiveness of oral squamous cell carcinoma. *Molecular Pharmaceutics,* 7**,** 2312-2323.

HUNT, I., MUERS, M. F. & TREASURE, T. 2009. *ABC of lung cancer,* Chichester, Wiley-Blackwell/BMJ.

HUNT, J. L., BARNES, L., LEWIS JR, J. S., MAHFOUZ, M. E., SLOOTWEG, P. J., THOMPSON, L. D., CARDESA, A., DEVANEY, K. O., GNEPP, D. R. & WESTRA, W. H. 2013. Molecular diagnostic alterations in squamous cell carcinoma of the head and neck and potential diagnostic applications. *European Archives of Oto-Rhino-Laryngology***,** 1-13.

HUNT, P. 2008. Diagnosing and managing patients with lung cancer. *Nurs Stand,* 22**,** 50-6; quiz 58, 60.

HUNTER, K. D., THURLOW, J. K., FLEMING, J., DRAKE, P. J., VASS, J. K., KALNA, G., HIGHAM, D. J., HERZYK, P., MACDONALD, D. G., PARKINSON, E. K. & HARRISON, P. R. 2006. Divergent routes to oral cancer. *Cancer Res,* 66**,** 7405-13.

HYNES, N. E. & LANE, H. A. 2005. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer,* 5**,** 341-54.

HYUN, S. H., CHOI, J. Y., KIM, K., KIM, J., SHIM, Y. M., UM, S. W., KIM, H., LEE, K. H. & KIM, B. T. 2013. Volume-based parameters of (18)F-fluorodeoxyglucose positron emission tomography/computed tomography improve outcome prediction in early-stage non-small cell lung cancer after surgical resection. *Ann Surg,* 257**,** 364-70.

INNIS, J. W., GOODMAN, F. R., BACCHELLI, C., WILLIAMS, T. M., MORTLOCK, D. P., SATEESH, P., SCAMBLER, P. J., MCKINNON, W. & GUTTMACHER, A. E. 2002. A HOXA13 allele with a missense mutation in the homeobox and a dinucleotide deletion in the promoter underlies Guttmacher syndrome. *Human mutation,* 19**,** 573-574.

JACKMAN, D. M. & JOHNSON, B. E. 2005. Small-cell lung cancer. *Lancet,* 366**,** 1385-96.
JAZDZEWSKI, K., LIYANARACHCHI, S., SWIERNIAK, M., PACHUCKI, J., RINGEL, M. D., JARZAB, B. & DE LA CHAPELLE, A. 2009. Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. *Proc Natl Acad Sci U S A,* 106**,** 1502-5.
JEANMOUGIN, M., DE REYNIES, A., MARISA, L., PACCARD, C., NUEL, G. & GUEDJ, M. 2010. Should we abandon the t-test in the analysis of gene expression microarray data: a comparison of variance modeling strategies. *PLoS One,* 5**,** e12336.
JECHLINGER, M., GRUNERT, S., TAMIR, I. H., JANDA, E., LUDEMANN, S., WAERNER, T., SEITHER, P., WEITH, A., BEUG, H. & KRAUT, N. 2003. Expression profiling of epithelial plasticity in tumor progression. *Oncogene,* 22**,** 7155-69.
JEFFERY, I. B., HIGGINS, D. G. & CULHANE, A. C. 2006. Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. *BMC Bioinformatics,* 7**,** 359.
JEMAL, A., BRAY, F., CENTER, M. M., FERLAY, J., WARD, E. & FORMAN, D. 2011. Global cancer statistics. *CA Cancer J Clin,* 61**,** 69-90.
JEMAL, A., CLEGG, L. X., WARD, E., RIES, L. A., WU, X., JAMISON, P. M., WINGO, P. A., HOWE, H. L., ANDERSON, R. N. & EDWARDS, B. K. 2004. Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival. *Cancer,* 101**,** 3-27.
JEMAL, A., SIEGEL, R., WARD, E., MURRAY, T., XU, J. & THUN, M. J. 2007. Cancer statistics, 2007. *CA: a cancer journal for clinicians,* 57**,** 43-66.
JEMAL, A., SIEGEL, R., XU, J. & WARD, E. 2010. Cancer statistics, 2010. *CA: a cancer journal for clinicians,* 60**,** 277-300.
JIANG, J. G., CHEN, C. L., CARD, J. W., YANG, S., CHEN, J. X., FU, X. N., NING, Y. G., XIAO, X., ZELDIN, D. C. & WANG, D. W. 2005. Cytochrome P450 2J2 promotes the neoplastic phenotype of carcinoma cells and is up-regulated in human tumors. *Cancer Res,* 65**,** 4707-15.
JIANG, J. G., FU, X. N., CHEN, C. L. & WANG, D. W. 2009. Expression of cytochrome P450 arachidonic acid epoxygenase 2J2 in human tumor tissues and cell lines. *Ai Zheng,* 28**,** 93-6.
JIANG, R., DENG, L., ZHAO, L., LI, X., ZHANG, F., XIA, Y., GAO, Y., WANG, X. & SUN, B. 2011. miR-22 promotes HBV-related hepatocellular carcinoma development in males. *Clin Cancer Res,* 17**,** 5593-603.
JIN, K., KONG, X., SHAH, T., PENET, M. F., WILDES, F., SGROI, D. C., MA, X. J., HUANG, Y., KALLIONIEMI, A., LANDBERG, G., BIECHE, I., WU, X., LOBIE, P. E., DAVIDSON, N. E., BHUJWALLA, Z. M., ZHU, T. & SUKUMAR, S. 2012. The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway. *Proc Natl Acad Sci U S A,* 109**,** 2736-41.
JIN, K. & SUKUMAR, S. 2010. BRCA1: linking HOX to breast cancer suppression. *Breast Cancer Res,* 12**,** 306.
JOHNSON, D., KAN, S. H., OLDRIDGE, M., TREMBATH, R. C., ROCHE, P., ESNOUF, R. M., GIELE, H. & WILKIE, A. O. 2003. Missense mutations in the homeodomain of HOXD13 are associated with brachydactyly types D and E. *Am J Hum Genet,* 72**,** 984-97.
JOHNSON, J. L., PILLAI, S. & CHELLAPPAN, S. P. 2012. Genetic and biochemical alterations in non-small cell lung cancer. *Biochem Res Int,* 2012**,** 940405.
JOHNSON, R. L. & TABIN, C. J. 1997. Molecular models for vertebrate limb development. *Cell,* 90**,** 979-90.
JUDD, N. P., WINKLER, A. E., MURILLO-SAUCA, O., BROTMAN, J. J., LAW, J. H., LEWIS, J. S., JR., DUNN, G. P., BUI, J. D., SUNWOO, J. B. & UPPALURI, R.

2012. ERK1/2 regulation of CD44 modulates oral cancer aggressiveness. *Cancer Res,* 72**,** 365-74.

JUNG, C., KIM, R.-S., ZHANG, H.-J., LEE, S.-J. & JENG, M.-H. 2004. HOXB13 induces growth suppression of prostate cancer cells as a repressor of hormone-activated androgen receptor signaling. *Cancer research,* 64**,** 9185-9192.

KALYANKRISHNA, S. & GRANDIS, J. R. 2006. Epidermal growth factor receptor biology in head and neck cancer. *J Clin Oncol,* 24**,** 2666-72.

KAMANGAR, F., DORES, G. M. & ANDERSON, W. F. 2006. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol,* 24**,** 2137-50.

KAMB, A., GRUIS, N. A., WEAVER-FELDHAUS, J., LIU, Q., HARSHMAN, K., TAVTIGIAN, S. V., STOCKERT, E., DAY, R. S., 3RD, JOHNSON, B. E. & SKOLNICK, M. H. 1994. A cell cycle regulator potentially involved in genesis of many tumor types. *Science,* 264**,** 436-40.

KAN, Z., JAISWAL, B. S., STINSON, J., JANAKIRAMAN, V., BHATT, D., STERN, H. M., YUE, P., HAVERTY, P. M., BOURGON, R., ZHENG, J., MOORHEAD, M., CHAUDHURI, S., TOMSHO, L. P., PETERS, B. A., PUJARA, K., CORDES, S., DAVIS, D. P., CARLTON, V. E., YUAN, W., LI, L., WANG, W., EIGENBROT, C., KAMINKER, J. S., EBERHARD, D. A., WARING, P., SCHUSTER, S. C., MODRUSAN, Z., ZHANG, Z., STOKOE, D., DE SAUVAGE, F. J., FAHAM, M. & SESHAGIRI, S. 2010. Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature,* 466**,** 869-73.

KANDIOLER, D., STAMATIS, G., EBERHARDT, W., KAPPEL, S., ZOCHBAUER-MULLER, S., KUHRER, I., MITTLBOCK, M., ZWRTEK, R., AIGNER, C., BICHLER, C., TICHY, V., HUDEC, M., BACHLEITNER, T., END, A., MULLER, M. R., ROTH, E. & KLEPETKO, W. 2008. Growing clinical evidence for the interaction of the p53 genotype and response to induction chemotherapy in advanced non-small cell lung cancer. *J Thorac Cardiovasc Surg,* 135**,** 1036-41.

KARI, C., CHAN, T. O., DE QUADROS, M. R. & RODECK, U. 2003. Targeting the Epidermal Growth Factor Receptor in Cancer Apoptosis Takes Center Stage. *Cancer research,* 63**,** 1-5.

KATAYAMA, M., IWAMATSU, A., MASUTANI, H., FURUKE, K., TAKEDA, K., WADA, H., MASUDA, T. & ISHII, K. 1997. Expression of neural cell adhesion molecule L1 in human lung cancer cell lines. *Cell Struct Funct,* 22**,** 511-6.

KAUFMAN, T. C., LEWIS, R. & WAKIMOTO, B. 1980. Cytogenetic Analysis of Chromosome 3 in DROSOPHILA MELANOGASTER: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. *Genetics,* 94**,** 115-33.

KERNSTINE, K. H. & RECKAMP, K. L. 2010. *Lung cancer: a multidisciplinary approach to diagnosis and management*, Demos Medical Publishing.

KERR, M. K., MARTIN, M. & CHURCHILL, G. A. 2000. Analysis of variance for gene expression microarray data. *J Comput Biol,* 7**,** 819-37.

KEYNES, M. & COX, T. M. 2008. William Bateson, the rediscoverer of Mendel. *J R Soc Med,* 101**,** 104.

KHUDER, S. A. 2001. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. *Lung Cancer,* 31**,** 139-48.

KIESSLICH, T., ALINGER, B., WOLKERSDÖRFER, G. W., OCKER, M., NEUREITER, D. & BERR, F. 2010. Active Wnt signalling is associated with low differentiation and high proliferation in human biliary tract cancer in vitro and in vivo and is sensitive to pharmacological inhibition. *International journal of oncology,* 36**,** 49-58.

KIM, H., WATKINSON, J., VARADAN, V. & ANASTASSIOU, D. 2010a. Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics,* 3**,** 51.

KIM, M. H., JIN, H., SEOL, E. Y., YOO, M. & PARK, H. W. 2000. Sequence analysis and tissue specific expression of human HOXA7. *Mol Biotechnol,* 14**,** 19-24.

KIM, Y.-R., OH, K.-J., PARK, R.-Y., XUAN, N. T., KANG, T.-W., KWON, D.-D., CHOI, C., KIM, M. S., NAM, K. I. & AHN, K. Y. 2010b. Research HOXB13 promotes androgen independent growth of LNCaP prostate cancer cells by the activation of E2F signaling.

KIMURA, S., NAGANUMA, S., SUSUKI, D., HIRONO, Y., YAMAGUCHI, A., FUJIEDA, S., SANO, K. & ITOH, H. 2010. Expression of microRNAs in squamous cell carcinoma of human head and neck and the esophagus: miR-205 and miR-21 are specific markers for HNSCC and ESCC. *Oncol Rep,* 23**,** 1625-33.

KING, N. 2010. *RT-PCR protocols,* New York, Humana.

KNUDSEN, S. 1999. Promoter2.0: for the recognition of PolII promoter sequences. *Bioinformatics,* 15**,** 356-61.

KOCHAN, K. J., AMARAL, M. E., AGARWALA, R., SCHAFFER, A. A. & RIGGS, P. K. 2008. Application of dissociation curve analysis to radiation hybrid panel marker scoring: generation of a map of river buffalo (B. bubalis) chromosome 20. *BMC Genomics,* 9**,** 544.

KOK, K., NAYLOR, S. L. & BUYS, C. H. 1997. Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. *Adv Cancer Res,* 71**,** 27-92.

KONDO, I. & SHIMIZU, N. 1983. Mapping of the human gene for epidermal growth factor receptor (EGFR) on the p13 leads to q22 region of chromosome 7. *Cytogenet Cell Genet,* 35**,** 9-14.

KOOP, D., HOLLAND, N. D., SÉMON, M., ALVAREZ, S., DE LERA, A. R., LAUDET, V., HOLLAND, L. Z. & SCHUBERT, M. 2010. Retinoic acid signaling targets< i> Hox</i> genes during the amphioxus gastrula stage: Insights into early anterior–posterior patterning of the chordate body plan. *Developmental biology,* 338**,** 98-106.

KORGAONKAR, C., ZHAO, L., MODESTOU, M. & QUELLE, D. E. 2002. ARF function does not require p53 stabilization or Mdm2 relocalization. *Mol Cell Biol,* 22**,** 196-206.

KRAMER, N., WALZL, A., UNGER, C., ROSNER, M., KRUPITZA, G., HENGSTSCHLAGER, M. & DOLZNIG, H. 2013. In vitro cell migration and invasion assays. *Mutat Res,* 752**,** 10-24.

KREIMER, A. R., CLIFFORD, G. M., BOYLE, P. & FRANCESCHI, S. 2005. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiol Biomarkers Prev,* 14**,** 467-75.

KROON, E., KROSL, J., THORSTEINSDOTTIR, U., BABAN, S., BUCHBERG, A. M. & SAUVAGEAU, G. 1998. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *The EMBO Journal,* 17**,** 3714-3725.

KRUMLAUF, R. 1994. Hox genes in vertebrate development. *Cell,* 78**,** 191-201.

KUES, W. A., ANGER, M., CARNWATH, J. W., PAUL, D., MOTLIK, J. & NIEMANN, H. 2000. Cell cycle synchronization of porcine fetal fibroblasts: effects of serum deprivation and reversible cell cycle inhibitors. *Biol Reprod,* 62**,** 412-9.

KUMAR, A., PETRI, E. T., HALMOS, B. & BOGGON, T. J. 2008a. Structure and clinical relevance of the epidermal growth factor receptor in human cancer. *J Clin Oncol,* 26**,** 1742-51.

KUMAR, B., CORDELL, K. G., LEE, J. S., WORDEN, F. P., PRINCE, M. E., TRAN, H. H., WOLF, G. T., URBA, S. G., CHEPEHA, D. B. & TEKNOS, T. N. 2008b. EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer. *Journal of clinical oncology,* 26**,** 3128.

KUSAYAMA, M., WADA, K., NAGATA, M., ISHIMOTO, S., TAKAHASHI, H., YONEDA, M., NAKAJIMA, A., OKURA, M., KOGO, M. & KAMISAKI, Y. 2011. Critical role of aquaporin 3 on growth of human esophageal and oral squamous cell carcinoma. *Cancer Sci,* 102**,** 1128-36.

KWONG, R. A., KALISH, L. H., NGUYEN, T. V., KENCH, J. G., BOVA, R. J., COLE, I. E., MUSGROVE, E. A. & SUTHERLAND, R. L. 2005. p14ARF protein expression is a predictor of both relapse and survival in squamous cell carcinoma of the anterior tongue. *Clinical cancer research,* 11**,** 4107-4116.

LABBAYE, C., SPINELLO, I., QUARANTA, M. T., PELOSI, E., PASQUINI, L., PETRUCCI, E., BIFFONI, M., NUZZOLO, E. R., BILLI, M., FOA, R., BRUNETTI, E., GRIGNANI, F., TESTA, U. & PESCHLE, C. 2008. A three-step pathway comprising PLZF/miR-146a/CXCR4 controls megakaryopoiesis. *Nat Cell Biol,* 10**,** 788-801.

LACKO, M., OUDE OPHUIS, M. B., PETERS, W. H. & MANNI, J. J. 2009. Genetic polymorphisms of smoking-related carcinogen detoxifying enzymes and head and neck cancer susceptibility. *Anticancer Res,* 29**,** 753-61.

LADOU, J., CASTLEMAN, B., FRANK, A., GOCHFELD, M., GREENBERG, M., HUFF, J., JOSHI, T. K., LANDRIGAN, P. J., LEMEN, R., MYERS, J., SOFFRITTI, M., SOSKOLNE, C. L., TAKAHASHI, K., TEITELBAUM, D., TERRACINI, B. & WATTERSON, A. 2010. The case for a global ban on asbestos. *Environ Health Perspect,* 118**,** 897-901.

LAJER, C. B., NIELSEN, F. C., FRIIS-HANSEN, L., NORRILD, B., BORUP, R., GARNAES, E., ROSSING, M., SPECHT, L., THERKILDSEN, M. H., NAUNTOFTE, B., DABELSTEEN, S. & VON BUCHWALD, C. 2011. Different miRNA signatures of oral and pharyngeal squamous cell carcinomas: a prospective translational study. *Br J Cancer,* 104**,** 830-40.

LAM, S., BOYLE, P., HEALEY, G. F., MADDISON, P., PEEK, L., MURRAY, A., CHAPMAN, C. J., ALLEN, J., WOOD, W. C. & SEWELL, H. F. 2011. EarlyCDT-Lung: an immunobiomarker test as an aid to early detection of lung cancer. *Cancer Prevention Research,* 4**,** 1126-1134.

LANCE-JONES, C., OMELCHENKO, N., BAILIS, A., LYNCH, S. & SHARMA, K. 2001. Hoxd10 induction and regionalization in the developing lumbosacral spinal cord. *Development,* 128**,** 2255-68.

LANGENDIJK, J. A. & PSYRRI, A. 2010. The prognostic significance of p16 overexpression in oropharyngeal squamous cell carcinoma: implications for treatment strategies and future clinical studies. *Ann Oncol,* 21**,** 1931-4.

LANTUÉJOUL, S., SALAMEIRE, D., SALON, C. & BRAMBILLA, E. 2009. Pulmonary preneoplasia–sequential molecular carcinogenetic events. *Histopathology,* 54**,** 43-54.

LANTZ, P. M., MENDEZ, D. & PHILBERT, M. A. 2013. Radon, smoking, and lung cancer: the need to refocus radon control policy. *Am J Public Health,* 103**,** 443-7.

LAPORTA, R. F. & TAICHMAN, L. B. 1982. Human papilloma viral DNA replicates as a stable episome in cultured epidermal keratinocytes. *Proc Natl Acad Sci U S A,* 79**,** 3393-7.

LAPPIN, T. R., GRIER, D. G., THOMPSON, A. & HALLIDAY, H. L. 2006. HOX genes: seductive science, mysterious mechanisms. *Ulster Med J,* 75**,** 23-31.

LARONDE-LEBLANC, N. A. & WOLBERGER, C. 2003. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev,* 17**,** 2060-72.

LASH, G. E., INNES, B. A., DRURY, J. A., ROBSON, S. C., QUENBY, S. & BULMER, J. N. 2012. Localization of angiogenic growth factors and their receptors in the human endometrium throughout the menstrual cycle and in recurrent miscarriage. *Human reproduction,* 27**,** 183-195.

LASSEN, P., ERIKSEN, J. G., HAMILTON-DUTOIT, S., TRAMM, T., ALSNER, J. & OVERGAARD, J. 2009. Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck. *Journal of clinical oncology,* 27**,** 1992-1998.

LASSIG, A. A., LINDGREN, B. R., FERNANDES, P., COOPER, S., ARDESHIPOUR, F., SCHOTZKO, C. & YUEH, B. 2013. The effect of young age on outcomes in head and neck cancer. *Laryngoscope,* 123**,** 1896-902.

LAUFFENBURGER, D. A. & HORWITZ, A. F. 1996. Cell migration: a physically integrated molecular process. *Cell,* 84**,** 359-69.

LAWRENCE, H. J., CHRISTENSEN, J., FONG, S., HU, Y. L., WEISSMAN, I., SAUVAGEAU, G., HUMPHRIES, R. K. & LARGMAN, C. 2005. Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood,* 106**,** 3988-94.

LAWRENCE, H. J., HELGASON, C. D., SAUVAGEAU, G., FONG, S., IZON, D. J., HUMPHRIES, R. K. & LARGMAN, C. 1997. Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis. *Blood,* 89**,** 1922-1930.

LAWRENCE, H. J. & LARGMAN, C. 1992. Homeobox genes in normal hematopoiesis and leukemia. *Blood,* 80**,** 2445-53.

LECHNER, J. F., WANG, Y., SIDDIQ, F., FUGARO, J. M., WALI, A., LONARDO, F., WILLEY, J. C., HARRIS, C. C. & PASS, H. I. 2002. Human lung cancer cells and tissues partially recapitulate the homeobox gene expression profile of embryonic lung. *Lung Cancer,* 37**,** 41-7.

LECHNER, M., FRAMPTON, G., FENTON, T., FEBER, A., PALMER, G., JAY, A., PILLAY, N., FORSTER, M., CRONIN, M. T. & LIPSON, D. 2013. Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV-tumors. *Genome medicine,* 5**,** 49.

LEE, J. M., YOON, T. J. & CHO, Y. S. 2013. Recent developments in nanoparticle-based siRNA delivery for cancer therapy. *Biomed Res Int,* 2013**,** 782041.

LEE, M.-Y., LIN, K.-D., HSIAO, P.-J. & SHIN, S.-J. 2012. The association of diabetes mellitus with liver, colon, lung, and prostate cancer is independent of hypertension, hyperlipidemia, and gout in Taiwanese patients. *Metabolism,* 61**,** 242-249.

LEMON, B. & TJIAN, R. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev,* 14**,** 2551-69.

LEMONS, D. & MCGINNIS, W. 2006. Genomic evolution of Hox gene clusters. *Science,* 313**,** 1918-22.

LEVINE, M. & HOEY, T. 1988. Homeobox proteins as sequence-specific transcription factors. *Cell,* 55**,** 537-540.

LEVINE, M., RUBIN, G. M. & TJIAN, R. 1984. Human DNA sequences homologous to a protein coding region conserved between homeotic genes of Drosophila. *Cell,* 38**,** 667-673.

LEWIS, E. B. 1978. A gene complex controlling segmentation in Drosophila. *Nature,* 276**,** 565-70.

LI, C.-L., NIE, H., WANG, M., SU, L.-P., LI, J.-F., YU, Y.-Y., YAN, M., QU, Q.-L., ZHU, Z.-G. & LIU, B.-Y. 2012a. microRNA-155 is downregulated in gastric cancer cells and involved in cell metastasis. *Oncology reports,* 27**,** 1960-1966.

LI, J., POI, M. J. & TSAI, M.-D. 2011. Regulatory mechanisms of tumor suppressor P16INK4A and their relevance to cancer. *Biochemistry,* 50**,** 5566-5582.

LI, J., ZUCKER, S., PULKOSKI-GROSS, A., KUSCU, C., KARAAYVAZ, M., JU, J., YAO, H., SONG, E. & CAO, J. 2012b. Conversion of stationary to invasive tumor initiating cells (TICs): role of hypoxia in membrane type 1-matrix metalloproteinase (MT1-MMP) trafficking. *PLoS One,* 7**,** e38403.

LI, L., ZHENG, P. & DEAN, J. 2010a. Maternal control of early mouse development. *Development,* 137**,** 859-70.

LI, Q., DING, C., CHEN, C., ZHANG, Z., XIAO, H., XIE, F., LEI, L., CHEN, Y., MAO, B., JIANG, M., LI, J., WANG, D. & WANG, G. 2013a. miR-224 promotes cell migration and invasion by modulating p-PAK4 and MMP-9 via targeting HOXD10 in human hepatocellular carcinoma. *J Gastroenterol Hepatol*.

LI, Y., BURNS, J. A., CHENEY, C. A., ZHANG, N., VITELLI, S., WANG, F., BETT, A., CHASTAIN, M., AUDOLY, L. P. & ZHANG, Z. Q. 2010b. Distinct expression profiles of Notch-1 protein in human solid tumors: Implications for development of targeted therapeutic monoclonal antibodies. *Biologics,* 4**,** 163-71.

LI, Y., VANDENBOOM, T. G., 2ND, WANG, Z., KONG, D., ALI, S., PHILIP, P. A. & SARKAR, F. H. 2010c. miR-146a suppresses invasion of pancreatic cancer cells. *Cancer Res,* 70**,** 1486-95.

LI, Z., ZHANG, Z., LI, Y., ARNOVITZ, S., CHEN, P., HUANG, H., JIANG, X., HONG, G.-M., KUNJAMMA, R. B. & REN, H. 2013b. PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood,* 121**,** 1422-1431.

LIANG, Z. & BIGGIN, M. D. 1998. Eve and ftz regulate a wide array of genes in blastoderm embryos: the selector homeoproteins directly or indirectly regulate most genes in Drosophila. *Development,* 125**,** 4471-4482.

LIN, F. & PRICHARD, J. 2011. *Handbook of practical immunohistochemistry : frequently asked questions,* New York, Springer.

LIN, S. L., CHIANG, A., CHANG, D. & YING, S. Y. 2008. Loss of mir-146a function in hormone-refractory prostate cancer. *RNA,* 14**,** 417-24.

LIN, W. J., HSUEH, H. M. & CHEN, J. J. 2010. Power and sample size estimation in microarray studies. *BMC Bioinformatics,* 11**,** 48.

LIN, Z., CHEN, Q., SHI, L., LEE, M., GIEHL, K. A., TANG, Z., WANG, H., ZHANG, J., YIN, J. & WU, L. 2012. Loss-of-Function Mutations in< i> HOXC13</i> Cause Pure Hair and Nail Ectodermal Dysplasia. *The American Journal of Human Genetics*.

LINDEMAN, R. E. & PELEGRI, F. 2010. Vertebrate maternal-effect genes: Insights into fertilization, early cleavage divisions, and germ cell determinant localization from studies in the zebrafish. *Molecular reproduction and development,* 77**,** 299-313.

LINDENBERGH-VAN DER PLAS, M., MARTENS-DE KEMP, S. R., DE MAAKER, M., VAN WIERINGEN, W. N., YLSTRA, B., AGAMI, R., CERISOLI, F., LEEMANS, C. R., BRAAKHUIS, B. J. & BRAKENHOFF, R. H. 2013. Identification of lethal microRNAs specific for head and neck cancer. *Clinical cancer research*.

LINDSEY, S. & WILKINSON, M. F. 1996. Homeobox genes and male reproductive development. *J Assist Reprod Genet,* 13**,** 182-92.

LIPPMAN, S. M., SUDBO, J. & HONG, W. K. 2005. Oral cancer prevention and the evolution of molecular-targeted drug development. *J Clin Oncol,* 23**,** 346-56.

LISSOWSKA, J., FORETOVA, L., DABEK, J., ZARIDZE, D., SZESZENIA-DABROWSKA, N., RUDNAI, P., FABIANOVA, E., CASSIDY, A., MATES, D., BENCKO, V., JANOUT, V., HUNG, R. J., BRENNAN, P. & BOFFETTA, P. 2010. Family history and lung cancer risk: international multicentre case-control study in Eastern and Central Europe and meta-analyses. *Cancer Causes Control,* 21**,** 1091-104.

LIU, B., ZHANG, H., DUAN, X., HAO, J., XIE, Y., ZHOU, Q., WANG, Y., TIAN, Y. & WANG, T. 2009. Adenovirus-mediated wild-type p53 transfer radiosensitizes H1299 cells to subclinical-dose carbon-ion irradiation through the restoration of p53 function. *Cancer biotherapy & radiopharmaceuticals,* 24**,** 57-66.

LIU, Z., ZHU, J., CAO, H., REN, H. & FANG, X. 2012. miR-10b promotes cell invasion through RhoC-AKT signaling pathway by targeting HOXD10 in gastric cancer. *International journal of oncology,* 40**,** 1553.

LODISH, H. F. 1999. *Molecular cell biology,* New York ; Basingstoke, W.H. Freeman.

LOOTS, G. G., CHAIN, P. S., MABERY, S., RASLEY, A., GARCIA, E. & OVCHARENKO, I. 2006. Array2BIO: from microarray expression data to functional annotation of co-regulated genes. *BMC Bioinformatics,* 7**,** 307.

LOPEZ, R., GARRIDO, E., VAZQUEZ, G., PINA, P., PEREZ, C., ALVARADO, I. & SALCEDO, M. 2006. A subgroup of HOX Abd-B gene is differentially expressed in cervical cancer. *International Journal of Gynecological Cancer,* 16**,** 1289-1296.

LOZANO, R., NAGHAVI, M., FOREMAN, K., LIM, S., SHIBUYA, K., ABOYANS, V., ABRAHAM, J., ADAIR, T., AGGARWAL, R., AHN, S. Y., ALVARADO, M., ANDERSON, H. R., ANDERSON, L. M., ANDREWS, K. G., ATKINSON, C., BADDOUR, L. M., BARKER-COLLO, S., BARTELS, D. H., BELL, M. L., BENJAMIN, E. J., BENNETT, D., BHALLA, K., BIKBOV, B., BIN ABDULHAK, A., BIRBECK, G., BLYTH, F., BOLLIGER, I., BOUFOUS, S., BUCELLO, C., BURCH, M., BURNEY, P., CARAPETIS, J., CHEN, H., CHOU, D., CHUGH, S. S., COFFENG, L. E., COLAN, S. D., COLQUHOUN, S., COLSON, K. E., CONDON, J., CONNOR, M. D., COOPER, L. T., CORRIERE, M., CORTINOVIS, M., DE VACCARO, K. C., COUSER, W., COWIE, B. C., CRIQUI, M. H., CROSS, M., DABHADKAR, K. C., DAHODWALA, N., DE LEO, D., DEGENHARDT, L., DELOSSANTOS, A., DENENBERG, J., DES JARLAIS, D. C., DHARMARATNE, S. D., DORSEY, E. R., DRISCOLL, T., DUBER, H., EBEL, B., ERWIN, P. J., ESPINDOLA, P., EZZATI, M., FEIGIN, V., FLAXMAN, A. D., FOROUZANFAR, M. H., FOWKES, F. G., FRANKLIN, R., FRANSEN, M., FREEMAN, M. K., GABRIEL, S. E., GAKIDOU, E., GASPARI, F., GILLUM, R. F., GONZALEZ-MEDINA, D., HALASA, Y. A., HARING, D., HARRISON, J. E., HAVMOELLER, R., HAY, R. J., HOEN, B., HOTEZ, P. J., HOY, D., JACOBSEN, K. H., JAMES, S. L., JASRASARIA, R., JAYARAMAN, S., JOHNS, N., KARTHIKEYAN, G., KASSEBAUM, N., KEREN, A., KHOO, J. P., KNOWLTON, L. M., KOBUSINGYE, O., KORANTENG, A., KRISHNAMURTHI, R., LIPNICK, M., LIPSHULTZ, S. E., OHNO, S. L., et al. 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet,* 380**,** 2095-128.

LUBIN, J. H., BOICE, J. D., JR., EDLING, C., HORNUNG, R. W., HOWE, G. R., KUNZ, E., KUSIAK, R. A., MORRISON, H. I., RADFORD, E. P., SAMET, J. M. & ET AL. 1995. Lung cancer in radon-exposed miners and estimation of risk from indoor exposure. *J Natl Cancer Inst,* 87**,** 817-27.

LUDOVINI, V., PISTOLA, L., GREGORC, V., FLORIANI, I., RULLI, E., PIATTONI, S., DI CARLO, L., SEMERARO, A., DARWISH, S. & TOFANETTI, F. R. 2008. Plasma DNA, microsatellite alterations, and p53 tumor mutations are associated with disease-free survival in radically resected non-small cell lung cancer patients: a study of the perugia multidisciplinary team for thoracic oncology. *Journal of Thoracic Oncology,* 3**,** 365-373.

LUMSDEN, A. & KRUMLAUF, R. 1996. Patterning the vertebrate neuraxis. *Science,* 274**,** 1109-15.

LUO, S. Y. & LAM, D. C. 2013. Oncogenic driver mutations in lung cancer. *Translational Respiratory Medicine,* 1**,** 1-8.

LYNCH, T. J., BELL, D. W., SORDELLA, R., GURUBHAGAVATULA, S., OKIMOTO, R. A., BRANNIGAN, B. W., HARRIS, P. L., HASERLAT, S. M., SUPKO, J. G., HALUSKA, F. G., LOUIS, D. N., CHRISTIANI, D. C., SETTLEMAN, J. & HABER, D. A. 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med,* 350**,** 2129-39.

MA, L., TERUYA-FELDSTEIN, J. & WEINBERG, R. A. 2007. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature,* 449**,** 682-8.

MACHADO, J., REIS, P. P., ZHANG, T., SIMPSON, C., XU, W., PEREZ-ORDONEZ, B., GOLDSTEIN, D. P., BROWN, D. H., GILBERT, R. W., GULLANE, P. J., IRISH, J. C. & KAMEL-REID, S. 2010. Low prevalence of human papillomavirus in oral cavity carcinomas. *Head Neck Oncol,* 2**,** 6.

MACK, J. A., ABRAMSON, S. R., BEN, Y., COFFIN, J. C., ROTHROCK, J. K., MAYTIN, E. V., HASCALL, V. C., LARGMAN, C. & STELNICKI, E. J. 2003. Hoxb13 knockout adult skin exhibits high levels of hyaluronan and enhanced wound healing. *FASEB J,* 17**,** 1352-4.

MACKINNON, A. C., KOPATZ, J. & SETHI, T. 2010. The molecular and cellular biology of lung cancer: identifying novel therapeutic strategies. *Br Med Bull,* 95**,** 47-61.

MACONOCHIE, M., NONCHEV, S., MORRISON, A. & KRUMLAUF, R. 1996. Paralogous Hox genes: function and regulation. *Annu Rev Genet,* 30**,** 529-56.

MAGLI, M. C., LARGMAN, C. & LAWRENCE, H. J. 1997. Effects of HOX homeobox genes in blood cell differentiation. *J Cell Physiol,* 173**,** 168-77.

MAGRINI, R., RUSSO, D., OTTAGGIO, L., FRONZA, G., INGA, A. & MENICHINI, P. 2008. PRIMA-1 synergizes with adriamycin to induce cell death in non-small cell lung cancer cells. *J Cell Biochem,* 104**,** 2363-73.

MALLO, M., WELLIK, D. M. & DESCHAMPS, J. 2010. Hox genes and regional patterning of the vertebrate body plan. *Dev Biol,* 344**,** 7-15.

MANN, R. S. & CHAN, S.-K. 1996. Extra specificity from< i> extradenticle</i>: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends in Genetics,* 12**,** 258-262.

MARSHALL, H., MORRISON, A., STUDER, M., POPPERL, H. & KRUMLAUF, R. 1996. Retinoids and Hox genes. *FASEB J,* 10**,** 969-78.

MARUR, S. & FORASTIERE, A. A. 2008. Head and neck cancer: changing epidemiology, diagnosis, and treatment. *Mayo Clin Proc,* 83**,** 489-501.

MASCAUX, C., IANNINO, N., MARTIN, B., PAESMANS, M., BERGHMANS, T., DUSART, M., HALLER, A., LOTHAIRE, P., MEERT, A. P., NOEL, S., LAFITTE, J. J. & SCULIER, J. P. 2005. The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer,* 92**,** 131-9.

MASSARELLI, E., VARELLA-GARCIA, M., TANG, X., XAVIER, A. C., OZBURN, N. C., LIU, D. D., BEKELE, B. N., HERBST, R. S. & WISTUBA, I. I. 2007. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non–small-cell lung cancer. *Clinical Cancer Research,* 13**,** 2890-2896.

MATAKIDOU, A., EISEN, T. & HOULSTON, R. 2005. Systematic review of the relationship between family history and lung cancer risk. *British journal of cancer,* 93**,** 825-833.

MAZIERES, J., ROUQUETTE, I., LEPAGE, B., MILIA, J., BROUCHET, L., GUIBERT, N., BEAU-FALLER, M., VALIDIRE, P., HOFMAN, P. & FOURET, P. 2013. Specificities of lung adenocarcinoma in women who have never smoked. *J Thorac Oncol,* 8**,** 923-9.

MCBRIDE, O., MERRY, D. & GIVOL, D. 1986. The gene for human p53 cellular tumor antigen is located on chromosome 17 short arm (17p13). *Proceedings of the National Academy of Sciences of the United States of America,* 83**,** 130.

MCCABE, C. D., SPYROPOULOS, D. D., MARTIN, D. & MORENO, C. S. 2008. Genome-wide analysis of the homeobox C6 transcriptional network in prostate cancer. *Cancer Res,* 68**,** 1988-96.

MCDONIELS-SILVERS, A. L., NIMRI, C. F., STONER, G. D., LUBET, R. A. & YOU, M. 2002. Differential gene expression in human lung adenocarcinomas and squamous cell carcinomas. *Clin Cancer Res,* 8**,** 1127-38.

MCGINNIS, W., GARBER, R. L., WIRZ, J., KUROIWA, A. & GEHRING, W. J. 1984a. A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell,* 37**,** 403-8.

MCGINNIS, W., LEVINE, M. S., HAFEN, E., KUROIWA, A. & GEHRING, W. J. 1984b. A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature,* 308**,** 428-33.

MCGREGOR, F., MUNTONI, A., FLEMING, J., BROWN, J., FELIX, D. H., MACDONALD, D. G., PARKINSON, E. K. & HARRISON, P. R. 2002. Molecular changes associated with oral dysplasia progression and acquisition of immortality: potential for its reversal by 5-azacytidine. *Cancer Res,* 62**,** 4757-66.

MCINTYRE, D. C., RAKSHIT, S., YALLOWITZ, A. R., LOKEN, L., JEANNOTTE, L., CAPECCHI, M. R. & WELLIK, D. M. 2007. Hox patterning of the vertebrate rib cage. *Development,* 134**,** 2981-9.

MCKIE, J. & RADOMISLI, T. 2010. Congenital vertical talus: a review. *Clinics in podiatric medicine and surgery,* 27**,** 145-156.

MCLAUGHLIN-DRUBIN, M. E. & MUNGER, K. 2009. Oncogenic activities of human papillomaviruses. *Virus Res,* 143**,** 195-208.

MEHANNA, H., PALERI, V., WEST, C. M. & NUTTING, C. 2010. Head and neck cancer--Part 1: Epidemiology, presentation, and prevention. *BMJ,* 341**,** c4684.

MEHRA, R., COHEN, R. B. & BURTNESS, B. A. 2008. The role of cetuximab for the treatment of squamous cell carcinoma of the head and neck. *Clin Adv Hematol Oncol,* 6**,** 742-50.

MEHROTRA, R. & YADAV, S. 2006. Oral squamous cell carcinoma: etiology, pathogenesis and prognostic value of genomic alterations. *Indian J Cancer,* 43**,** 60-6.

MEI, J., BACHOO, R. & ZHANG, C. L. 2011. MicroRNA-146a inhibits glioma development by targeting Notch1. *Mol Cell Biol,* 31**,** 3584-92.

MESSEGUER, X., ESCUDERO, R., FARRE, D., NUNEZ, O., MARTINEZ, J. & ALBA, M. M. 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics,* 18**,** 333-4.

MEYER, L. R., ZWEIG, A. S., HINRICHS, A. S., KAROLCHIK, D., KUHN, R. M., WONG, M., SLOAN, C. A., ROSENBLOOM, K. R., ROE, G., RHEAD, B., RANEY, B. J., POHL, A., MALLADI, V. S., LI, C. H., LEE, B. T., LEARNED, K., KIRKUP, V., HSU, F., HEITNER, S., HARTE, R. A., HAEUSSLER, M., GURUVADOO, L., GOLDMAN, M., GIARDINE, B. M., FUJITA, P. A., DRESZER, T. R., DIEKHANS, M., CLINE, M. S., CLAWSON, H., BARBER, G. P., HAUSSLER, D. & KENT, W. J. 2013. The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res,* 41**,** D64-9.

MEYERS, R. A. 2004. *Encyclopedia of molecular cell biology and molecular medicine,* Weinheim ; [Chichester], Wiley-VCH.

MILLER, G. J., MILLER, H. L., VAN BOKHOVEN, A., LAMBERT, J. R., WERAHERA, P. N., SCHIRRIPA, O., LUCIA, M. S. & NORDEEN, S. K. 2003. Aberrant HOXC expression accompanies the malignant phenotype in human prostate. *Cancer Res,* 63**,** 5879-88.

MITSUDOMI, T., HAMAJIMA, N., OGAWA, M. & TAKAHASHI, T. 2000. Prognostic significance of p53 alterations in patients with non-small cell lung cancer: a meta-analysis. *Clinical Cancer Research,* 6**,** 4055-4063.

MITSUDOMI, T., KOSAKA, T., ENDOH, H., HORIO, Y., HIDA, T., MORI, S., HATOOKA, S., SHINODA, M., TAKAHASHI, T. & YATABE, Y. 2005. Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. *J Clin Oncol,* 23**,** 2513-20.

MITSUDOMI, T. & YATABE, Y. 2007. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer science,* 98**,** 1817-1824.

MOENS, C. B. & SELLERI, L. 2006. Hox cofactors in vertebrate development. *Dev Biol,* 291**,** 193-206.

MOGI, A. & KUWANO, H. 2011. TP53 mutations in nonsmall cell lung cancer. *J Biomed Biotechnol,* 2011**,** 583929.

MOON, S.-M., KIM, S.-A., YOON, J.-H. & AHN, S.-G. 2012. HOXC6 is deregulated in human head and neck squamous cell carcinoma and modulates Bcl-2 expression. *Journal of Biological Chemistry,* 287**,** 35678-35688.

MORGAN, R. 2006. Hox genes: a continuation of embryonic patterning? *Trends Genet,* 22**,** 67-9.

MORGAN, R., PIRARD, P. M., SHEARS, L., SOHAL, J., PETTENGELL, R. & PANDHA, H. S. 2007. Antagonism of HOX/PBX dimer formation blocks the in vivo proliferation of melanoma. *Cancer Res,* 67**,** 5806-13.

MORGAN, R., PLOWRIGHT, L., HARRINGTON, K. J., MICHAEL, A. & PANDHA, H. S. 2010. Targeting HOX and PBX transcription factors in ovarian cancer. *BMC Cancer,* 10**,** 89.

MORGENSZTERN, D., NG, S. H., GAO, F. & GOVINDAN, R. 2010. Trends in stage distribution for patients with non-small cell lung cancer: a National Cancer Database survey. *J Thorac Oncol,* 5**,** 29-33.

MORTLOCK, D. P. & INNIS, J. W. 1997. Mutation of HOXA13 in hand-foot-genital syndrome. *Nat Genet,* 15**,** 179-80.

MOSSMAN, B. T., BIGNON, J., CORN, M., SEATON, A. & GEE, J. B. 1990. Asbestos: scientific developments and implications for public policy. *Science,* 247**,** 294-301.

MUNGER, K., WERNESS, B. A., DYSON, N., PHELPS, W. C., HARLOW, E. & HOWLEY, P. M. 1989. Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J,* 8**,** 4099-105.

MUNOZ, N., BOSCH, F. X., DE SANJOSE, S., HERRERO, R., CASTELLSAGUE, X., SHAH, K. V., SNIJDERS, P. J. & MEIJER, C. J. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med,* 348**,** 518-27.

MUTTAGI, S. S., CHATURVEDI, P., GAIKWAD, R., SINGH, B. & PAWAR, P. 2012. Head and neck squamous cell carcinoma in chronic areca nut chewing Indian women: Case series and review of literature. *Indian journal of medical and paediatric oncology: official journal of Indian Society of Medical & Paediatric Oncology,* 33**,** 32.

MYERS, C., CHARBONEAU, A., CHEUNG, I., HANKS, D. & BOUDREAU, N. 2002. Sustained expression of homeobox D10 inhibits angiogenesis. *Am J Pathol,* 161**,** 2099-109.

NA, I.-K., SCHEIBENBOGEN, C., ADAM, C., STROUX, A., GHADJAR, P., THIEL, E., KEILHOLZ, U. & COUPLAND, S. E. 2008. Nuclear expression of CXCR4 in tumor cells of non–small cell lung cancer is correlated with lymph node metastasis. *Human pathology,* 39**,** 1751-1755.

NACHT, M., DRACHEVA, T., GAO, Y., FUJII, T., CHEN, Y., PLAYER, A., AKMAEV, V., COOK, B., DUFAULT, M., ZHANG, M., ZHANG, W., GUO, M., CURRAN, J., HAN, S., SIDRANSKY, D., BUETOW, K., MADDEN, S. L. & JEN, J. 2001. Molecular characteristics of non-small cell lung cancer. *Proc Natl Acad Sci U S A,* 98**,** 15203-8.

NADON, R. & SHOEMAKER, J. 2002. Statistical issues with microarrays: processing and analysis. *Trends Genet,* 18**,** 265-71.

NAKASHIMA, T., KURATOMI, Y., YASUMATSU, R., MASUDA, M., KOIKE, K., UMEZAKI, T., CLAYMAN, G. L., NAKAGAWA, T. & KOMUNE, S. 2005. The effect of cyclin D1 overexpression in human head and neck cancer cells. *Eur Arch Otorhinolaryngol,* 262**,** 379-83.

NAKAYAMA, I., SHIBAZAKI, M., YASHIMA-ABO, A., MIURA, F., SUGIYAMA, T., MASUDA, T. & MAESAWA, C. 2013. Loss of HOXD10 expression induced by upregulation of miR-10b accelerates the migration and invasion activities of ovarian cancer cells. *Int J Oncol,* 43**,** 63-71.

NEVILLE, S. E., BAIGENT, S. M., BICKNELL, A. B., LOWRY, P. J. & GLADWELL, R. T. 2002. Hox gene expression in adult tissues with particular reference to the adrenal gland. *Endocr Res,* 28**,** 669-73.

NICHOLS, A. C., CHAN-SENG-YUE, M., YOO, J., XU, W., DHALIWAL, S., BASMAJI, J., SZETO, C. C., DOWTHWAITE, S., TODOROVIC, B. & STARMANS, M. H. 2012. A pilot study comparing HPV-positive and HPV-negative head and neck squamous cell carcinomas by whole exome sequencing. *ISRN oncology,* 2012.

NICHOLSON, W. J. 1986. *Airborne asbestos health assessment update,* Research Triangle Park, N.C., Environmental Criteria and Assessment Office, Office of Health and Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency.

NORIEGA, E. & RAMOS, E. 2013. New mutation in periaxin gene causing charcot marie tooth disease in a puerto rican young male. *J Clin Neuromuscul Dis,* 15**,** 63-8.

NORMANNO, N., DE LUCA, A., BIANCO, C., STRIZZI, L., MANCINO, M., MAIELLO, M. R., CAROTENUTO, A., DE FEO, G., CAPONIGRO, F. & SALOMON, D. S. 2006. Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene,* 366**,** 2-16.

NURWIDYA, F., TAKAHASHI, F., MURAKAMI, A. & TAKAHASHI, K. 2012. Epithelial mesenchymal transition in drug resistance and metastasis of lung cancer. *Cancer Res Treat,* 44**,** 151-6.

OHTA, S., UEMURA, H., MATSUI, Y., ISHIGURO, H., FUJINAMI, K., KONDO, K., MIYAMOTO, H., YAZAWA, T., DANENBERG, K. & DANENBERG, P. V. 2009. Alterations of p16 and p14ARF genes and their 9p21 locus in oral squamous cell carcinoma. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology,* 107**,** 81-91.

OLIVIER, M., EELES, R., HOLLSTEIN, M., KHAN, M. A., HARRIS, C. C. & HAINAUT, P. 2002. The IARC TP53 database: new online mutation analysis and recommendations to users. *Human mutation,* 19**,** 607-614.

ONG, C. C., JUBB, A. M., ZHOU, W., HAVERTY, P. M., HARRIS, A. L., BELVIN, M., FRIEDMAN, L. S., KOEPPEN, H. & HOEFLICH, K. P. 2011. p21-activated kinase 1: PAK'ed with potential. *Oncotarget,* 2**,** 491-6.

OSBORNE, J., HU, C., HAWLEY, C., UNDERWOOD, L. J., O'BRIEN, T. J. & BAKER, V. V. 1998. Expression of HOXD10 gene in normal endometrium and endometrial adenocarcinoma. *J Soc Gynecol Investig,* 5**,** 277-80.

OVERDEVEST, J. B., KNUBEL, K. H., DUEX, J. E., THOMAS, S., NITZ, M. D., HARDING, M. A., SMITH, S. C., FRIERSON, H. F., CONAWAY, M. & THEODORESCU, D. 2012. CD24 expression is important in male urothelial tumorigenesis and metastasis in mice and is androgen regulated. *Proceedings of the National Academy of Sciences,* 109**,** E3588-E3596.

PAIK, J. H., JANG, J.-Y., JEON, Y. K., KIM, W. Y., KIM, T. M., HEO, D. S. & KIM, C.-W. 2011. MicroRNA-146a downregulates NFκB activity via targeting TRAF6 and functions as a tumor suppressor having strong prognostic implications in NK/T cell lymphoma. *Clinical Cancer Research,* 17**,** 4761-4771.

PANKOV, R. & YAMADA, K. M. 2002. Fibronectin at a glance. *Journal of cell science,* 115**,** 3861-3863.

PAO, W., WANG, T. Y., RIELY, G. J., MILLER, V. A., PAN, Q., LADANYI, M., ZAKOWSKI, M. F., HEELAN, R. T., KRIS, M. G. & VARMUS, H. E. 2005. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS medicine,* 2**,** e17.

PAOLO, M., ASSUNTA, S., ANTONIO, R., PAOLA CLAUDIA, S., MARIA ANNA, B., CLORINDA, S., FRANCESCA, C., FORTUNATO, C. & CESARE, G. 2013. Selumetinib in Advanced Non Small Cell Lung Cancer (NSCLC) Harbouring KRAS Mutation: Endless Clinical Challenge to KRAS-mutant NSCLC. *Reviews on recent clinical trials,* 8**,** 93-100.

PAPAGEORGIOU, S. 2007. *HOX gene expression,* Austin,Texas, Landes Bioscience/Eurekah.com ; New York : Springer Science + Business Media.
PARAMASIVAM, M., SARKESHIK, A., YATES, J. R., 3RD, FERNANDES, M. J. & MCCOLLUM, D. 2011. Angiomotin family proteins are novel activators of the LATS2 kinase tumor suppressor. *Mol Biol Cell,* 22**,** 3725-33.
PAREYSON, D. 1999. Charcot-marie-tooth disease and related neuropathies: molecular basis for distinction and diagnosis. *Muscle Nerve,* 22**,** 1498-509.
PAREYSON, D. & MARCHESI, C. 2009. Diagnosis, natural history, and management of Charcot-Marie-Tooth disease. *Lancet Neurol,* 8**,** 654-67.
PASS, H. I., MITCHELL, J. B., JOHNSON, D. H., TURRISI, A. T. & MINNA, J. D. 2005. *Lung cancer: principles and practice*, Lippincott Williams & Wilkins.
PASTORE, G., ZNAOR, A., SPREAFICO, F., GRAF, N., PRITCHARD-JONES, K. & STELIAROVA-FOUCHER, E. 2006. Malignant renal tumours incidence and survival in European children (1978-1997): report from the Automated Childhood Cancer Information System project. *Eur J Cancer,* 42**,** 2103-14.
PAUL, D., BRIDOUX, L., REZSOHAZY, R. & DONNAY, I. 2011. HOX genes are expressed in bovine and mouse oocytes and early embryos. *Mol Reprod Dev,* 78**,** 436-49.
PEARSON, J. C., LEMONS, D. & MCGINNIS, W. 2005. Modulating Hox gene functions during animal body patterning. *Nat Rev Genet,* 6**,** 893-904.
PENG, X.-H., KARNA, P., CAO, Z., JIANG, B.-H., ZHOU, M. & YANG, L. 2006. Cross-talk between epidermal growth factor receptor and hypoxia-inducible factor-1α signal pathways increases resistance to apoptosis by up-regulating survivin gene expression. *Journal of Biological Chemistry,* 281**,** 25903-25914.
PEREZ-ORDONEZ, B. 2007. An update on Epstein-Barr virus and nasopharyngeal carcinogenesis. *Head and neck pathology,* 1**,** 141-145.
PETRINI, M., FELICETTI, F., BOTTERO, L., ERRICO, M. C., MORSILLI, O., BOE, A., DE FEO, A. & CARE, A. 2013. HOXB1 restored expression promotes apoptosis and differentiation in the HL60 leukemic cell line. *Cancer Cell Int,* 13**,** 101.
PETRUZZI, M. J. & GREEN, D. M. 1997. Wilms' tumor. *Pediatr Clin North Am,* 44**,** 939-52.
PINEAULT, N., HELGASON, C. D., LAWRENCE, H. J. & HUMPHRIES, R. K. 2002. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Exp Hematol,* 30**,** 49-57.
PLOWRIGHT, L., HARRINGTON, K. J., PANDHA, H. S. & MORGAN, R. 2009. HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). *Br J Cancer,* 100**,** 470-5.
POETA, M. L., MANOLA, J., GOLDWASSER, M. A., FORASTIERE, A., BENOIT, N., CALIFANO, J. A., RIDGE, J. A., GOODWIN, J., KENADY, D., SAUNDERS, J., WESTRA, W., SIDRANSKY, D. & KOCH, W. M. 2007. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. *N Engl J Med,* 357**,** 2552-61.
POMERANTZ, J., SCHREIBER-AGUS, N., LIEGEOIS, N. J., SILVERMAN, A., ALLAND, L., CHIN, L., POTES, J., CHEN, K., ORLOW, I., LEE, H. W., CORDON-CARDO, C. & DEPINHO, R. A. 1998. The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell,* 92**,** 713-23.
POSCHL, G. & SEITZ, H. K. 2004. Alcohol and cancer. *Alcohol Alcohol,* 39**,** 155-65.
PRADIDARCHEEP, W., LABRUYERE, W. T., DABHOIWALA, N. F. & LAMERS, W. H. 2008. Lack of specificity of commercially available antisera: better specifications needed. *J Histochem Cytochem,* 56**,** 1099-111.
PRESTRIDGE, D. S. 1995. Predicting Pol II promoter sequences using transcription factor binding sites. *J Mol Biol,* 249**,** 923-32.
PRIME, S. S., NIXON, S. V., CRANE, I. J., STONE, A., MATTHEWS, J. B., MAITLAND, N. J., REMNANT, L., POWELL, S. K., GAME, S. M. & SCULLY, C. 1990. The

behaviour of human oral squamous cell carcinoma in cell culture. *J Pathol,* 160**,** 259-69.
PROCHIANTZ, A. & JOLIOT, A. 2003. Can transcription factors function as cell–cell signalling molecules? *Nature Reviews Molecular Cell Biology,* 4**,** 814-819.
PROVENZANO, M. & MOCELLIN, S. 2007. Complementary techniques: validation of gene expression data by quantitative real time PCR. *Adv Exp Med Biol,* 593**,** 66-73.
PSYRRI, A., GOUVERIS, P. & VERMORKEN, J. B. 2009. Human papillomavirus-related head and neck tumors: clinical and research implication. *Curr Opin Oncol,* 21**,** 201-5.
QIAN, Y. Q., OTTING, G., BILLETER, M., MULLER, M., GEHRING, W. & WUTHRICH, K. 1993. Nuclear magnetic resonance spectroscopy of a DNA complex with the uniformly 13C-labeled Antennapedia homeodomain and structure determination of the DNA-bound homeodomain. *J Mol Biol,* 234**,** 1070-83.
RABEN, D., HELFRICH, B. & BUNN, P. A., JR. 2004. Targeted therapies for non-small-cell lung cancer: biology, rationale, and preclinical results from a radiation oncology perspective. *Int J Radiat Oncol Biol Phys,* 59**,** 27-38.
RAMACHANDRAN, S., LIU, P., YOUNG, A. N., YIN-GOEN, Q., LIM, S. D., LAYCOCK, N., AMIN, M. B., CARNEY, J. K., MARSHALL, F. F., PETROS, J. A. & MORENO, C. S. 2005. Loss of HOXC6 expression induces apoptosis in prostate cancer cells. *Oncogene,* 24**,** 188-98.
RAMAN, V., MARTENSEN, S. A., REISMAN, D., EVRON, E., ODENWALD, W. F., JAFFEE, E., MARKS, J. & SUKUMAR, S. 2000. Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature,* 405**,** 974-978.
RAMI-PORTA, R., CROWLEY, J. J. & GOLDSTRAW, P. 2009. The revised TNM staging system for lung cancer. *Ann Thorac Cardiovasc Surg,* 15**,** 4-9.
RANAHAN, W. P., HAN, Z., SMITH-KINNAMAN, W., NABINGER, S. C., HELLER, B., HERBERT, B. S., CHAN, R. & WELLS, C. D. 2011. The adaptor protein AMOT promotes the proliferation of mammary epithelial cells via the prolonged activation of the extracellular signal-regulated kinases. *Cancer Res,* 71**,** 2203-11.
RANCE, T. 1814. Case of fungus haematodes of the kidneys. *Med Phys J,* 32**,** 19-25.
RATUSHNY, V., ASTSATUROV, I., BURTNESS, B. A., GOLEMIS, E. A. & SILVERMAN, J. S. 2009. Targeting EGFR resistance networks in head and neck cancer. *Cellular signalling,* 21**,** 1255-1268.
RAUCH, T., WANG, Z., ZHANG, X., ZHONG, X., WU, X., LAU, S. K., KERNSTINE, K. H., RIGGS, A. D. & PFEIFER, G. P. 2007. Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay. *Proc Natl Acad Sci U S A,* 104**,** 5527-32.
RAWAT, V. P., THOENE, S., NAIDU, V. M., ARSENI, N., HEILMEIER, B., METZELER, K., PETROPOULOS, K., DESHPANDE, A., QUINTANILLA-MARTINEZ, L. & BOHLANDER, S. K. 2008. Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood,* 111**,** 309-319.
READ, C., JANES, S., GEORGE, J. & SPIRO, S. 2006. Early lung cancer: screening and detection. *Prim Care Respir J,* 15**,** 332-6.
RECK, M., HEIGENER, D. F., MOK, T., SORIA, J. C. & RABE, K. F. 2013. Management of non-small-cell lung cancer: recent developments. *Lancet,* 382**,** 709-19.
RED BREWER, M., YUN, C. H., LAI, D., LEMMON, M. A., ECK, M. J. & PAO, W. 2013. Mechanism for activation of mutated epidermal growth factor receptors in lung cancer. *Proc Natl Acad Sci U S A,* 110**,** E3595-604.
REDDY, S. D., OHSHIRO, K., RAYALA, S. K. & KUMAR, R. 2008. MicroRNA-7, a homeobox D10 target, inhibits p21-activated kinase 1 and regulates its functions. *Cancer Res,* 68**,** 8195-200.

REDLINE, R. W., HUDOCK, P., MACFEE, M. & PATTERSON, P. 1994. Expression of AbdB-type homeobox genes in human tumors. *Lab Invest,* 71**,** 663-70.

REDLINE, R. W., WILLIAMS, A. J., PATTERSON, P. & COLLINS, T. 1992. Human HOX4E: a gene strongly expressed in the adult male and female urogenital tracts. *Genomics,* 13**,** 425-30.

REED, A. L., CALIFANO, J., CAIRNS, P., WESTRA, W. H., JONES, R. M., KOCH, W., AHRENDT, S., EBY, Y., SEWELL, D., NAWROZ, H., BARTEK, J. & SIDRANSKY, D. 1996. High frequency of p16 (CDKN2/MTS-1/INK4A) inactivation in head and neck squamous cell carcinoma. *Cancer Res,* 56**,** 3630-3.

REYNOLDS, P. A., SIGAROUDINIA, M., ZARDO, G., WILSON, M. B., BENTON, G. M., MILLER, C. J., HONG, C., FRIDLYAND, J., COSTELLO, J. F. & TLSTY, T. D. 2006. Tumor suppressor p16INK4A regulates polycomb-mediated DNA hypermethylation in human mammary epithelial cells. *J Biol Chem,* 281**,** 24790-802.

RIELY, G. J., KRIS, M. G., ROSENBAUM, D., MARKS, J., LI, A., CHITALE, D. A., NAFA, K., RIEDEL, E. R., HSU, M., PAO, W., MILLER, V. A. & LADANYI, M. 2008. Frequency and distinctive spectrum of KRAS mutations in never smokers with lung adenocarcinoma. *Clin Cancer Res,* 14**,** 5731-4.

RIETHOVEN, J.-J. M. 2010. Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Computational Biology of Transcription Factor Binding.* Springer.

ROBERT, F., EZEKIEL, M. P., SPENCER, S. A., MEREDITH, R. F., BONNER, J. A., KHAZAELI, M. B., SALEH, M. N., CAREY, D., LOBUGLIO, A. F., WHEELER, R. H., COOPER, M. R. & WAKSAL, H. W. 2001. Phase I study of anti--epidermal growth factor receptor antibody cetuximab in combination with radiation therapy in patients with advanced head and neck cancer. *J Clin Oncol,* 19**,** 3234-43.

ROBERTSON, A., ALLEN, J., LANEY, R. & CURNOW, A. 2013. The cellular and molecular carcinogenic effects of radon exposure: a review. *Int J Mol Sci,* 14**,** 14024-63.

ROCCO, J. W., LI, D., LIGGETT, W., DUAN, L., SAUNDERS, J., SIDRANSKY, D. & O'MALLEY, B. 1998. p16INK4A adenovirus-mediated gene therapy for human head and neck squamous cell cancer. *Clinical cancer research,* 4**,** 1697-1704.

ROCHER, H. 1913. Les raideurs articulaires congénitales multiples. *I Med Bordeaux,* 84**,** 722.

RODINI, C. O., XAVIER, F. C. A., PAIVA, K. B. S., DE SOUZA SETÚBAL DESTRO, M. F., MOYSES, R. A., MICHALUARTE, P., CARVALHO, M. B., FUKUYAMA, E. E., GENCAPO, H. & TAJARA, E. H. 2012. Homeobox gene expression profile indicates HOXA5 as a candidate prognostic marker in oral squamous cell carcinoma. *International journal of oncology,* 40**,** 1180-1188.

RODRIGUEZ, T., ALTIERI, A., CHATENOUD, L., GALLUS, S., BOSETTI, C., NEGRI, E., FRANCESCHI, S., LEVI, F., TALAMINI, R. & LA VECCHIA, C. 2004. Risk factors for oral and pharyngeal cancer in young adults. *Oral Oncol,* 40**,** 207-13.

ROEMER, R., TAYLOR, A. & LARIVIERE, J. 2005. Origins of the WHO framework convention on tobacco control. *Journal Information,* 95.

ROGALLA, P., DRECHSLER, K., KAZMIERCZAK, B., RIPPE, V., BONK, U. & BULLERDIEK, J. 1997. Expression of HMGI-C, a member of the high mobility group protein family, in a subset of breast cancers: relationship to histologic grade. *Mol Carcinog,* 19**,** 153-6.

ROKITA, M., STEC, R., BODNAR, L., CHARKIEWICZ, R., KORNILUK, J., SMOTER, M., CICHOWICZ, M., CHYCZEWSKI, L., NIKLIŃSKI, J. & KOZŁOWSKI, W. 2013. Overexpression of epidermal growth factor receptor as a prognostic factor in colorectal cancer on the basis of the Allred scoring system. *OncoTargets and therapy,* 6**,** 967.

RONG, M., HE, R., DANG, Y. & CHEN, G. 2013. Expression and clinicopathological significance of miR-146a in hepatocellular carcinoma tissues. *Ups J Med Sci*.
ROSELL, R., MORAN, T., QUERALT, C., PORTA, R., CARDENAL, F., CAMPS, C., MAJEM, M., LOPEZ-VIVANCO, G., ISLA, D., PROVENCIO, M., INSA, A., MASSUTI, B., GONZALEZ-LARRIBA, J. L., PAZ-ARES, L., BOVER, I., GARCIA-CAMPELO, R., MORENO, M. A., CATOT, S., ROLFO, C., REGUART, N., PALMERO, R., SANCHEZ, J. M., BASTUS, R., MAYO, C., BERTRAN-ALAMILLO, J., MOLINA, M. A., SANCHEZ, J. J. & TARON, M. 2009. Screening for epidermal growth factor receptor mutations in lung cancer. *N Engl J Med,* 361**,** 958-67.
ROSENBERG, P. S., ALTER, B. P. & EBELL, W. 2008. Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica,* 93**,** 511-7.
ROSSI, A., MAIONE, P., BARESCHINO, M., SCHETTINO, C., SACCO, P., FERRARA, M., CASTALDO, V. & GRIDELLI, C. 2010. The emerging role of histology in the choice of first-line treatment of advanced non-small cell lung cancer: implication in the clinical decision-making. *Current medicinal chemistry,* 17**,** 1030-1038.
ROSSI DEGL'INNOCENTI, D., CASTIGLIONE, F., BUCCOLIERO, A. M., BECHI, P., TADDEI, G. L., FRESCHI, G. & TADDEI, A. 2007. Quantitative expression of the homeobox and integrin genes in human gastric carcinoma. *Int J Mol Med,* 20**,** 621-9.
ROTH, J. A., COX, J. D. & HONG, W. K. 2008. *Lung cancer,* Oxford, Blackwell.
ROYER-POKORA, B., BUSCH, M., BEIER, M., DUHME, C., DE TORRES, C., MORA, J., BRANDT, A. & ROYER, H.-D. 2010. Wilms tumor cells with WT1 mutations have characteristic features of mesenchymal stem cells and express molecular markers of paraxial mesoderm. *Human molecular genetics,* 19**,** 1651-1668.
RUBIN GRANDIS, J., MELHEM, M. F., GOODING, W. E., DAY, R., HOLST, V. A., WAGENER, M. M., DRENNING, S. D. & TWEARDY, D. J. 1998. Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. *J Natl Cancer Inst,* 90**,** 824-32.
RUPNIAK, H. T., ROWLATT, C., LANE, E. B., STEELE, J. G., TREJDOSIEWICZ, L. K., LASKIEWICZ, B., POVEY, S. & HILL, B. T. 1985. Characteristics of four new human cell lines derived from squamous cell carcinomas of the head and neck. *J Natl Cancer Inst,* 75**,** 621-35.
RUSSELL, P. J., HERTZ, P. E. & MCMILLAN, B. 2010. *Biology: The dynamic science*, Cengage Learning.
SAKURAI, Y., HATAKEYAMA, H., SATO, Y., HYODO, M., AKITA, H. & HARASHIMA, H. 2013. Gene Silencing via RNAi and siRNA Quantification in Tumor Tissue Using MEND, a Liposomal siRNA Delivery System. *Molecular Therapy*.
SALGADO, R. A., SNOECKX, A., SPINHOVEN, M., OP DE BEECK, B., CORTHOUTS, B. & PARIZEL, P. M. 2013. Update in non small-cell lung cancer staging. *JBR-BTR,* 96**,** 118-22.
SAMARAJEEWA, N. U., YANG, F., DOCANTO, M. M., SAKURAI, M., MCNAMARA, K. M., SASANO, H., FOX, S. B., SIMPSON, E. R. & BROWN, K. A. 2013. HIF-1alpha stimulates aromatase expression driven by prostaglandin E2 in breast adipose stroma. *Breast Cancer Res,* 15**,** R30.
SAMET, J. M., AVILA-TANG, E., BOFFETTA, P., HANNAN, L. M., OLIVO-MARSTON, S., THUN, M. J. & RUDIN, C. M. 2009. Lung cancer in never smokers: clinical epidemiology and environmental risk factors. *Clin Cancer Res,* 15**,** 5626-45.
SANTOS, E., MARTIN-ZANCA, D., REDDY, E. P., PIEROTTI, M. A., DELLA PORTA, G. & BARBACID, M. 1984. Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. *Science,* 223**,** 661-664.
SATO, M., TAKAHASHI, K., NAGAYAMA, K., ARAI, Y., ITO, N., OKADA, M., MINNA, J. D., YOKOTA, J. & KOHNO, T. 2005. Identification of chromosome arm 9p as the most frequent target of homozygous deletions in lung cancer. *Genes, Chromosomes and Cancer,* 44**,** 405-414.

SAUVAGEAU, G., LANSDORP, P. M., EAVES, C. J., HOGGE, D. E., DRAGOWSKA, W. H., REID, D. S., LARGMAN, C., LAWRENCE, H. J. & HUMPHRIES, R. K. 1994. Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proc Natl Acad Sci U S A,* 91**,** 12223-7.

SCARIA, P. V. & SHAFER, R. H. 1991. Binding of ethidium bromide to a DNA triple helix. Evidence for intercalation. *J Biol Chem,* 266**,** 5417-23.

SCHEFFNER, M., HUIBREGTSE, J. M., VIERSTRA, R. D. & HOWLEY, P. M. 1993. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell,* 75**,** 495-505.

SCHEFFNER, M., WERNESS, B. A., HUIBREGTSE, J. M., LEVINE, A. J. & HOWLEY, P. M. 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell,* 63**,** 1129-36.

SCHELTER, F., GRANDL, M., SEUBERT, B., SCHATEN, S., HAUSER, S., GERG, M., BOCCACCIO, C., COMOGLIO, P. & KRÜGER, A. 2011. Tumor cell-derived Timp-1 is necessary for maintaining metastasis-promoting Met-signaling via inhibition of Adam-10. *Clinical & experimental metastasis,* 28**,** 793-802.

SCHIESSL, B., INNES, B., BULMER, J., OTUN, H., CHADWICK, T., ROBSON, S. & LASH, G. 2009. Localization of angiogenic growth factors and their receptors in the human placental bed throughout normal human pregnancy. *Placenta,* 30**,** 79-87.

SCHMITTGEN, T. D. & LIVAK, K. J. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc,* 3**,** 1101-8.

SCHNABEL, C. L., WAGNER, S., WAGNER, B., DURAN, M. C., BABASYAN, S., NOLTE, I., PFARRER, C., FEIGE, K., MURUA ESCOBAR, H. & CAVALLERI, J. M. 2013. Evaluation of the reactivity of commercially available monoclonal antibodies with equine cytokines. *Vet Immunol Immunopathol,* 156**,** 1-19.

SCHNEIDER, M. R. & WOLF, E. 2009. The epidermal growth factor receptor ligands at a glance. *J Cell Physiol,* 218**,** 460-6.

SCHNEUWLY, S., KLEMENZ, R. & GEHRING, W. J. 1987. Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia. *Nature,* 325**,** 816-8.

SCHWARTZ, A. G., PRYSAK, G. M., BOCK, C. H. & COTE, M. L. 2007. The molecular epidemiology of lung cancer. *Carcinogenesis,* 28**,** 507-18.

SCIAN, M. J., STAGLIANO, K. E., ELLIS, M. A., HASSAN, S., BOWMAN, M., MILES, M. F., DEB, S. P. & DEB, S. 2004. Modulation of gene expression by tumor-derived p53 mutants. *Cancer Res,* 64**,** 7447-54.

SCOTT, M. P. 1992. Vertebrate homeobox gene nomenclature. *Cell,* 71**,** 551-553.

SCOTT, M. P. 1993. A rational nomenclature for vertebrate homeobox (HOX) genes. *Nucleic Acids Res,* 21**,** 1687-8.

SCOTT, M. P. & WEINER, A. J. 1984. Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. *Proceedings of the National Academy of Sciences,* 81**,** 4115-4119.

SENGUPTA, A. 2012. Recent advances in head and neck cancer. *Apollo Medicine,* 9**,** 96-103.

SEQUIST, L. V., WALTMAN, B. A., DIAS-SANTAGATA, D., DIGUMARTHY, S., TURKE, A. B., FIDIAS, P., BERGETHON, K., SHAW, A. T., GETTINGER, S., COSPER, A. K., AKHAVANFARD, S., HEIST, R. S., TEMEL, J., CHRISTENSEN, J. G., WAIN, J. C., LYNCH, T. J., VERNOVSKY, K., MARK, E. J., LANUTI, M., IAFRATE, A. J., MINO-KENUDSON, M. & ENGELMAN, J. A. 2011. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med,* 3**,** 75ra26.

SERRANO, M., LEE, H., CHIN, L., CORDON-CARDO, C., BEACH, D. & DEPINHO, R. A. 1996. Role of the INK4a locus in tumor suppression and cell mortality. *Cell,* 85**,** 27-37.

SETH, S., JOHNS, R. & TEMPLIN, M. V. 2012. Delivery and biodistribution of siRNA for cancer therapy: challenges and future prospects. *Ther Deliv,* 3**,** 245-61.

SEVERINO, P., BRUGGEMANN, H., ANDREGHETTO, F. M., CAMPS, C., KLINGBEIL, M. D., DE PEREIRA, W. O., SOARES, R. M., MOYSES, R., WUNSCH-FILHO, V., MATHOR, M. B., NUNES, F. D., RAGOUSSIS, J. & TAJARA, E. H. 2013. MicroRNA expression profile in head and neck cancer: HOX-cluster embedded microRNA-196a and microRNA-10b dysregulation implicated in cell proliferation. *BMC Cancer,* 13**,** 533.

SHAH, N. & SUKUMAR, S. 2010. The Hox genes and their roles in oncogenesis. *Nature Reviews Cancer,* 10**,** 361-371.

SHAH, V., DRILL, E. & LANCE-JONES, C. 2004. Ectopic expression of Hoxd10 in thoracic spinal segments induces motoneurons with a lumbosacral molecular profile and axon projections to the limb. *Dev Dyn,* 231**,** 43-56.

SHANMUGAM, K., FEATHERSTONE, M. S. & SARAGOVI, H. U. 1997. Residues flanking the HOX YPWM motif contribute to cooperative interactions with PBX. *J Biol Chem,* 272**,** 19081-7.

SHANMUGAM, K., GREEN, N. C., RAMBALDI, I., SARAGOVI, H. U. & FEATHERSTONE, M. S. 1999. PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. *Molecular and cellular biology,* 19**,** 7577-7588.

SHARAFINSKI, M. E., FERRIS, R. L., FERRONE, S. & GRANDIS, J. R. 2010. Epidermal growth factor receptor targeted therapy of squamous cell carcinoma of the head and neck. *Head Neck,* 32**,** 1412-21.

SHEARS, L., PLOWRIGHT, L., HARRINGTON, K., PANDHA, H. S. & MORGAN, R. 2008. Disrupting the interaction between HOX and PBX causes necrotic and apoptotic cell death in the renal cancer lines CaKi-2 and 769-P. *The Journal of urology,* 180**,** 2196-2201.

SHEN, J., AMBROSONE, C. B., DICIOCCIO, R. A., ODUNSI, K., LELE, S. B. & ZHAO, H. 2008. A functional polymorphism in the miR-146a gene and age of familial breast/ovarian cancer diagnosis. *Carcinogenesis,* 29**,** 1963-6.

SHEN, W. F., ROZENFELD, S., KWONG, A., KOM VES, L. G., LAWRENCE, H. J. & LARGMAN, C. 1999. HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. *Mol Cell Biol,* 19**,** 3051-61.

SHETH, R., BASTIDA, M. F., KMITA, M. & ROS, M. 2013. “Self-regulation,” a new facet of Hox genes' function. *Developmental Dynamics*.

SHI, J., CHATTERJEE, N., ROTUNNO, M., WANG, Y., PESATORI, A. C., CONSONNI, D., LI, P., WHEELER, W., BRODERICK, P. & HENRION, M. 2012. Inherited variation at chromosome 12p13. 33, including RAD52, influences the risk of squamous cell lung carcinoma. *Cancer discovery,* 2**,** 131-139.

SHIGEMATSU, H., LIN, L., TAKAHASHI, T., NOMURA, M., SUZUKI, M., WISTUBA, II, FONG, K. M., LEE, H., TOYOOKA, S., SHIMIZU, N., FUJISAWA, T., FENG, Z., ROTH, J. A., HERZ, J., MINNA, J. D. & GAZDAR, A. F. 2005. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst,* 97**,** 339-46.

SHIN, D. M., RO, J. Y., HONG, W. K. & HITTELMAN, W. N. 1994. Dysregulation of epidermal growth factor receptor expression in premalignant lesions during head and neck tumorigenesis. *Cancer Res,* 54**,** 3153-9.

SHINTANI, S., NAKAHARA, Y., MIHARA, M., UEYAMA, Y. & MATSUMURA, T. 2001. Inactivation of the p14< sup> ARF</sup>, p15< sup> INK4B</sup> and p16< sup> INK4A</sup> genes is a frequent event in human oral squamous cell carcinomas. *Oral oncology,* 37**,** 498-504.

SHRIMPTON, A. E., LEVINSOHN, E. M., YOZAWITZ, J. M., PACKARD, D. S., JR., CADY, R. B., MIDDLETON, F. A., PERSICO, A. M. & HOOTNICK, D. R. 2004. A HOX gene mutation in a family with isolated congenital vertical talus and Charcot-Marie-Tooth disease. *Am J Hum Genet,* 75**,** 92-6.

SIMEN-KAPEU, A., SURCEL, H.-M., KOSKELA, P., PUKKALA, E. & LEHTINEN, M. 2010. Lack of association between human papillomavirus type 16 and 18 infections and female lung cancer. *Cancer Epidemiology Biomarkers & Prevention,* 19**,** 1879-1881.

SKRZYPSKI, M., DZIADZIUSZKO, R., JASSEM, E., SZYMANOWSKA-NARLOCH, A., GULIDA, G., RZEPKO, R., BIERNAT, W., TARON, M., JELITTO-GORSKA, M., MARJANSKI, T., RZYMAN, W., ROSELL, R. & JASSEM, J. 2013. Main Histologic Types of Non-small-cell Lung Cancer Differ in Expression of Prognosis-related Genes. *Clin Lung Cancer*.

SLACK, J. M. W. 2006. *Essential developmental biology,* Oxford, Blackwell Publishing.

SMITH, E., LIN, C. & SHILATIFARD, A. 2011. The super elongation complex (SEC) and MLL in development and disease. *Genes Dev,* 25**,** 661-72.

SMITH, P. K., KROHN, R. I., HERMANSON, G. T., MALLIA, A. K., GARTNER, F. H., PROVENZANO, M. D., FUJIMOTO, E. K., GOEKE, N. M., OLSON, B. J. & KLENK, D. C. 1985. Measurement of protein using bicinchoninic acid. *Anal Biochem,* 150**,** 76-85.

SNEAD, C. M., SMITH, S. M., SADEGHEIN, N., LACRUZ, R. S., HU, P., KURTZ, I. & PAINE, M. L. 2011. Identification of a pH-responsive DNA region upstream of the transcription start site of human NBCe1-B. *Eur J Oral Sci,* 119 Suppl 1**,** 136-41.

SOBIN, L. H., GOSPODAROWICZ, M. K. & WITTEKIND, C. 2011. *TNM classification of malignant tumours*, Wiley. com.

SOSHNIKOVA, N. & DUBOULE, D. 2009. Epigenetic regulation of vertebrate Hox genes: a dynamic equilibrium. *Epigenetics,* 4**,** 537-40.

SOTTILE, J. & HOCKING, D. C. 2002. Fibronectin polymerization regulates the composition and stability of extracellular matrix fibrils and cell-matrix adhesions. *Mol Biol Cell,* 13**,** 3546-59.

STADLER, M. E., PATEL, M. R., COUCH, M. E. & HAYES, D. N. 2008. Molecular biology of head and neck cancer: risks and pathways. *Hematology/oncology clinics of North America,* 22**,** 1099-1124.

STELNICKI, E. J., KOMUVES, L. G., KWONG, A. O., HOLMES, D., KLEIN, P., ROZENFELD, S., LAWRENCE, H. J., ADZICK, N. S., HARRISON, M. & LARGMAN, C. 1998. HOX homeobox genes exhibit spatial and temporal changes in expression during human skin development. *J Invest Dermatol,* 110**,** 110-5.

STOTT, F. J., BATES, S., JAMES, M. C., MCCONNELL, B. B., STARBORG, M., BROOKES, S., PALMERO, I., RYAN, K., HARA, E. & VOUSDEN, K. H. 1998. The alternative product from the human CDKN2A locus, p14ARF, participates in a regulatory feedback loop with p53 and MDM2. *The EMBO Journal,* 17**,** 5001-5014.

STRANSKY, N., EGLOFF, A. M., TWARD, A. D., KOSTIC, A. D., CIBULSKIS, K., SIVACHENKO, A., KRYUKOV, G. V., LAWRENCE, M. S., SOUGNEZ, C. & MCKENNA, A. 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science,* 333**,** 1157-1160.

STRATHDEE, G., HOLYOAKE, T. L., SIM, A., PARKER, A., OSCIER, D. G., MELO, J. V., MEYER, S., EDEN, T., DICKINSON, A. M. & MOUNTFORD, J. C. 2007. Inactivation of HOXA genes by hypermethylation in myeloid and lymphoid malignancy is frequent and associated with poor prognosis. *Clinical Cancer Research,* 13**,** 5048-5055.

SU, P.-F., HUANG, W.-L., WU, H.-T., WU, C.-H., LIU, T.-Y. & KAO, S.-Y. 2010. < i> p16INK4A</i> promoter hypermethylation is associated with invasiveness and prognosis of oral squamous cell carcinoma in an age-dependent manner. *Oral oncology,* 46**,** 734-739.

SUN, M., SONG, C. X., HUANG, H., FRANKENBERGER, C. A., SANKARASHARMA, D., GOMES, S., CHEN, P., CHEN, J., CHADA, K. K., HE, C. & ROSNER, M. R. 2013. HMGA2/TET1/HOXA9 signaling pathway regulates breast cancer growth and metastasis. *Proc Natl Acad Sci U S A,* 110**,** 9920-5.

SUN, S., SCHILLER, J. H. & GAZDAR, A. F. 2007. Lung cancer in never smokers—a different disease. *Nature Reviews Cancer,* 7**,** 778-790.

SUTINEN, M., KAINULAINEN, T., HURSKAINEN, T., VESTERLUND, E., ALEXANDER, J. P., OVERALL, C. M., SORSA, T. & SALO, T. 1998. Expression of matrix metalloproteinases (MMP-1 and -2) and their inhibitors (TIMP-1, -2 and -3) in oral lichen planus, dysplasia, squamous cell carcinoma and lymph node metastasis. *Br J Cancer,* 77**,** 2239-45.

SYRIGOS, K., VANSTEENKISTE, J., PARIKH, P., VON PAWEL, J., MANEGOLD, C., MARTINS, R., SIMMS, L., SUGARMAN, K., VISSEREN-GRUL, C. & SCAGLIOTTI, G. 2010. Prognostic and predictive factors in a randomized phase III trial comparing cisplatin–pemetrexed versus cisplatin–gemcitabine in advanced non-small-cell lung cancer. *Annals of oncology,* 21**,** 556-561.

TABIN, C. 1995. The initiation of the limb bud: growth factors, Hox genes, and retinoids. *Cell,* 80**,** 671-4.

TAKAHASHI, O., HAMADA, J., ABE, M., HATA, S., ASANO, T., TAKAHASHI, Y., TADA, M., MIYAMOTO, M., KONDO, S. & MORIUCHI, T. 2007. Dysregulated expression of HOX and ParaHOX genes in human esophageal squamous cell carcinoma. *Oncol Rep,* 17**,** 753-60.

TAKAHASHI, Y., HAMADA, J., MURAKAWA, K., TAKADA, M., TADA, M., NOGAMI, I., HAYASHI, N., NAKAMORI, S., MONDEN, M., MIYAMOTO, M., KATOH, H. & MORIUCHI, T. 2004. Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. *Exp Cell Res,* 293**,** 144-53.

TAKES, R., BAATENBURG DE JONG, R., SCHUURING, E., LITVINOV, S., HERMANS, J. & VAN KRIEKEN, J. 1997. Differences in expression of oncogenes and tumor suppressor genes in different sites of head and neck squamous cell. *Anticancer research,* 18**,** 4793-4800.

TAKES, R. P., RINALDO, A., SILVER, C. E., PICCIRILLO, J. F., HAIGENTZ, M., JR., SUAREZ, C., VAN DER POORTEN, V., HERMANS, R., RODRIGO, J. P., DEVANEY, K. O. & FERLITO, A. 2010. Future of the TNM classification and staging system in head and neck cancer. *Head Neck,* 32**,** 1693-711.

TAKEUCHI, S., BARTRAM, C., SERIU, T., MILLER, C., TOBLER, A., JANSSEN, J., REITER, A., LUDWIG, W., ZIMMERMANN, M. & SCHWALLER, J. 1995. Analysis of a family of cyclin-dependent kinase inhibitors: p15/MTS2/INK4B, p16/MTS1/INK4A, and p18 genes in acute lymphoblastic leukemia of childhood. *Blood,* 86**,** 755-760.

TANIA, N., REBECCA, E. H. & STEPHEN, A. B. 2006. Quantification of mRNA using real-time RT-PCR. *Nat Protoc,* 1**,** 1559-1582.

TARCHINI, B., DUBOULE, D. & KMITA, M. 2006. Regulatory constraints in the evolution of the tetrapod limb anterior-posterior polarity. *Nature,* 443**,** 985-8.

TASHKIN, D. P. 2013. Effects of marijuana smoking on the lung. *Ann Am Thorac Soc,* 10**,** 239-47.

TAUBE, J. H., HERSCHKOWITZ, J. I., KOMUROV, K., ZHOU, A. Y., GUPTA, S., YANG, J., HARTWELL, K., ONDER, T. T., GUPTA, P. B., EVANS, K. W., HOLLIER, B. G., RAM, P. T., LANDER, E. S., ROSEN, J. M., WEINBERG, R. A. & MANI, S. A. 2010. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci U S A,* 107**,** 15449-54.

TAYLOR, H. S. 2000. The role of HOX genes in the development and function of the female reproductive tract. *Semin Reprod Med,* 18**,** 81-9.

TEMAM, S., FLAHAULT, A., PÉRIÉ, S., MONCEAUX, G., COULET, F., CALLARD, P., BERNAUDIN, J. F., ST GUILY, J. L. & FOURET, P. 2000. p53 gene status as a predictor of tumor response to induction chemotherapy of patients with locoregionally advanced squamous cell carcinomas of the head and neck. *Journal of clinical oncology,* 18**,** 385.

TERAI, M., HASHIMOTO, K., YODA, K. & SATA, T. 1999. High prevalence of human papillomaviruses in the normal oral cavity of adults. *Oral Microbiol Immunol,* 14**,** 201-5.

TESSEMA, M., WILLINK, R., DO, K., YU, Y. Y., YU, W., MACHIDA, E. O., BROCK, M., VAN NESTE, L., STIDLEY, C. A., BAYLIN, S. B. & BELINSKY, S. A. 2008. Promoter methylation of genes in and around the candidate lung cancer susceptibility locus 6q23-25. *Cancer Res,* 68**,** 1707-14.

THOMAS, S., QUINN, B. A., DAS, S. K., DASH, R., EMDAD, L., DASGUPTA, S., WANG, X. Y., DENT, P., REED, J. C., PELLECCHIA, M., SARKAR, D. & FISHER, P. B. 2013. Targeting the Bcl-2 family for cancer therapy. *Expert Opin Ther Targets,* 17**,** 61-75.

THOMPSON, R. C., VARDINOGIANNIS, I. & GILMORE, T. D. 2013. Identification of an NF-kappaB p50/p65-responsive site in the human MIR155HG promoter. *BMC Mol Biol,* 14**,** 24.

THORSTEINSDOTTIR, U., KROON, E., JEROME, L., BLASI, F. & SAUVAGEAU, G. 2001. Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. *Mol Cell Biol,* 21**,** 224-34.

THORSTEINSDOTTIR, U., SAUVAGEAU, G., HOUGH, M. R., DRAGOWSKA, W., LANSDORP, P. M., LAWRENCE, H. J., LARGMAN, C. & HUMPHRIES, R. K. 1997. Overexpression of HOXA10 in murine hematopoietic cells perturbs both myeloid and lymphoid differentiation and leads to acute myeloid leukemia. *Molecular and Cellular Biology,* 17**,** 495-505.

TIRET, L., LE MOUELLIC, H., MAURY, M. & BRULET, P. 1998. Increased apoptosis of motoneurons and altered somatotopic maps in the brachial spinal cord of Hoxc-8-deficient mice. *Development,* 125**,** 279-91.

TISCHFIELD, M. A., BOSLEY, T. M., SALIH, M. A., ALORAINY, I. A., SENER, E. C., NESTER, M. J., OYSTRECK, D. T., CHAN, W. M., ANDREWS, C., ERICKSON, R. P. & ENGLE, E. C. 2005. Homozygous HOXA1 mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. *Nat Genet,* 37**,** 1035-7.

TOSSAVAINEN, A. 2000. International expert meeting on new advances in the radiology and screening of asbestos-related diseases. *Scand J Work Environ Health,* 26**,** 449-54.

TOYOOKA, S., TSUDA, T. & GAZDAR, A. F. 2003. The TP53 gene, tobacco exposure, and lung cancer. *Hum Mutat,* 21**,** 229-39.

TRAVIS, W. D. 2004. *Pathology and genetics of tumours of the lung, pleura, thymus and heart,* Lyon ; [Great Britain], IARC Press.

TRAVIS, W. D. 2011. Pathology of Lung Cancer. *Clinics in chest medicine,* 32**,** 669-692.

TROYANOVSKY, B., LEVCHENKO, T., MÅNSSON, G., MATVIJENKO, O. & HOLMGREN, L. 2001. Angiomotin An Angiostatin Binding Protein That Regulates Endothelial Cell Migration and Tube Formation. *The Journal of cell biology,* 152**,** 1247-1254.

TRUMBO, T. A., SCHULTZ, E., BORLAND, M. G. & PUGH, M. E. 2013. Applied spectrophotometry: analysis of a biochemical mixture. *Biochem Mol Biol Educ,* 41**,** 242-50.

TSAI, L. H., CHEN, P. M., CHENG, Y. W., CHEN, C. Y., SHEU, G. T., WU, T. C. & LEE, H. 2013. LKB1 loss by alteration of the NKX2-1/p53 pathway promotes tumor malignancy and predicts poor survival and relapse in lung adenocarcinomas. *Oncogene*.

TSIM, S., O'DOWD, C., MILROY, R. & DAVIDSON, S. 2010. Staging of non-small cell lung cancer (NSCLC): a review. *Respiratory medicine,* 104**,** 1767-1774.

TU, H.-F., LIN, S.-C. & CHANG, K.-W. 2013. MicroRNA aberrances in head and neck cancer: pathogenetic and clinical significance. *Current opinion in otolaryngology & head and neck surgery,* 21**,** 104-111.

UNGER, M. A. 2002. Microarray analysis identifies HoxA9 as a novel breast cancer progression gene.

VALLAT, J. M., MATHIS, S. & FUNALOT, B. 2013. The various Charcot-Marie-Tooth diseases. *Curr Opin Neurol,* 26**,** 473-80.

VAN BOERDONK, R. A., DANIELS, J. M., BLOEMENA, E., KRIJGSMAN, O., STEENBERGEN, R. D., BRAKENHOFF, R. H., GRUNBERG, K., YLSTRA, B., MEIJER, C. J., SMIT, E. F., SNIJDERS, P. J. & HEIDEMAN, D. A. 2013. High-risk human papillomavirus-positive lung cancer: molecular evidence for a pattern of pulmonary metastasis. *J Thorac Oncol,* 8**,** 711-8.

VAN DER SCHROEFF, M. P. & BAATENBURG DE JONG, R. J. 2009. Staging and prognosis in head and neck cancer. *Oral Oncol,* 45**,** 356-60.

VAN DYKE, D. L., WORSHAM, M. J., DRUMHELLER, T., BENNINGER, M. S., KRAUSE, C. J., BAKER, S. R., WOLF, G. T., CAREY, T. E. & TILLEY, B. C. 1994. Recurrent cytogenetic abnormalities in squamous cell carcinomas of the head and neck region. *Genes, Chromosomes and Cancer,* 9**,** 192-206.

VAQUERIZAS, J. M., KUMMERFELD, S. K., TEICHMANN, S. A. & LUSCOMBE, N. M. 2009. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet,* 10**,** 252-63.

VARELLA-GARCIA, M. 2010. Chromosomal and genomic changes in lung cancer. *Cell Adh Migr,* 4**,** 100-6.

VASANTHI, D. & MISHRA, R. K. 2008. Epigenetic regulation of genes during development: a conserved theme from flies to mammals. *J Genet Genomics,* 35**,** 413-29.

VAUGHAN, C. A., SINGH, S., WINDLE, B., YEUDALL, W. A., FRUM, R., GROSSMAN, S. R., DEB, S. P. & DEB, S. 2012. Gain-of-Function Activity of Mutant p53 in Lung Cancer through Up-Regulation of Receptor Protein Tyrosine Kinase Axl. *Genes & cancer,* 3**,** 491-502.

VERMORKEN, J. B. & SPECENIER, P. 2010. Optimal treatment for recurrent/metastatic head and neck cancer. *Ann Oncol,* 21 Suppl 7**,** vii252-vii261.

VILLAVICENCIO-LORINI, P., KUSS, P., FRIEDRICH, J., HAUPT, J., FAROOQ, M., TURKMEN, S., DUBOULE, D., HECHT, J. & MUNDLOS, S. 2010. Homeobox genes d11-d13 and a13 control mouse autopod cortical bone and joint formation. *J Clin Invest,* 120**,** 1994-2004.

VINAGRE, T., MONCAUT, N., CARAPUÇO, M., NÓVOA, A., BOM, J. & MALLO, M. 2010. Evidence for a myotomal Hox/Myf cascade governing nonautonomous control of rib specification within global vertebral domains. *Developmental cell,* 18**,** 655-661.

VINEIS, P., FORASTIERE, F., HOEK, G. & LIPSETT, M. 2004. Outdoor air pollution and lung cancer: recent epidemiologic evidence. *Int J Cancer,* 111**,** 647-52.

VISWANATHAN, M., TSUCHIDA, N. & SHANMUGAM, G. 2003. Promoter hypermethylation profile of tumor-associated genes p16, p15, hMLH1, MGMT and E-cadherin in oral squamous cell carcinoma. *International journal of cancer,* 105**,** 41-46.

VITIELLO, D., KODAMAN, P. H. & TAYLOR, H. S. 2007. HOX genes in implantation. *Semin Reprod Med,* 25**,** 431-6.

VOLPE, M. V., ARCHAVACHOTIKUL, K., BHAN, I., LESSIN, M. S. & NIELSEN, H. C. 2000. Association of bronchopulmonary sequestration with expression of the homeobox protein Hoxb-5. *J Pediatr Surg,* 35**,** 1817-9.

VOLPE, M. V., PHAM, L., LESSIN, M., RALSTON, S. J., BHAN, I., CUTZ, E. & NIELSEN, H. C. 2003. Expression of Hoxb-5 during human lung development and in congenital lung malformations. *Birth Defects Res A Clin Mol Teratol,* 67**,** 550-6.

WAIN, H. M., BRUFORD, E. A., LOVERING, R. C., LUSH, M. J., WRIGHT, M. W. & POVEY, S. 2002. Guidelines for human gene nomenclature. *Genomics,* 79**,** 464-70.

WAKAI, K., NAGATA, C., MIZOUE, T., TANAKA, K., NISHINO, Y., TSUJI, I., INOUE, M. & TSUGANE, S. 2007. Alcohol drinking and lung cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population. *Japanese journal of clinical oncology,* 37**,** 168-174.

WAKELEE, H. A., CHANG, E. T., GOMEZ, S. L., KEEGAN, T. H., FESKANICH, D., CLARKE, C. A., HOLMBERG, L., YONG, L. C., KOLONEL, L. N., GOULD, M. K. & WEST, D. W. 2007. Lung cancer incidence in never smokers. *J Clin Oncol,* 25**,** 472-8.

WALBOOMERS, J. M., JACOBS, M. V., MANOS, M. M., BOSCH, F. X., KUMMER, J. A., SHAH, K. V., SNIJDERS, P. J., PETO, J., MEIJER, C. J. & MUNOZ, N. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol,* 189**,** 12-9.

WALTER, V., YIN, X., WILKERSON, M. D., CABANSKI, C. R., ZHAO, N., DU, Y., ANG, M. K., HAYWARD, M. C., SALAZAR, A. H., HOADLEY, K. A., FRITCHIE, K., SAILEY, C. G., WEISSLER, M. C., SHOCKLEY, W. W., ZANATION, A. M., HACKMAN, T., THORNE, L. B., FUNKHOUSER, W. D., MULDREW, K. L., OLSHAN, A. F., RANDELL, S. H., WRIGHT, F. A., SHORES, C. G. & HAYES, D. N. 2013. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. *PLoS One,* 8**,** e56823.

WALTREGNY, D., ALAMI, Y., CLAUSSE, N., DE LEVAL, J. & CASTRONOVO, V. 2002. Overexpression of the homeobox gene HOXC8 in human prostate cancer correlates with loss of tumor differentiation. *Prostate,* 50**,** 162-9.

WAN, Y. W., RAESE, R. A., FORTNEY, J. E., XIAO, C., LUO, D., CAVENDISH, J., GIBSON, L. F., CASTRANOVA, V., QIAN, Y. & GUO, N. L. 2012. A smoking-associated 7-gene signature for lung cancer diagnosis and prognosis. *Int J Oncol,* 41**,** 1387-96.

WANG, B. D., YANG, Q., CENICCOLA, K., BIANCO, F., ANDRAWIS, R., JARRETT, T., FRAZIER, H., PATIERNO, S. R. & LEE, N. H. 2013. Androgen receptor-target genes in african american prostate cancer disparities. *Prostate Cancer,* 2013**,** 763569.

WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008a. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.

WANG, K., YAMAMOTO, H., CHIN, J. R., WERB, Z. & VU, T. H. 2004. Epidermal growth factor receptor-deficient mice have delayed primary endochondral ossification because of defective osteoclast recruitment. *J Biol Chem,* 279**,** 53848-56.

WANG, L., CHEN, S., XUE, M., ZHONG, J., WANG, X., GAN, L., LAM, E. K., LIU, X., ZHANG, J. & ZHOU, T. 2012. Homeobox D10 gene, a candidate tumor suppressor, is downregulated through promoter hypermethylation and associated with gastric carcinogenesis. *Molecular Medicine,* 18**,** 389.

WANG, X., TANG, S., LE, S.-Y., LU, R., RADER, J. S., MEYERS, C. & ZHENG, Z.-M. 2008b. Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth. *PloS one,* 3**,** e2557.

WANG, Y., WEI, X., XIAO, X., HUI, R., CARD, J. W., CAREY, M. A., WANG, D. W. & ZELDIN, D. C. 2005. Arachidonic acid epoxygenase metabolites stimulate endothelial cell growth and angiogenesis via mitogen-activated protein kinase and phosphatidylinositol 3-kinase/Akt signaling pathways. *Journal of Pharmacology and Experimental Therapeutics,* 314**,** 522-532.

WARNAKULASURIYA, S., REIBEL, J., BOUQUOT, J. & DABELSTEEN, E. 2008. Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *J Oral Pathol Med,* 37**,** 127-33.

WARTH, A., MULEY, T., MEISTER, M., STENZINGER, A., THOMAS, M., SCHIRMACHER, P., SCHNABEL, P. A., BUDCZIES, J., HOFFMANN, H. & WEICHERT, W. 2012. The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification system of lung adenocarcinoma is a stage-independent predictor of survival. *Journal of clinical oncology,* 30**,** 1438-1446.

WEBER, A., BELLMANN, U., BOOTZ, F., WITTEKIND, C. & TANNAPFEL, A. 2002. INK4a-ARF alterations and p53 mutations in primary and consecutive squamous cell carcinoma of the head and neck. *Virchows Arch,* 441**,** 133-42.

WEINBERG, R. A. 1991. Tumor suppressor genes. *Science,* 254**,** 1138-46.

WILKENING, S., BERMEJO, J. L. & HEMMINKI, K. 2007. MDM2 SNP309 and cancer risk: a combined analysis. *Carcinogenesis,* 28**,** 2262-2267.

WILLINGHAM, M. C. 1999. Conditional epitopes: is your antibody always specific? *Journal of Histochemistry & Cytochemistry,* 47**,** 1233-1235.

WILMS, M. 1902. *Die mischgeschwuelste*, A. Georgi.

WOENCKHAUS, M., STOEHR, R., DIETMAIER, W., WILD, P. J., ZIEGLMEIER, U., FOERSTER, J., MERK, J., BLASZYK, H., PFEIFER, M. & HOFSTAEDTER, F. 2003. Microsatellite instability at chromosome 8p in non-small cell lung cancer is associated with lymph node metastasis and squamous differentiation. *International journal of oncology,* 23**,** 1357.

WRAY, J. A., SUGDEN, M. C., ZELDIN, D. C., GREENWOOD, G. K., SAMSUDDIN, S., MILLER-DEGRAFF, L., BRADBURY, J. A., HOLNESS, M. J., WARNER, T. D. & BISHOP-BAILEY, D. 2009. The epoxygenases CYP2J2 activates the nuclear receptor PPARalpha in vitro and in vivo. *PLoS One,* 4**,** e7421.

WU, C.-T., CHANG, Y.-L., SHIH, J.-Y. & LEE, Y.-C. 2005. The significance of estrogen receptor β in 301 surgically treated non–small cell lung cancers. *The Journal of thoracic and cardiovascular surgery,* 130**,** 979-986.

WU, X., CHEN, H., PARKER, B., RUBIN, E., ZHU, T., LEE, J. S., ARGANI, P. & SUKUMAR, S. 2006. HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer research,* 66**,** 9527-9534.

XIAO, F. M., CHEN, Z. Z., ZENG, X. P., BAI, Y. F., GUO, L. L. & LI, Y. F. 2013. [Role of homeobox gene A5 in multidrug resistance of human small cell lung cancer cells]. *Nan Fang Yi Ke Da Xue Xue Bao,* 33**,** 1665-8.

XU, B., FENG, N. H., LI, P. C., TAO, J., WU, D., ZHANG, Z. D., TONG, N., WANG, J. F., SONG, N. H. & ZHANG, W. 2010. A functional polymorphism in Pre-miR-146a gene is associated with prostate cancer risk and mature miR-146a expression in vivo. *The Prostate,* 70**,** 467-472.

XU, B., WANG, N., WANG, X., TONG, N., SHAO, N., TAO, J., LI, P., NIU, X., FENG, N., ZHANG, L., HUA, L., WANG, Z. & CHEN, M. 2012. MiR-146a suppresses tumor growth and progression by targeting EGFR pathway and in a p-ERK-dependent manner in castration-resistant prostate cancer. *Prostate,* 72**,** 1171-8.

YAMAMOTO, M., TAKAI, D. & YAMAMOTO, F. 2003. Comprehensive expression profiling of highly homologous 39 hox genes in 26 different human adult tissues by the modified systematic multiplex RT-pCR method reveals tissue-specific expression pattern that suggests an important role of chromosomal structure in the regulation of hox gene expression in adult tissues. *Gene Expr,* 11**,** 199-210.

YAMATOJI, M., KASAMATSU, A., YAMANO, Y., SAKUMA, K., OGOSHI, K., IYODA, M., SHINOZUKA, K., OGAWARA, K., TAKIGUCHI, Y. & SHIIBA, M. 2010. State of homeobox A10 expression as a putative prognostic marker for oral squamous cell carcinoma. *Oncology reports,* 23**,** 61-67.

YANG, F., HUANG, W., LI, Y., LIU, S., JIN, M., WANG, Y., JIA, L. & GAO, Z. 2013. Anti-tumor effects in mice induced by survivin-targeted siRNA delivered through polysaccharide nanoparticles. *Biomaterials*.

YANO, T., MIURA, N., TAKENAKA, T., HARO, A., OKAZAKI, H., OHBA, T., KOUSO, H., KOMETANI, T., SHOJI, F. & MAEHARA, Y. 2008. Never-smoking nonsmall cell lung cancer as a separate entity. *Cancer,* 113**,** 1012-1018.

YARBROUGH, W. G., BESSHO, M., ZANATION, A., BISI, J. E. & XIONG, Y. 2002. Human tumor suppressor ARF impedes S-phase progression independent of p53. *Cancer research,* 62**,** 1171-1177.

YARDEN, Y. 2001. The EGFR family and its ligands in human cancer: signalling mechanisms and therapeutic opportunities. *European journal of cancer,* 37**,** 3-8.

YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T. L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics,* 13**,** 134.

YI, C., TROUTMAN, S., FERA, D., STEMMER-RACHAMIMOV, A., AVILA, J. L., CHRISTIAN, N., PERSSON, N. L., SHIMONO, A., SPEICHER, D. W., MARMORSTEIN, R., HOLMGREN, L. & KISSIL, J. L. 2011. A tight junction-associated Merlin-angiomotin complex mediates Merlin's regulation of mitogenic signaling and tumor suppressive functions. *Cancer Cell,* 19**,** 527-40.

YOSHIDA, R., NAGATA, M., NAKAYAMA, H., NIIMORI-KITA, K., HASSAN, W., TANAKA, T., SHINOHARA, M. & ITO, T. 2013. The pathological significance of Notch1 in oral squamous cell carcinoma. *Lab Invest,* 93**,** 1068-81.

YOU, M., WANG, D., LIU, P., VIKIS, H., JAMES, M., LU, Y., WANG, Y., WANG, M., CHEN, Q. & JIA, D. 2009. Fine mapping of chromosome 6q23-25 region in familial lung cancer families reveals RGS17 as a likely candidate gene. *Clinical cancer research,* 15**,** 2666-2674.

YOUSSEF, E. M., LOTAN, D., ISSA, J. P., WAKASA, K., FAN, Y. H., MAO, L., HASSAN, K., FENG, L., LEE, J. J., LIPPMAN, S. M., HONG, W. K. & LOTAN, R. 2004. Hypermethylation of the retinoic acid receptor-beta(2) gene in head and neck carcinogenesis. *Clin Cancer Res,* 10**,** 1733-42.

YUEN, P. W., MAN, M., LAM, K. Y. & KWONG, Y. L. 2002. Clinicopathological significance of p16 gene expression in the surgical treatment of head and neck squamous cell carcinomas. *J Clin Pathol,* 55**,** 58-60.

ZACCHETTI, G., DUBOULE, D. & ZAKANY, J. 2007. Hox gene function in vertebrate gut morphogenesis: the case of the caecum. *Development,* 134**,** 3967-73.

ZAIN, R. B., IKEDA, N., GUPTA, P. C., WARNAKULASURIYA, S., WYK, C. W., SHRESTHA, P. & AXÉLL, T. 1999. Oral mucosal lesions associated with betel quid, areca nut and tobacco chewing habits: consensus from a workshop held in Kuala Lumpur, Malaysia, November 25-27, 1996. *Journal of oral pathology & medicine,* 28**,** 1-4.

ZAKANY, J. & DUBOULE, D. 2007. The role of Hox genes during vertebrate limb development. *Curr Opin Genet Dev,* 17**,** 359-66.

ZELTSER, L., DESPLAN, C. & HEINTZ, N. 1996. Hoxb-13: a new Hox gene in a distant region of the HOXB cluster maintains colinearity. *Development,* 122**,** 2475-84.

ZHANG, H., WANG, Y. S., HAN, G. & SHI, Y. 2007. TIMP-3 gene transfection suppresses invasive and metastatic capacity of human hepatocarcinoma cell line HCC-7721. *Hepatobiliary Pancreat Dis Int,* 6**,** 487-91.

ZHANG, M., DAI, C., ZHU, H., CHEN, S., WU, Y., LI, Q., ZENG, X., WANG, W., ZUO, J., ZHOU, M., XIA, Z., JI, G., SAIYIN, H., QIN, L. & YU, L. 2011a. Cyclophilin A promotes human hepatocellular carcinoma cell metastasis via regulation of MMP3 and MMP9. *Mol Cell Biochem,* 357**,** 387-95.

ZHANG, M. Q. 2003. Prediction, annotation, and analysis of human promoters. *Cold Spring Harb Symp Quant Biol,* 68**,** 217-25.

ZHANG, T. H., LIU, H. C., ZHU, L. J., CHU, M., LIANG, Y. J., LIANG, L. Z. & LIAO, G. Q. 2011b. Activation of Notch signaling in human tongue carcinoma. *J Oral Pathol Med,* 40**,** 37-45.

ZHANG, W., SHMULEVICH, I. & ASTOLA, J. 2004. *Microarray quality control*, Wiley-Liss.

ZHANG, W. C., LIU, J., XU, X. & WANG, G. 2013a. The role of microRNAs in lung cancer progression. *Med Oncol,* 30**,** 675.

ZHANG, X., ZHU, T., CHEN, Y., MERTANI, H. C., LEE, K.-O. & LOBIE, P. E. 2003. Human growth hormone-regulated HOXA1 is a human mammary epithelial oncogene. *Journal of Biological Chemistry,* 278**,** 7580-7590.

ZHANG, Y., CHENG, J. C., HUANG, H. F. & LEUNG, P. C. 2013b. Homeobox A7 stimulates breast cancer cell proliferation by up-regulating estrogen receptor-alpha. *Biochem Biophys Res Commun,* 440**,** 652-7.

ZHANG, Z., LEE, J. C., LIN, L., OLIVAS, V., AU, V., LAFRAMBOISE, T., ABDEL-RAHMAN, M., WANG, X., LEVINE, A. D., RHO, J. K., CHOI, Y. J., CHOI, C. M., KIM, S. W., JANG, S. J., PARK, Y. S., KIM, W. S., LEE, D. H., LEE, J. S., MILLER, V. A., ARCILA, M., LADANYI, M., MOONSAMY, P., SAWYERS, C., BOGGON, T. J., MA, P. C., COSTA, C., TARON, M., ROSELL, R., HALMOS, B. & BIVONA, T. G. 2012. Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. *Nat Genet,* 44**,** 852-60.

ZHANG, Z. F., MORGENSTERN, H., SPITZ, M. R., TASHKIN, D. P., YU, G. P., HSU, T. C. & SCHANTZ, S. P. 2000. Environmental tobacco smoking, mutagen sensitivity, and head and neck squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev,* 9**,** 1043-9.

ZHAO, B., LI, L., LU, Q., WANG, L. H., LIU, C.-Y., LEI, Q. & GUAN, K.-L. 2011. Angiomotin is a novel Hippo pathway component that inhibits YAP oncoprotein. *Genes & development,* 25**,** 51-63.

ZHAO, C. Y., GRINKEVICH, V. V., NIKULENKOV, F., BAO, W. & SELIVANOVA, G. 2010. Rescue of the apoptotic-inducing function of mutant p53 by small molecule RITA. *Cell Cycle,* 9**,** 1847-55.

ZHAO, X., SUN, M., ZHAO, J., LEYVA, J. A., ZHU, H., YANG, W., ZENG, X., AO, Y., LIU, Q., LIU, G., LO, W. H., JABS, E. W., AMZEL, L. M., SHAN, X. & ZHANG, X. 2007. Mutations in HOXD13 underlie syndactyly type V and a novel brachydactyly-syndactyly syndrome. *Am J Hum Genet,* 80**,** 361-71.

ZHENG, Q. & WANG, X. J. 2008. GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Res,* 36**,** W358-63.

ZHOU, Q., ZHANG, X. C., CHEN, Z. H., YIN, X. L., YANG, J. J., XU, C. R., YAN, H. H., CHEN, H. J., SU, J., ZHONG, W. Z., YANG, X. N., AN, S. J., WANG, B. C., HUANG, Y. S., WANG, Z. & WU, Y. L. 2011. Relative abundance of EGFR mutations predicts benefit from gefitinib treatment for advanced non-small-cell lung cancer. *J Clin Oncol,* 29**,** 3316-21.

ZIEGLER, U. & GROSCURTH, P. 2004. Morphological features of cell death. *News Physiol Sci,* 19**,** 124-8.

APPENDICES

Appendix 1: Lab equipment.

Equipment	Manufacturer
Applied Biosystems 7900HT Fast Real-Time PCR System	Life technologies
Avanti J-26 XP Centrifuge	Beckman Coulter
Class II Safety Cabinet	Walker
Compact X4 Automatic X-ray Film Processor	Xograph
Deltawave microwave	Toshiba
DNA Engine Thermal Cycler	Dyad
Dri-Block Heater	Techne
Electromagnetic stirrer	FALC
G:Box gel imaging system	Syngene
Galaxy CO2 incubators	Eppendorf
High-Speed Centrifuge	Sigma-Aldrich
Inverted Tissue Culture Microscope	AmScope
Microplate spectrophotometer	Tecan
Microtube centrifuge	Sigma-Aldrich
Mini-PROTEAN Tetra Electrophoresis System	Bio-Rad
NanoDrop 1000	Thermo
PureLab Option Water Deionizer	Elga
Shaking Incubator	Labnet
Slide Stainer	Leica
Sorvall Legend XT centrifuge	Thermo
Spiramix roller mixer	Denley
SteriMate Autoclave	Astell
Sub-Cell GT System	Bio-Rad
2100 Bioanalyzer instrument	Agilent
Microarray C Scanner	Agilent
GloMax Microplate Luminometer	Promega

Appendix 2: Buffers, solutions and dyes.

Buffer	Ingredients
Tris Buffered Saline with Tween-20 (TBST) buffer	19 mM Tris-base, 2.7 mM KCl, 137 mM NaCl, 0.1% Tween-20, (pH 7)
Lower Tris buffer	1.5 M Tris-base, 0.4% SDS, (pH 8.8)
Upper Tris uffer	0.5 M Tris-base, 0.4% SDS, (pH 6.8)
SDS-PAGE running buffer	250 mM Tris, 1.92 M glycine, 1% SDS
WB transfer buffer	25% Methanol, 25 mM Tris, 195 mM Glycine.
5% Blocking buffer	5% skimmed milk powder in TBST
Tris Acetate-EDTA buffer (TAE)	40 mM Tris-acetate, 0.01% glacial acetic acid, 1mM EDTA, (pH 8.0)
2x SDS protein sample loading dye	1 M DTT, 20% SDS, 0.02% Glycerol, 0.5 M Tris-HCl, 0.01% bromophenol blue dye, (pH to 6.8)
0.1% Crystal Violet dye	2.45 mM Crystal Violet powder
IHC washing PBS buffer	9.7 mM Dipotassium Phosphate, 145 mM Sodium Chloride, 1.84 mM Potassium Dihydrogen Phosphate
Basic antigen retrieval buffer	10 mM Tris-base, 1.26 mM EDTA, 0.1% Tween-20 (pH 9)
Acidic antigen retrieval buffer	10 mM Sodium citrate dehydrate (pH 6)
0.1% Acid alcohol	0.1% 0.1M Hydrochloric acid, 70% Isopropanol
Scott's tap water	41.7 mM Sodium bicarbonate, 166.2 mM Magnesium Sulphate

N.B: Buffers and solutions were completed to their final volumes using dH_2O, unless otherwise stated.

Appendix 3: Cell culture media.

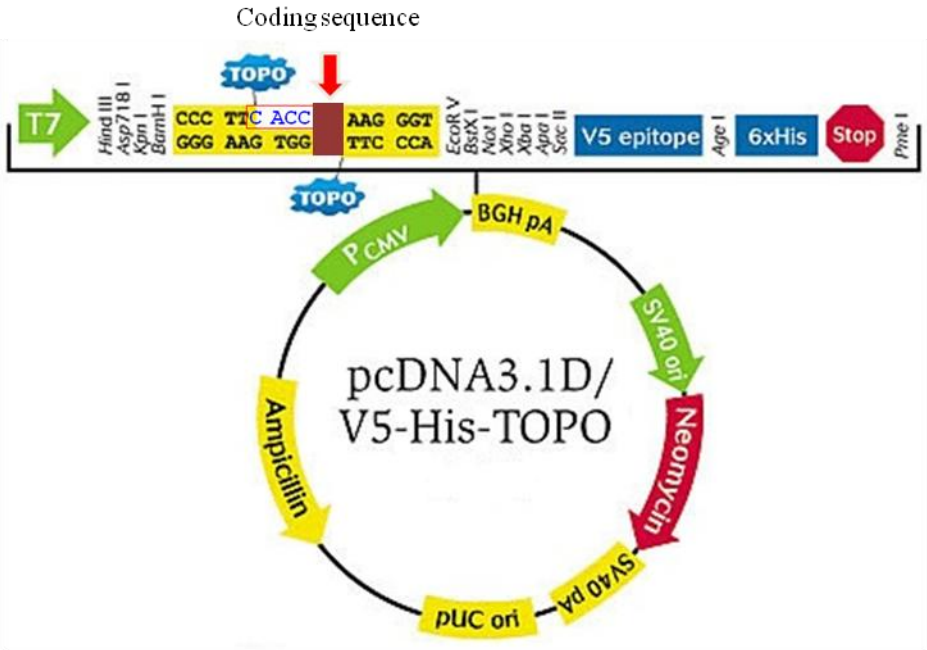
Medium	Supplements	Final concentration/ percentage
KGM	Dulbecco's modified Eagle's medium (DMEM)	67%
	Ham's F-12 media	27%
	Foetal Calf Serum (FCS)	10%
	L-glutamine	2 mM/mL
	Adenine	$1.8\times e^{-04}$ M
	Hydrocortisone	0.5 µg/mL
	Epidermal growth factor (EGF)	10 ng/mL
	Insulin	5 µg/mL
RPMI	RPMI-1640	89%
	Foetal Calf Serum (FCS)	10%
	Hydrocortisone	0.5 µg/mL
	L-glutamine	2 mM/mL

Appendix 4: Primers designed in house for SYBR® Green-based qPCR reaction.

Symbol	Gene name	NCBI Accession #	Forward primer (5'→ 3')	Reverse primer (5'→ 3')
ACTG2	*actin, gamma 2, smooth muscle, enteric*	NM_001615	TTTGTGACCACAGCTGAGAGAGAAATTGTGC	GAAGAGGAAGCTGCTGTGGCCATCTCAT
ALDH1A2	*aldehyde dehydrogenase 1 family, member A2*	NM_003888	GTACACCAAGATCTTTATAAACAACG	TTGTCTATATCTGCCTTGTCTGCT
ALDH3A1	*aldehyde dehydrogenase 3 family, member A1*	NM_001135168	ACCTGCACAAGAATGAATGGAAC	TCAGGGAGCTTCTGGATCATGTA
ALKBH8	*alkB, alkylation repair homolog 8 (E. coli)*	NM_138775	ATCTGGGGGTCTTCCTGACATT	AATATGAGCGGGAATTCCTTGC
AMD1	*adenosylmethionine decarboxylase 1*	NM_001634	GTCACTCGTGAGAGTGGAATTCGTGA	CATTCCATTCATCGAATACCCACAAGG
AMOT	*angiomotin*	NM_133265	CATGGAGGGCAGGATTAAGACC	CGACAGCTGCTCTGTCTTGCT
APAF1	*apoptotic peptidase activating factor 1*	NM_181861	TGGAATAACTTCGTATGTAAGGAC	CTTTCAATTTGGAGAGCTTCT
AQP3	*aquaporin 3*	NM_004925	GCTGTATTATGATGCAATCTGGC	CTGTGCCTATGAACTGGTCAAAG
ARNT2	*aryl-hydrocarbon receptor nuclear translocator 2*	NM_014862	ATGACAGGCCGGATCTTGGAC	CCACACCTCATCCTGCAGATGA
BIRC6	*baculoviral IAP repeat containing 6*	NM_016252	TTTATCATCAGCCTGCCTCATCT	CGTGTTCAGACCAAGGTTCATC
BST2	*bone marrow stromal cell antigen 2*	NM_004335	CAACCACACTGTGATGGCCCTAATGG	AGTGATCTCTCCCTCAAGCTCCTCCAC
CADM1	*cell adhesion molecule 1*	NM_001098517	TCCTCTACAAGGCTTAACCCGGG	TACCATCACAGGCTGGGGCTT
CALM3	*calmodulin 3 (phosphorylase kinase, delta)*	NM_005184	CAGATGGGAACGGGACCATTGAC	ACGCCTCTCGGATCTCCTCCTCA
CCR10	*chemokine (C-C motif) receptor 10*	NM_016602	TTCCTGGCCTGTATCAGCGCCG	AGCGCCAGGAGCAGTGACAGCAG
CD24	*cluster of differentiation 24*	NM_013230	TGCAGAAGAGAGAGTGAGACCAC	AAATCCAACTAATGCCACCACC
CDC42	*cell division cycle 42*	NM_044472	TGTTTGATGAGGCTATCCTAGC	GAAGGGAAGGAGAAAACAGTTTAG
CHST10	*carbohydrate sulfotransferase 10*	NM_004854	TCTAAATGGAGCATTTTCTTCC	ATTTCTGCATCACTGAAGGAA
CLIC3	*chloride intracellular channel 3*	NM_004669	CTGCACATCGTCGACACGGTGTG	CCTGCATCGCGCTGTCCAGGTA
CNTN1	*contactin 1*	NM_001843	TGCTGCACCAAATGTGGCTCC	CTCTTGACAAAGGCGCCCATGTT
CRIP1	*cysteine-rich protein 1*	NM_001311	CACGCTGAGCACGAAGGCAAACC	CGCCCCGCCCAAAGCCTTTAG
CXCL9	*chemokine (C-X-C motif) ligand 9*	NM_002416	ATTGAAATCATTGCTACACTGAAGAATGG	TCTTTTGGCTGACCTGTTTCTCCC
CXCR4	*chemokine (C-X-C motif) receptor 4*	NM_001008540	CCAAGGAAGCTGTTGGCTGAAA	TGTCATCTGCCTCACTGACGTTG
CYP2J2	*cytochrome P450, family 2, subfamily J, polypeptide 2*	NM_000775	AACGGACAGCCTTTTGACCCTC	CAAAGCGTTCTCCGAAGGTGATG
DMBT1	*deleted in malignant brain tumors 1*	NM_007329	ATCTGCTCAGCCACCCAAATAA	AAGCCACCACAATTTGAAGAGG
DNAH5	*dynein, axonemal, heavy chain 5*	NM_001369	ACTAAATACCAGGGCATTGTG	GCAACTCGTTATGAAGGTCAT
EDNRA	*endothelin receptor type A*	NM_001957	GTGTTGACAGGTACAGAGCAGTTGCCTCCT	TCAGGAATGGCCAGGATAAAGGACAGGATC
EGLN1	*egl nine homolog 1 (C. elegans)*	NM_022051	GGGATGCCAAGGTAAGTGGAGGTA	TGCTGGTTGTACTTCATGAGGGTTG
EYA2	*eyes absent homolog 2 (Drosophila)*	NM_005244	ACCTTCCACACCAGCGAAAGA	CACGAACACACGCTCAATCTCAT
FAS	*Fas (TNF receptor superfamily, member 6)*	NM_000043	TGTTTGGGTGAAGAGAAAGGAA	TGCCACTGTTTCAGGATTTAAAG
FMN2	*forminutes 2*	NM_020066	GACCTTTTTCAGGCCTCACAGA	CCTGCTTCAACTTCACAGGCT
FOXA2	*orkhead box A2*	NM_021784	CGTACATGAGCATGTCGGCGGCC	TCATGCCAGCGCCCACGTACG
FOXQ1	*forkhead box Q1*	NM_033260	AAGCAGGGCAGTGACCTGGAGGGC	CTGTTCGCCGTCGCCCTGCGTAT
GBP2	*guanylate binding protein 2, interferon-inducible*	NM_004120	TAAACTTCAGGAACAGGAACG	GAGTATGTTACATATTGGCTCCA
GREB1	*growth regulation by estrogen in breast cancer 1*	NM_014668	CCGTCACCGGAAAATACCAAG	AATGTCACCCTCTGGACAGAG G
HOXA11	*homeobox A11*	NM_005523	ACAAGGCCGGCGGCTCCAGTGG	GAGCATGCGGGACAGTTGCAGGCG
HOXC10	*homeobox C10*	NM_017409	TAACGAAGCGAAAGAGGAGATAAAGGC	GCTCTCGCGTCAAATACATATTGAACAG
HTATIP2	*HIV-1 Tat interactive protein 2*	NM_006410	CAAGTTAAGGGAGAAGTAGAAGC	ATAACAGAACTCCAGGCCTAAAT
IFI44	*interferon-induced protein 44*	NM_006417	GGGAGTTGGTAAACGCTGGTGTGGTACAT	TGGACTTCCTCTAGCTTGGACCTCACAGG
IGFBP3	*insulin-like growth factor binding protein 3*	NM_001013398	GACAGAATATGGTCCCTGCCGTA	AAGGGCGACACTGCTTTTTCTTA
IL18R1	*interleukin 18 receptor 1*	NM_003855	CAATAGTGGAAGATCGCAGTAAT	CTTACGTTTTTTCCTAATTCCACT
IL19	*interleukin 19*	NM_153758	TCTGCGGCAATGTCAGGAACAG	GCGTGGACCTCCAGCTGATCA

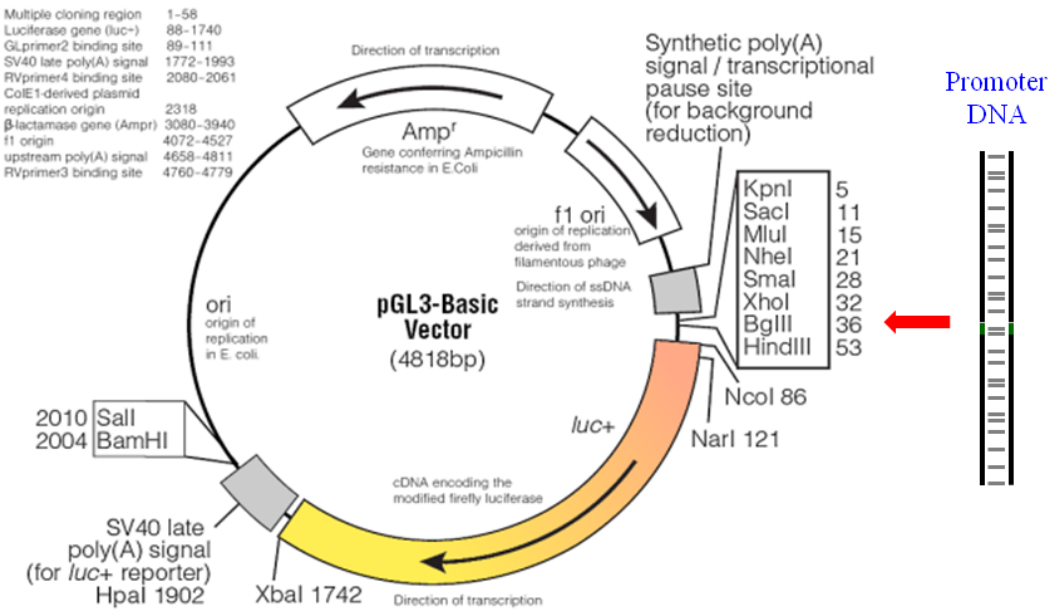
IRAK2	*interleukin-1 receptor-associated kinase 2*	NM_001570	GCTGTGGAATAGTGTTGGCCG	AAGTCCTTCAGGTAAACCGGGC
ITGA5	*integrin, alpha 5 (fibronectin receptor, alpha polypeptide)*	NM_002205	ACAGCTACCTAGGATACTCTGTGG	CATTAAGGATGGTGACATAGCC
ITGB2	*integrin, beta 2*	NM_000211	GAGCAACGAATTCGACTACCCATC	CGGTGAGTTTCTCGTAGGTCTTCAC
JAM2	*junctional adhesion molecule 2*	NM_021219	CCAAGAGGCTATTTTAGCCTGCAAAA	CAGCTCGATTTTTAAAATCACCTTGAAGAG
L1CAM	*L1 cell adhesion molecule*	NM_000425	TCTACGTTGTCCAGCTGCCAGC	CATCCTCGTCCAGCCACTGAAC
LAMA4	*aminin, alpha 4*	NM_001105206	GCCAAGAACTGTGCAGTGTGCAAC	GCAGTCCATGCCTGTAGGGGG
MMP11	*matrix metallopeptidase 11*	NM_001105206	ATTTGGTTCTTCCAAGGTGCTCAGTA	GAAGTAGATCTTGTTCTTCTCGGGACC
NANOG	*Nanog homeobox*	NM_024865	CTACAAACAGGTGAAGACCTGG	GTAGGAAGAGTAAAGGCTGGGG
NANOS1	*nanos homolog 1 (Drosophila)*	NM_199461	GAGCTGCAGGTGTGCGTGTTCTG	AGAGCGGGCAGTACTTGATGGTGTG
NTS	*neurotensin*	NM_006183	TCTGTGCTCAGATTCAGAAGA	CATGTGCTTTACTAATCTTTGATG
ORAI1	*ORAI calcium release-activated calcium modulator 1*	NM_032790	GACCACCATCATGGTGCCCTTCGGC	CGCCAGCTCGTTGAGCTCCTGGAAC
PDGFRL	*platelet-derived growth factor receptor-like*	NM_006207	AAGTGGGGACGACATCAGTGTGCTC	TCACAGGCCTTTCATCCTTCTGCCCT
PDPN	*podoplanin*	NM_198389	GGCCGCGGTGCTTTTTAATTT	TTCCCAAAACGAAGAGCAGAGCT
PEBP4	*phosphatidylethanolamine-binding protein 4*	NM_144962	AGGAGTTATCAGCCTACCAGG	GAGAGAGATGACTTTTCCTTCCT
PF4	*platelet factor 4*	NM_002619	CCTCGGTGTCCACTTCAGGCTTCC	CTCAGTGCGATGGGAAACTCGGG
PHF19	*PHD finger protein 19*	NM_001009936	CCCTGCGATGGGTGGATGTGGT	CTGGGGTGCTGGTGAGCTTGCC
PIR	*pirin (iron-binding nuclear protein)*	NM_003662	GAGAACAAGGATCCCAAGAGA	TTCATCACAAATGGACCATGT
PLXND1	*plexin D1*	NM_015103	CTGATGCGCACAGATACCAGCATC	GGTTCTGCATGTACCAGAAGGTGAGG
PPARGC1A	*peroxisome proliferator-activated receptor gamma, coactivator 1 alpha*	NM_013261	GAAGGCAATTGAAGAGCGCCG	TAGCTGTCTCCATCATCCCGCAG
PTCH1	*patched 1*	NM_001083602	CATGTACAACAGGCAGTGGAAA	TGATCCATGTAACCTGTTTCTGTG
RAC2	*ras-related C3 botulinum toxin substrate 2*	NM_002872	CCTGGAGTGCTCAGCTCTCACCCAG	TAGAGGAGGCTGCAGGCGCGCTT
RHOC	*ras homolog gene family, member C*	NM_175744	CCTGACAGCCTGGAAAACATTCCT	GAACGGGCTCCTGCTTCATCTTG
RUNX3	*runt-related transcription factor 3*	NM_001031680	TGACACCGAGCACACCCAGCC	CAGTTCCGAGGTGCCTTGGATTGG
SOD2	*superoxide dismutase 2, mitochondrial*	NM_001024465	CGTTGGCCAAGGGAGATGTTACAGCC	CCAGCAACTCCCCTTTGGGTTCTCC
STAT3	*signal transducer and activator of transcription 3*	NM_213662	AGAGAGTGCAGGATCTAGAACAG	GATCTTGCATGTCTCCTTGAC
TF	*transferrin*	NM_001063	CTTGAGAAAGCAGTGGCCAATTTC	CACCCTGGACACAGTTGACACA
TNFSF12	*tumor necrosis factor (ligand) superfamily, member 12*	NM_003809	AGGAATTCTCAGCCACTGCGGC	AGGAAGGGGGCAGCCTTGAGAT
TRIM17	*tripartite motif containing 17*	NM_001134855	CACGAGCCCCTCAAGCTTTTCT	CCTCCAGCTTCAACTTGTACCCCT
TUSC3	*tumor suppressor candidate 3*	NM_006765	TTATTCTGGTACTGAATGCCGCTA	AATCCCACTAGGCAAATTATCCGT
USP18	*ubiquitin specific peptidase 18*	NM_017414	GCTTCAATGACTCCAATATTTGCTTGGTGTCC	AAACCAGAAGATATGCAGTTTCCTGCCAGT
USP6	*ubiquitin specific peptidase 6 (Tre-2 oncogene)*	NM_004505	CCCAGGATCGTGATAACTGTAT	CAGCCTCTGCAAAATCTATACTG
VASH2	*vasohibin 2*	NM_024749	ACGTGGCCAAGGTGCACCCTA	CTGGGGTATTGAAGGCTTTGCCAA
VEGFC	*vascular endothelial growth factor C*	NM_005429	ACTACCACAGTGTCAGGCAGC	GGAATCCATCTGTTGAGTCATCTC
WEE1	*WEE1 homolog (S. pombe)*	NM_003390	GTAATGGTGGAAGTTTAGCTGAT	CCAAAGACATTGAATGAATATACC
ZNF554	*zinc finger protein 554*	NM_001102651	CTCAAGCCTCCTGTTCAGATTGGATGACT	TGTCCTCTAACTGCTTCCATCCCCCAT
CAV1	*caveolin 1, caveolae protein*	NM_001753	GCAAATACGTAGACTCGGAGGGA	TCAATCTTGACCACGTCATCGTT
HCLS1	*hematopoietic cell-specific Lyn substrate 1*	NM_005335	AAACACGAGTCCCAGAGAGATTATGCCAA	TCATTGAAGCCGACAGCGCTCTTATCCA
JAG2	*jagged 2*	NM_002226	CTCTCTGTGAGGTGGATGTCGACC	ACGGAGCAGTTCTTGCCACCAA
NOTCH1	*notch 1*	NM_017617	CTGCCTGCCAGGCTTCACCGGCCAGAA	TCGGTACAGTACTGACCTGTCCACTCTGG
OCLN	*occludin*	NM_002538	CTGGATCAGGGAATATCCACCTATCA	CAATTCTTTATCCAAACGGGAGAGTTC
PPIA	*peptidylprolyl isomerase A (cyclophilin A)*	NM_021130	CTTTGAGCTGTTTGCAGACAAGGTC	GGAATAATTCTGTGAAAGCAGGAACCC
PLAC8	*placenta-specific 8*	NM_001130716	CGATATGGCATCCCTGGATCTATTTGTGATGA	GCATGGCTCTCCTTCTGTTGATATTCTCTTG
ROR2	*receptor tyrosine kinase-like orphan receptor 2*	NM_004560	GGAATTGCCTGTGCACGCTTCA	GATCATGGTGAAGGCCGCTGTG
WISP2	*WNT1 inducible signaling pathway protein 2*	NM_003881	CCTCTCTCAAAGGTGCGTACCCAGCT	GTGCACATACCCGGCAGCAGCCA
RNU6	*U6 small nuclear RNA*	NR_004394	GTGCTCGCTTCGGCAGCACATAT	AAAAATATGGAACGCTTCACGAA

Appendix 5: A diagram of pcDNA3.1 TOPO vector.



Primers were designed to amplify the target ORF and adding extra 5'-CACC-3' sequence (marked with red arrow) needed for introducing the ORF into a mammalian expressing vector, pcDNA3.1D/V5-His-TOPO. The vector's TOPO enzyme ligates the CACC base pairs at the ORF 5'-end with its complementary sequence on the vector, in addition to ligating the ORF 3' blunt-end with the vector's other side to form a circular sequence. *The sketch was taken from the manufacturer's manual.*

Appendix 6: A diagram of pGL3-basic vector.



Cloning primers were designed to specifically amplify a promoter DNA sequence but with the addition of upstream and downstream restriction site sequences to allow for cloning into the vector's multiple cloning site (marked with red arrow). *The sketch was taken from the manufacturer's manual.*

Appendix 7: Bacterial culture plate.

In a 1 L conical flask, 20 g of (Luria-Bertani) LB agar granulated medium (Fisher, Loughborough, UK) were dissolved in 500 mL dH_2O. Then, the flask was covered with aluminum foil and autoclaved at 121ºC for 15 minutes under 15 pounds per square inch (psi) air pressure (121ºC/15 min/15 psi). Directly after the autoclaving, the medium was allowed to cool down in a 50 °C water bath. This temperature was unharmful to the ampicillin, which was added in an amount sufficient to give a final concentration of 100 µg/mL. The agar was then poured into clean petri dish plates in a quantity of 20 mL/plate. Finally, the agar plates were allowed to completely solidify before using directly or storing at 4 °C for future use.

Appendix 8: LB broth medium .

12 g LB broth granulated medium (Fisher, Loughborough, UK) were added to 500 mL dH_2O, then autoclaved as before. The broth was stored at room temperature. The aluminium foil was not removed unless next to a flame to prevent airborne contamination. A miniprep culture was prepared by adding 4 mL of the LB broth medium to a 20 mL sterile universal tube and a sufficient amount of ampicillin to give a final concentration of 100 µg/mL.

Appendix 9: The 414 putative targets of HOXD10 identified by ANOVA test.

Probe ID	Accession number	**Gene symbol**	*p* - value (ANOVA)	**Effect of HOXD10 over-expression**	**Effect of HOXD10 silencing**	Fold change
A_23_P23074	NM_006417	**IFI44**	4.38E-06	-	+	4.6
A_23_P65518	NM_016651	**DACT1**	5.92E-06	-	+	2.7
A_23_P133408	NM_000758	**CSF2**	6.36E-06	-	+	6
A_23_P114185	NM_004615	**TSPAN7**	1.20E-05	+	-	3.1
A_23_P10232	NM_017935	**BANK1**	1.29E-05	+	-	3.3
A_23_P14083	NM_181847	**AMIGO2**	1.41E-05	-	+	2.8
A_33_P3249534	NM_005382	**NEFM**	1.44E-05	+	-	3.2
A_23_P125815	NM_005676	**RBM10**	2.51E-05	-	+	2.7
A_23_P36611	NM_181861	**APAF1**	2.76E-05	-	+	2.4
A_33_P3221748	NM_001031680	**RUNX3**	2.89E-05	-	+	3
A_32_P95914	NM_198468	**MMS22L**	2.95E-05	+	-	2.6
A_23_P23279	NM_052862	**RCSD1**	3.01E-05	+	-	5.3
A_23_P19020	NM_005460	**SNCAIP**	3.11E-05	+	-	3.1
A_24_P344961	NM_133265	**AMOT**	4.03E-05	+	-	3.9
A_33_P3243887	NM_000641	**IL11**	4.37E-05	-	+	6.3
A_33_P3224331	NM_001122665	**DDX3Y**	4.56E-05	+	-	4.2
A_23_P215037	NM_152858	**WTAP**	4.65E-05	-	+	2.6
A_33_P3745146	NM_001098517	**CADM1**	4.85E-05	-	+	2.5
A_33_P3308744	NM_001105206	**LAMA4**	4.97E-05	+	-	3.1
A_23_P103201	NM_017761	**PNRC2**	5.63E-05	-	+	2.6
A_33_P3336686	NM_004669	**CLIC3**	5.67E-05	+	-	3
A_24_P45379	NM_014412	**CACYBP**	6.35E-05	+	-	2.8
A_33_P3251369	NM_001081573	**GAB3**	6.72E-05	+	-	6.7
A_33_P3336282	NM_005640	**TAF4B**	6.93E-05	+	-	3.2
A_23_P337270	ENST00000383770	**NEK10**	7.41E-05	-	+	2.9
A_23_P362148	NM_139072	**DNER**	8.56E-05	-	+	3.7
A_23_P399078	NM_000362	**TIMP3**	8.61E-05	-	+	2.8
A_33_P3263890	NM_006902	**PRRX1**	9.02E-05	-	+	2.7
A_33_P3243702	NM_198582	**KLHL30**	9.86E-05	+	-	3
A_24_P320665	NM_032581	**FAM126A**	9.93E-05	-	+	2.8
A_23_P120170	NM_145702	**TIGD1**	1.00E-04	+	-	2.6
A_33_P3327479	NM_016598	**ZDHHC3**	1.09E-04	+	-	2.7
A_23_P205007	NM_002271	**IPO5**	1.18E-04	+	-	2.6
A_23_P134176	NM_001024465	**SOD2**	1.29E-04	-	+	3.4
A_24_P377144	NM_058172	**ANTXR2**	1.33E-04	-	+	2.8
A_23_P121806	NM_021204	**ENOPH1**	1.51E-04	-	+	2.7
A_33_P3281695	NM_001243133	**NLRP3**	1.62E-04	-	+	3
A_23_P125423	NM_001733	**C1R**	1.75E-04	-	+	3.4
A_23_P155939	NM_182524	**ZNF595**	1.81E-04	+	-	3.7
A_33_P3236813	NM_006143	**GPR19**	1.85E-04	+	-	2.7
A_23_P106389	NM_003612	**SEMA7A**	1.86E-04	-	+	4
A_23_P24784	NM_003282	**TNNI2**	2.07E-04	+	-	4
A_33_P3290567	NM_003390	**WEE1**	2.17E-04	-	+	2.6
A_23_P136116	NM_001004320	**AGMO**	2.18E-04	-	+	3.1
A_23_P300056	NM_044472	**CDC42**	2.18E-04	-	+	2.5
A_32_P21579	NM_004412	**TRDMT1**	2.22E-04	+	-	2.5
A_23_P85693	NM_004120	**GBP2**	2.44E-04	+	-	2.9
A_24_P310256	NM_139284	**LGI4**	2.62E-04	+	-	4.6
A_23_P36187	NM_138567	**SYT8**	2.68E-04	+	-	4.2
A_33_P3237110	NM_001099652	**GPR137C**	2.96E-04	-	+	2.8
A_32_P164246	NM_033260	**FOXQ1**	3.02E-04	+	-	3.2
A_33_P3376636	NM_001017989	**OPA3**	3.04E-04	-	+	2.7
A_23_P162300	NM_007199	**IRAK3**	3.41E-04	-	+	3.4
A_23_P94533	NM_001912	**CTSL1**	3.50E-04	-	+	2.5
A_24_P406986	NM_199329	**SLC43A3**	3.85E-04	-	+	2.6
A_23_P381368	NM_002148	**HOXD10**	3.86E-04	+	-	2.9
A_33_P3401902	NM_001012421	**ANKRD20A2**	3.94E-04	+	-	3.5
A_33_P3718269	NR_029701	**MIR146A**	4.11E-04	-	+	2.8
A_32_P34920	NM_004472	**FOXD1**	4.29E-04	-	+	2.6
A_23_P84448	NM_006000	**TUBA4A**	4.29E-04	-	+	2.6
A_23_P55256	NM_014897	**ZNF652**	4.54E-04	+	-	2.6
A_23_P8913	NM_000067	**CA2**	4.67E-04	+	-	2.7
A_32_P193080	NM_052905	**FMNL2**	4.74E-04	-	+	2.8
A_24_P916141	NM_005828	**DCAF7**	5.92E-04	+	-	2.5
A_33_P3420224	NM_001033113	**ENTPD8**	5.96E-04	+	-	3.8
A_23_P121956	NM_017872	**THG1L**	6.18E-04	-	+	2.8
A_33_P3357658	NM_003484	**HMGA2**	6.21E-04	-	+	2.8
A_23_P212655	NM_130446	**KLHL6**	6.24E-04	+	-	4.7
A_23_P42257	NM_003897	**IER3**	6.43E-04	-	+	2.7
A_23_P70794	NM_016277	**RAB23**	6.44E-04	-	+	2.6

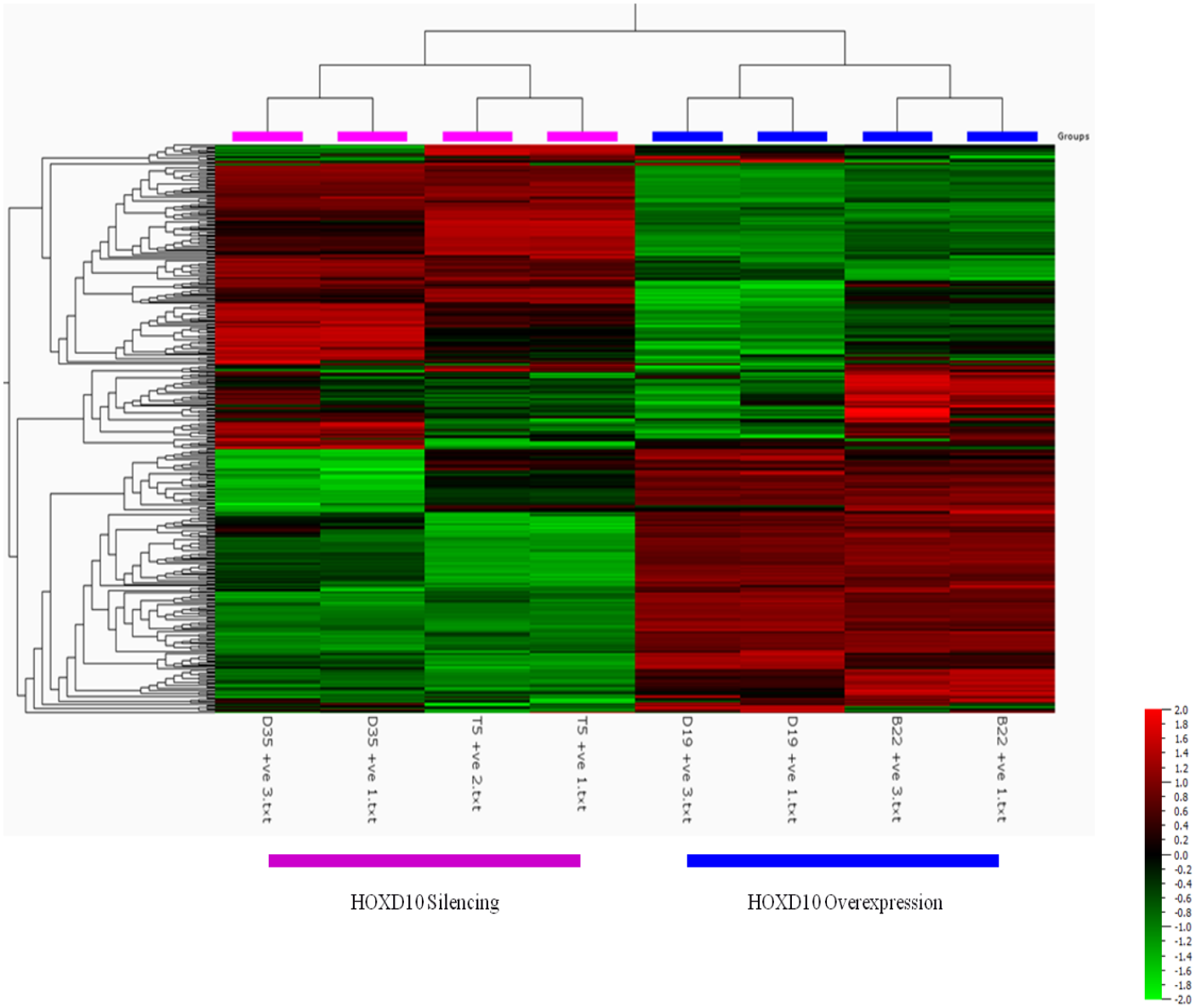
A_33_P3278362	NM_020349	**ANKRD2**	7.17E-04	+	-	2.7
A_33_P3320159	NM_031415	**GSDMC**	7.23E-04	-	+	3.5
A_32_P32739	NM_153006	**NAGS**	7.26E-04	-	+	2.6
A_23_P212508	NM_001063	**TF**	7.46E-04	+	-	3.7
A_24_P45367	NM_020448	**NIPAL3**	7.51E-04	+	-	2.7
A_23_P57474	NM_030758	**OSBP2**	8.01E-04	-	+	3.4
A_23_P2492	NM_001734	**C1S**	8.39E-04	-	+	3.8
A_23_P52676	NM_053054	**CATSPER1**	8.73E-04	-	+	2.9
A_23_P16469	NM_001005377	**PLAUR**	8.84E-04	-	+	2.8
A_33_P3379644	NM_000689	**ALDH1A1**	8.98E-04	+	-	2.6
A_24_P56281	NM_001002259	**CAPRIN2**	9.83E-04	-	+	2.6
A_23_P127565	NM_178834	**LAYN**	1.01E-03	-	+	2.6
A_24_P82106	NM_004995	**MMP14**	1.08E-03	-	+	2.7
A_32_P108156	NR_001458	**MIR155HG**	1.08E-03	-	+	4.3
A_23_P155057	NM_013385	**CYTH4**	1.11E-03	-	+	3.1
A_32_P114284	NM_001079526	**IKZF2**	1.16E-03	+	-	2.6
A_23_P143512	NM_007031	**HSF2BP**	1.16E-03	+	-	2.8
A_32_P59302	NM_024503	**HIVEP3**	1.16E-03	-	+	2.8
A_33_P3393927	NM_181873	**MTMR11**	1.16E-03	-	+	2.5
A_33_P3844650	NM_012098	**ANGPTL2**	1.20E-03	-	+	3
A_23_P11697	NM_145205	**HMGB4**	1.21E-03	+	-	3.3
A_32_P358887	NM_003759	**SLC4A4**	1.22E-03	-	+	3.3
A_23_P211428	NM_134269	**SMTN**	1.22E-03	-	+	2.9
A_33_P3299000	NM_031952	**SPATA9**	1.29E-03	+	-	4
A_23_P203475	NM_145040	**PRKCDBP**	1.35E-03	-	+	2.8
A_23_P1029	NM_017459	**MFAP2**	1.43E-03	-	+	3.2
A_23_P218111	NM_001002236	**SERPINA1**	1.48E-03	-	+	3.1
A_33_P3295358	NM_139314	**ANGPTL4**	1.50E-03	-	+	5.1
A_32_P153388	NM_016315	**GULP1**	1.53E-03	-	+	2.9
A_24_P406006	NM_024830	**LPCAT1**	1.56E-03	-	+	2.6
A_23_P166823	NM_003280	**TNNC1**	1.62E-03	+	-	3.3
A_23_P92517	NM_031956	**TTC29**	1.62E-03	+	-	3.2
A_33_P3404126	ENST00000430644	**PCGF3**	1.62E-03	-	+	2.8
A_32_P96719	NM_024745	**SHCBP1**	1.63E-03	-	+	2.7
A_24_P375205	ENST00000389126	**MKL2**	1.65E-03	+	-	3
A_33_P3417695	NM_001014440	**ODF3B**	1.69E-03	-	+	2.7
A_32_P45738	NM_002629	**PGAM1**	1.76E-03	-	+	3.1
A_23_P143029	NM_021192	**HOXD11**	1.81E-03	+	-	3.4
A_23_P114349	NM_130776	**XAGE3**	1.84E-03	-	+	3
A_24_P367645	NM_152780	**MAP7D2**	1.97E-03	+	-	3.4
A_32_P83049	NM_014971	**EFR3B**	2.13E-03	-	+	2.8
A_33_P3277198	NM_004854	**CHST10**	2.15E-03	+	-	2.5
A_33_P3288189	NM_014899	**RHOBTB3**	2.18E-03	+	-	2.8
A_33_P3627001	NM_144962	**PEBP4**	2.29E-03	+	-	2.3
A_32_P140898	NM_002158	**FOXN2**	2.29E-03	-	+	2.5
A_23_P73429	NM_005335	**HCLS1**	2.30E-03	-	+	2.3
A_33_P3332724	NM_015120	**ALMS1**	2.43E-03	+	-	2.8
A_23_P3956	NM_198594	**C1QTNF1**	2.49E-03	-	+	3.5
A_23_P30020	NM_030821	**PLA2G12A**	2.55E-03	+	-	2.8
A_23_P218770	NM_002872	**RAC2**	2.55E-03	-	+	2.2
A_33_P3310274	ENST00000420582	**LGR6**	2.64E-03	+	-	3
A_24_P79403	NM_002619	**PF4**	2.75E-03	-	+	2.7
A_23_P167040	NM_006810	**PDIA5**	2.78E-03	-	+	2.6
A_32_P42197	NM_031157	**HNRNPA1**	2.87E-03	+	-	2.6
A_23_P161624	NM_005438	**FOSL1**	2.97E-03	-	+	3.3
A_33_P3390539	NM_152312	**GYLTL1B**	3.00E-03	-	+	3.6
A_23_P18739	NM_018241	**TMEM184C**	3.05E-03	+	-	2.7
A_33_P3244808	NM_153274	**BEST4**	3.11E-03	-	+	2.6
A_24_P217572	NM_001957	**EDNRA**	3.17E-03	-	+	3.9
A_23_P120125	NM_199235	**COLEC11**	3.21E-03	-	+	3.3
A_23_P39251	NM_001013706	**PLIN5**	3.29E-03	-	+	2.5
A_23_P15174	NM_005949	**MT1F**	3.35E-03	+	-	2.6
A_24_P388786	NM_001369	**DNAH5**	3.43E-03	+	-	3.3
A_23_P392470	NM_000901	**NR3C2**	3.46E-03	+	-	3.1
A_24_P85258	NM_001080484	**KIAA1751**	3.75E-03	+	-	3.2
A_33_P3356462	NM_207322	**C2CD4A**	3.83E-03	-	+	4.9
A_33_P3216890	NM_018440	**PAG1**	3.84E-03	-	+	2.8
A_23_P208334	NM_006202	**PDE4A**	3.87E-03	+	-	2.7
A_24_P257579	NM_022140	**EPB41L4A**	3.87E-03	+	-	4.5
A_24_P330518	NM_001218	**CA12**	3.96E-03	-	+	6.1
A_23_P401904	NM_001009936	**PHF19**	3.96E-03	-	+	2.5
A_23_P114839	NM_004468	**FHL3**	4.05E-03	-	+	2.6
A_33_P3318357	NM_138775	**ALKBH8**	4.11E-03	+	-	2.7
A_24_P389916	NM_005512	**LRRC32**	4.12E-03	-	+	2.8
A_23_P155979	NM_001963	**EGF**	4.19E-03	-	+	3.2
A_33_P3420900	NM_212555	**PATE2**	4.27E-03	-	+	3
A_24_P85557	NM_005302	**GPR37**	4.40E-03	-	+	3.5

A_33_P3278941	NM_001048205	**REC8**	4.41E-03	-	+	2.7
A_24_P416346	NM_001079675	**ETV4**	4.57E-03	-	+	2.8
A_23_P62953	NM_002585	**PBX1**	4.59E-03	+	-	3.4
A_23_P25503	NM_001079673	**FNDC3A**	4.78E-03	-	+	2.5
A_33_P3386429	NM_015571	**SENP6**	4.79E-03	+	-	3.6
A_24_P16730	NM_030650	**KIAA1715**	4.80E-03	+	-	2.9
A_23_P134953	NM_001122	**PLIN2**	4.96E-03	-	+	3
A_23_P114983	NM_032588	**TRIM63**	5.20E-03	-	+	3.3
A_24_P146892	NM_032790	**ORAI1**	5.30E-03	-	+	2.7
A_33_P3320994	ENST00000234668	**SYDE2**	5.32E-03	+	-	2.8
A_33_P3257714	NM_001025	**RPS23**	5.43E-03	+	-	2.9
A_23_P65918	NM_002220	**ITPKA**	5.53E-03	-	+	2.8
A_23_P42718	NM_004289	**NFE2L3**	5.55E-03	-	+	3
A_23_P253958	NM_005824	**LRRC17**	5.55E-03	-	+	3
A_33_P3230259	NM_015341	**NCAPH**	5.82E-03	+	-	2.6
A_23_P131990	NM_014588	**VSX1**	5.84E-03	+	-	5.1
A_33_P3322085	NM_001102651	**ZNF554**	5.86E-03	+	-	5.8
A_33_P3387616	NM_052924	**RHPN1**	5.88E-03	+	-	3.4
A_23_P107116	NM_007148	**RNF112**	5.89E-03	+	-	2.6
A_33_P3416331	NM_001010893	**SLC10A5**	5.93E-03	+	-	3
A_23_P148473	NM_000206	**IL2RG**	5.97E-03	-	+	3.2
A_23_P161507	NM_004923	**MTL5**	5.98E-03	+	-	2.7
A_24_P157926	NM_006290	**TNFAIP3**	5.98E-03	-	+	2.8
A_23_P408830	NM_172215	**CAMKK2**	6.09E-03	-	+	2.7
A_23_P118150	NM_015161	**ARL6IP1**	6.35E-03	-	+	2.9
A_24_P191067	NM_001009566	**CLSTN1**	6.67E-03	-	+	2.7
A_33_P3285299	NM_014696	**GPRIN2**	6.70E-03	+	-	3.9
A_23_P75220	NM_031212	**SLC25A28**	7.00E-03	-	+	2.6
A_33_P3257993	NM_017831	**RNF125**	7.15E-03	+	-	2.7
A_24_P270728	NM_001042483	**N+R1**	7.35E-03	+	-	2.8
A_23_P216679	NM_033331	**CDC14B**	7.37E-03	+	-	2.7
A_33_P3358712	NR_027390	**POLR2M**	7.50E-03	+	-	4.2
A_33_P3281850	ENST00000440612	**CGREF1**	7.62E-03	-	+	2.8
A_23_P203558	NM_000518	**HBB**	7.72E-03	+	-	3.4
A_23_P45871	NM_006820	**IFI44L**	7.78E-03	-	+	4.5
A_33_P3252809	ENST00000405673	**FAM118A**	7.82E-03	+	-	2.8
A_33_P3365193	NM_001008219	**AMY1C**	7.94E-03	+	-	2.8
A_23_P374689	NM_000817	**GAD1**	7.99E-03	+	-	3.8
A_23_P161698	NM_002422	**MMP3**	8.09E-03	-	+	3.3
A_23_P65532	NM_021255	**PELI2**	8.22E-03	+	-	3.3
A_23_P43810	NM_206943	**LTBP1**	8.26E-03	-	+	2.8
A_33_P3299220	NM_025008	**ADAMTSL4**	8.37E-03	-	+	2.7
A_23_P118158	NM_006043	**HS3ST2**	8.43E-03	+	-	3.7
A_23_P7282	NM_153702	**ELMOD2**	8.50E-03	+	-	2.6
A_33_P3223678	NM_014564	**LHX3**	8.53E-03	+	-	2.9
A_23_P319859	NM_005244	**EYA2**	8.87E-03	+	-	2.7
A_33_P3355014	NM_182526	**TMEM229B**	8.90E-03	-	+	3.3
A_23_P2322	NM_001100917	**TSPAN19**	9.10E-03	-	+	4
A_23_P428298	NM_173561	**UNC5CL**	9.12E-03	+	-	5.3
A_24_P383609	NM_199461	**NANOS1**	9.13E-03	+	-	3
A_33_P3344956	ENST00000540155	**TXNL1**	9.22E-03	+	-	3.7
A_33_P3266550	NM_020469	**ABO**	9.39E-03	+	-	3
A_24_P941643	NM_182734	**PLCB1**	9.49E-03	+	-	4.6
A_23_P165239	NM_007153	**ZNF208**	9.51E-03	-	+	3.5
A_33_P3209229	NM_014353	**RAB26**	9.91E-03	+	-	3.2
A_24_P149704	NM_138709	**DAB2IP**	1.00E-02	+	-	4
A_33_P3223592	NM_000041	**APOE**	1.01E-02	-	+	2.8
A_24_P605190	NM_005023	**PGGT1B**	1.02E-02	+	-	2.6
A_23_P47806	NM_012064	**MIP**	1.02E-02	+	-	3.9
A_33_P3292769	NM_145912	**NFAM1**	1.02E-02	+	-	3.3
A_23_P82065	NM_022726	**ELOVL4**	1.04E-02	-	+	2.8
A_33_P3227857	NM_173827	**COX18**	1.11E-02	+	-	2.8
A_32_P228501	NM_139286	**CDC26**	1.12E-02	+	-	2.6
A_23_P360754	NM_005099	**ADAMTS4**	1.13E-02	-	+	5
A_33_P3407384	ENST00000373413	**ZSCAN20**	1.14E-02	+	-	2.8
A_23_P364837	NM_015378	**VPS13D**	1.14E-02	+	-	2.8
A_33_P3386671	NM_005060	**RORC**	1.16E-02	-	+	2.8
A_23_P78867	NM_000540	**RYR1**	1.18E-02	+	-	3.1
A_23_P69020	NM_001105580	**GABRR3**	1.18E-02	+	-	3.1
A_23_P106024	NM_002226	**JAG2**	1.19E-02	+	-	2.8
A_23_P60146	NM_006207	**PDGFRL**	1.19E-02	-	+	2.2
A_23_P206626	NM_001039	**SCNN1G**	1.20E-02	+	-	2.5
A_24_P14974	NM_182537	**HTR3D**	1.20E-02	+	-	5.5
A_23_P134454	NM_001753	**CAV1**	1.20E-02	+	-	2.3
A_23_P83028	NM_021111	**RECK**	1.20E-02	-	+	4.2
A_24_P230948	NM_016831	**PER3**	1.21E-02	+	-	2.8
A_24_P316430	NM_002526	**NT5E**	1.21E-02	-	+	3.5

A_23_P63896	NM_000043	**FAS**	1.21E-02	-	+	2.7
A_23_P38346	NM_024119	**DHX58**	1.22E-02	-	+	3
A_23_P71433	NM_001001481	**UBE2W**	1.23E-02	-	+	2.6
A_24_P284805	NM_001002837	**INPP5J**	1.24E-02	-	+	3.1
A_24_P175176	NM_020432	**PHTF2**	1.24E-02	-	+	2.6
A_23_P70818	NM_005631	**SMO**	1.25E-02	+	-	3.1
A_24_P362904	NM_004567	**PFKFB4**	1.25E-02	-	+	2.8
A_23_P414913	NM_022343	**GLIPR2**	1.25E-02	-	+	2.8
A_33_P3415623	ENST00000525971	**LRRIQ1**	1.26E-02	-	+	2.8
A_23_P207213	NM_000691	**ALDH3A1**	1.27E-02	+	-	4.1
A_23_P11081	NM_003886	**AKAP4**	1.27E-02	+	-	2.8
A_23_P39955	NM_001615	**ACTG2**	1.30E-02	+	-	4.3
A_23_P334798	NM_024512	**LRRC2**	1.33E-02	+	-	3.4
A_23_P6771	NM_014583	**LMCD1**	1.33E-02	-	+	3
A_23_P51646	NM_004073	**PLK3**	1.34E-02	-	+	2.7
A_23_P500542	NM_147192	**DMBX1**	1.36E-02	+	-	3.6
A_23_P322756	NM_030625	**TET1**	1.36E-02	-	+	3.8
A_33_P3311145	NM_014956	**CEP164**	1.38E-02	+	-	2.8
A_23_P161659	NM_020826	**SYT13**	1.38E-02	+	-	6.1
A_23_P165707	NM_174898	**LYG1**	1.39E-02	-	+	2.7
A_23_P419107	NM_152772	**TCP11L2**	1.42E-02	+	-	2.6
A_33_P3393684	NM_001013661	**VSIG8**	1.47E-02	-	+	2.8
A_33_P3350553	ENST00000376964	**UXT**	1.48E-02	+	-	3.7
A_33_P3422968	NM_001370	**DNAH6**	1.48E-02	+	-	5
A_33_P3271350	NM_003640	**IKBKAP**	1.48E-02	+	-	3
A_33_P3366431	NM_002297	**LCN1**	1.49E-02	+	-	3.3
A_23_P171107	NM_033031	**CCNB3**	1.49E-02	-	+	2.6
A_23_P121545	NM_201591	**GPM6A**	1.53E-02	+	-	2.7
A_23_P58082	NM_199511	**CCDC80**	1.58E-02	+	-	2.8
A_33_P3266964	NM_000257	**MYH7**	1.59E-02	+	-	2.8
A_33_P3401647	NM_033256	**PPP1R14A**	1.60E-02	+	-	3.2
A_33_P3318796	NM_005860	**FSTL3**	1.61E-02	-	+	2.9
A_24_P49106	NM_152278	**TCEAL7**	1.62E-02	+	-	8.1
A_33_P3228510	NM_177996	**EPB41L1**	1.62E-02	+	-	3.8
A_33_P3303212	NM_207310	**CCDC74B**	1.66E-02	+	-	3
A_23_P58228	NM_017855	**ODAM**	1.67E-02	+	-	4.3
A_32_P14762	NM_001080507	**OOEP**	1.67E-02	-	+	3.8
A_23_P46351	NM_006862	**TDRKH**	1.68E-02	-	+	2.8
A_24_P135753	ENST00000383657	**PTPLB**	1.72E-02	+	-	2.7
A_33_P3264926	NM_015589	**SAMD4A**	1.74E-02	-	+	2.6
A_33_P3366484	NM_198047	**HIBCH**	1.76E-02	+	-	2.9
A_33_P3370424	NM_017617	**NOTCH1**	1.78E-02	+	-	2.1
A_23_P379789	NM_013305	**ST8SIA5**	1.81E-02	+	-	2.6
A_23_P62115	NM_003254	**TIMP1**	1.81E-02	-	+	3
A_33_P3235009	NM_001242415	**WDR20**	1.82E-02	+	-	4.7
A_33_P3209772	NM_006237	**POU4F1**	1.83E-02	+	-	3
A_23_P169092	NM_004820	**CYP7B1**	1.84E-02	+	-	3.1
A_23_P356070	NM_002020	**FLT4**	1.85E-02	+	-	2.8
A_23_P381478	NM_152492	**CCDC27**	1.86E-02	+	-	4.7
A_23_P155786	NM_005420	**SULT1E1**	1.88E-02	+	-	3.7
A_23_P102950	NM_080860	**RSPH1**	1.90E-02	+	-	3.4
A_23_P348264	NM_144652	**LETM2**	1.91E-02	-	+	2.8
A_23_P86599	NM_007329	**DMBT1**	1.97E-02	-	+	2.9
A_23_P10062	NM_013244	**MGAT4C**	1.98E-02	+	-	2.7
A_23_P160668	NM_005376	**MYCL1**	2.01E-02	+	-	3.6
A_24_P277673	NM_003547	**HIST1H4G**	2.05E-02	+	-	2.8
A_23_P32175	NM_014368	**LHX6**	2.06E-02	+	-	3.3
A_33_P3369844	NM_013230	**CD24**	2.15E-02	+	-	3
A_23_P374695	NM_000459	**TEK**	2.17E-02	-	+	6
A_23_P137035	NM_003662	**PIR**	2.18E-02	+	-	2.4
A_24_P103264	NM_003360	**UGT8**	2.21E-02	-	+	2.9
A_33_P3423941	NM_003641	**IFITM1**	2.25E-02	-	+	4.6
A_33_P3216008	NM_145061	**SKA3**	2.28E-02	+	-	2.6
A_23_P104471	NM_001007271	**DUSP13**	2.29E-02	+	-	2.7
A_33_P3294446	NM_024567	**HMBOX1**	2.31E-02	+	-	2.8
A_23_P120594	NM_032501	**ACSS1**	2.32E-02	+	-	2.6
A_23_P140256	NM_000270	**PNP**	2.33E-02	-	+	2.7
A_24_P370670	NM_001195156	**ZMYM6NB**	2.33E-02	-	+	2.7
A_33_P3292719	NM_001013658	**PTX4**	2.34E-02	+	-	3.3
A_33_P3263533	NM_001202435	**SCN1A**	2.40E-02	-	+	4
A_33_P3599591	NM_002581	**PAPPA**	2.40E-02	-	+	2.9
A_23_P204269	NM_006313	**USP15**	2.40E-02	-	+	2.7
A_23_P433369	NM_015473	**HEATR5A**	2.41E-02	+	-	2.6
A_33_P3395947	ENST00000380922	**IL4R**	2.46E-02	+	-	3.2
A_23_P254212	NM_013347	**RPA4**	2.46E-02	-	+	2.8
A_23_P406424	NM_175744	**RHOC**	2.48E-02	-	+	2.1
A_33_P3222689	NM_001164416	**H2BFM**	2.58E-02	+	-	2.7

A_23_P420209	NM_004751	**GCNT3**	2.58E-02	+	-	3.2
A_33_P3327462	NM_023943	**TMEM108**	2.60E-02	+	-	2.8
A_33_P3265394	ENST00000538098	**WDR74**	2.61E-02	+	-	2.6
A_23_P389102	NM_015194	**MYO1D**	2.62E-02	+	-	2.5
A_33_P3258091	NM_020216	**RNPEP**	2.63E-02	+	-	2.5
A_32_P193322	NM_152756	**RICTOR**	2.66E-02	-	+	2.5
A_23_P106806	NM_031948	**PRSS27**	2.67E-02	+	-	3.5
A_23_P351148	NM_053282	**SH2D1B**	2.67E-02	+	-	2.8
A_23_P58407	NM_001076	**UGT2B15**	2.70E-02	+	-	3.1
A_23_P112482	NM_004925	**AQP3**	2.70E-02	+	-	3
A_23_P354151	NM_005546	**ITK**	2.70E-02	+	-	3.3
A_33_P3232692	NM_001185156	**IL24**	2.71E-02	-	+	4.1
A_33_P3214720	NM_025079	**ZC3H12A**	2.75E-02	-	+	2.7
A_24_P121406	NM_015070	**ZC3H13**	2.77E-02	+	-	2.6
A_24_P14634	NM_133455	**EMID1**	2.78E-02	+	-	4.9
A_23_P215227	NM_005494	**DNAJB6**	2.79E-02	-	+	3.7
A_23_P167096	NM_005429	**VEGFC**	2.81E-02	-	+	3.6
A_23_P101131	NM_002091	**GRP**	2.82E-02	+	-	2.8
A_24_P361457	NM_173627	**ENDOV**	2.82E-02	+	-	3.1
A_33_P3412678	NM_001037498	**DEFB112**	2.90E-02	+	-	2.9
A_23_P396858	NM_031866	**FZD8**	2.90E-02	-	+	3.2
A_33_P3323902	NM_001037804	**DEFB130**	2.91E-02	+	-	3.4
A_23_P36397	NM_000785	**CYP27B1**	2.96E-02	-	+	4.4
A_32_P83811	NM_001242936	**FAM47E**	2.97E-02	+	-	3.2
A_24_P260699	NM_003958	**RNF8**	2.97E-02	+	-	3
A_23_P398854	NM_173660	**DOK7**	2.98E-02	-	+	3.2
A_24_P160969	NM_001076787	**TP53I11**	3.00E-02	+	-	2.9
A_24_P380061	NM_001025616	**ARHGAP24**	3.01E-02	+	-	2.8
A_32_P300427	NM_153360	**APCDD1L**	3.03E-02	-	+	3.6
A_23_P92536	NM_000087	**CNGA1**	3.04E-02	+	-	3.9
A_33_P3314659	NM_144722	**SPEF2**	3.07E-02	+	-	2.9
A_33_P3315679	NM_005511	**MLANA**	3.11E-02	+	-	2.6
A_32_P136351	NM_052910	**SLITRK1**	3.11E-02	+	-	3.1
A_23_P308058	NM_172367	**TUSC5**	3.12E-02	+	-	4
A_23_P122924	NM_002192	**INHBA**	3.13E-02	-	+	3.5
A_23_P123488	NM_024504	**PRDM14**	3.14E-02	+	-	3.2
A_23_P397391	NM_005306	**FFAR2**	3.14E-02	-	+	3.1
A_33_P3345831	NM_001010874	**TECRL**	3.17E-02	+	-	4.9
A_33_P3407013	ENST00000420258	**IL6**	3.17E-02	-	+	3.4
A_33_P3251518	ENST00000366823	**PYCR2**	3.32E-02	+	-	3
A_23_P94840	NM_130897	**DYNLRB2**	3.33E-02	+	-	3.6
A_23_P38830	NM_024762	**ZNF552**	3.40E-02	+	-	3.1
A_33_P3245238	NM_176825	**SULT1C2**	3.43E-02	+	-	3.1
A_23_P395555	NM_001032372	**ZNF226**	3.49E-02	+	-	2.7
A_23_P416011	NM_058203	**SPAG11B**	3.51E-02	+	-	3.1
A_33_P3286036	ENST00000395264	**TMPRSS9**	3.55E-02	+	-	4.4
A_33_P3290403	NM_014214	**IMPA2**	3.56E-02	-	+	2.5
A_33_P3351346	ENST00000523711	**ARHGEF10**	3.58E-02	+	-	3.3
A_33_P3235078	NM_001134855	**TRIM17**	3.59E-02	+	-	5.4
A_23_P76749	NM_020692	**GALNTL1**	3.61E-02	-	+	2.9
A_33_P3404779	NM_001013632	**TCTEX1D4**	3.62E-02	-	+	2.6
A_23_P154488	NM_033109	**PNPT1**	3.66E-02	-	+	2.8
A_23_P129332	NM_001076780	**PKD1L2**	3.67E-02	+	-	2.7
A_24_P257478	NM_198721	**COL25A1**	3.67E-02	+	-	4.5
A_33_P3257891	NM_000847	**GSTA3**	3.73E-02	+	-	2.7
A_23_P18447	NM_013261	**PPARGC1A**	3.79E-02	+	-	3.5
A_23_P9075	NM_152413	**GOT1L1**	3.80E-02	+	-	2.7
A_32_P181077	NM_203447	**DOCK8**	3.83E-02	+	-	2.8
A_23_P30200	NM_020957	**PCDHB16**	3.86E-02	-	+	3.4
A_23_P204640	NM_024865	**NANOG**	3.87E-02	+	-	2.9
A_33_P3424062	NM_002236	**KCNF1**	3.89E-02	+	-	4.3
A_23_P361381	NM_007068	**DMC1**	3.89E-02	+	-	3
A_33_P3309130	NM_182833	**GDPD4**	3.92E-02	+	-	2.8
A_33_P3332492	NM_145235	**FANK1**	3.94E-02	+	-	2.8
A_33_P3371175	NM_177538	**CYP20A1**	3.94E-02	+	-	3.1
A_23_P151710	NM_000956	**PTGER2**	3.94E-02	-	+	3.4
A_23_P401774	NM_018712	**ELMOD1**	3.95E-02	+	-	2.7
A_23_P160214	NM_001080494	**TTC39A**	3.96E-02	+	-	2.9
A_24_P80135	NM_014369	**PTPN18**	3.97E-02	+	-	3
A_33_P3265374	NM_001145722	**HOMER3**	4.03E-02	-	+	2.7
A_33_P3211666	NM_003855	**IL18R1**	4.03E-02	-	+	2.8
A_33_P3249072	ENST00000370642	**HOGA1**	4.04E-02	+	-	2.7
A_24_P65722	ENST00000391736	**LILRB4**	4.08E-02	+	-	3.1
A_23_P105705	NM_020996	**FGF6**	4.09E-02	+	-	3
A_23_P42868	NM_000596	**IGFBP1**	4.09E-02	-	+	2.8
A_23_P326009	NM_020813	**ZNF471**	4.11E-02	-	+	3
A_23_P6413	NM_080430	**SELM**	4.13E-02	-	+	2.6

A_33_P3215550	NM_001010939	**LIPJ**	4.14E-02	-	+	2.8
A_24_P174793	NM_000439	**PCSK1**	4.20E-02	-	+	2.8
A_23_P47058	NM_022034	**CUZD1**	4.24E-02	+	-	2.7
A_23_P126167	NM_002249	**KCNN3**	4.26E-02	-	+	2.8
A_23_P717	NM_018252	**TMEM206**	4.26E-02	-	+	2.6
A_33_P3369505	NM_001031740	**MANEAL**	4.28E-02	+	-	3
A_32_P16007	NM_207355	**POTEB**	4.38E-02	+	-	3.2
A_33_P3252083	NM_021150	**GRIP1**	4.42E-02	+	-	3.1
A_33_P3314902	NM_016252	**BIRC6**	4.44E-02	+	-	2.2
A_24_P365515	NM_021784	**FOXA2**	4.45E-02	+	-	2.8
A_33_P3256914	ENST00000442458	**ITFG3**	4.46E-02	+	-	3.4
A_33_P3292043	NM_001005214	**LRRC52**	4.50E-02	+	-	3.2
A_23_P166884	NM_000172	**GNAT1**	4.53E-02	+	-	3.5
A_24_P332081	NM_001105521	**JAKMIP3**	4.53E-02	+	-	6
A_24_P102053	NM_002538	**OCLN**	4.56E-02	+	-	3.9
A_24_P147242	NM_145893	**RBFOX1**	4.57E-02	+	-	3
A_23_P109913	NM_006564	**CXCR6**	4.62E-02	+	-	3.1
A_23_P157914	NM_153267	**MAMDC2**	4.66E-02	-	+	3.2
A_23_P115444	NM_005092	**TNFSF18**	4.67E-02	+	-	2.6
A_24_P942481	NM_180989	**GPR180**	4.73E-02	-	+	2.7
A_23_P250251	NM_000605	**IFNA2**	4.74E-02	+	-	2.5
A_23_P139527	NM_002150	**HPD**	4.78E-02	+	-	4.5
A_24_P299685	NM_198389	**PDPN**	4.79E-02	-	+	3.4
A_33_P3409944	NM_005272	**GNAT2**	4.80E-02	+	-	2.8
A_23_P55400	NM_003771	**KRT36**	4.82E-02	+	-	3.6
A_32_P423691	NM_001018082	**ADIG**	4.84E-02	-	+	3.3
A_23_P110957	NM_001452	**FOXF2**	4.87E-02	+	-	2.6
A_24_P923945	ENST00000375218	**UBR4**	4.87E-02	-	+	3.1
A_23_P96291	NM_004988	**MAGEA1**	4.89E-02	-	+	2.8
A_23_P93213	NM_153320	**SLC22A7**	4.92E-02	+	-	3.9
A_23_P209700	NM_006056	**NMUR1**	4.93E-02	-	+	4.1
A_24_P297551	NM_178539	**FAM19A2**	4.94E-02	+	-	3
A_23_P111888	NM_138455	**CTHRC1**	4.95E-02	-	+	3.3
A_23_P353574	NM_133494	**NEK7**	4.97E-02	-	+	2.8
A_23_P33664	NM_022142	**ELSPBP1**	4.98E-02	+	-	5.7
A_23_P216756	NM_004108	**FCN2**	4.99E-02	+	-	3.5



A heat-map of the 414 differentially expressed genes identified by ANOVA statistical test.

Appendix 10: Sixty five genes out of the 414 were detected shared between NSCLC and HNSCC according to literature mining.

Probe ID	Accession number	**Gene symbol**	*p* - value (ANOVA)	**Effect of HOXD10 over-expression**	**Effect of HOXD10 silencing**	Fold change
A_23_P23074	NM_006417	IFI44	4.38E-06	-	+	4.6
A_23_P133408	NM_000758	CSF2	6.36E-06	-	+	6
A_23_P36611	NM_181861	APAF1	2.76E-05	-	+	2.4
A_33_P3221748	NM_001031680	RUNX3	2.89E-05	-	+	3
A_33_P3243887	NM_000641	IL11	4.37E-05	-	+	6.3
A_33_P3745146	NM_001098517	CADM1	4.85E-05	-	+	2.5
A_23_P399078	NM_000362	TIMP3	8.61E-05	-	+	2.8
A_23_P134176	NM_001024465	SOD2	1.29E-04	-	+	3.4
A_33_P3290567	NM_003390	WEE1	2.17E-04	-	+	2.6
A_23_P300056	NM_044472	CDC42	2.18E-04	-	+	2.5
A_23_P94533	NM_001912	CTSL1	3.50E-04	-	+	2.5
A_33_P3718269	NR_029701	MIR146A	4.11E-04	-	+	2.8
A_23_P8913	NM_000067	CA2	4.67E-04	+	-	2.7
A_33_P3357658	NM_003484	HMGA2	6.21E-04	-	+	2.8
A_23_P212508	NM_001063	TF	7.46E-04	+	-	3.7
A_23_P16469	NM_001005377	PLAUR	8.84E-04	-	+	2.8
A_33_P3379644	NM_000689	ALDH1A1	8.98E-04	+	-	2.6
A_24_P82106	NM_004995	MMP14	1.08E-03	-	+	2.7
A_23_P218111	NM_001002236	SERPINA1	1.48E-03	-	+	3.1
A_24_P79403	NM_002619	PF4	2.75E-03	-	+	2.7
A_23_P161624	NM_005438	FOSL1	2.97E-03	-	+	3.3
A_24_P217572	NM_001957	EDNRA	3.17E-03	-	+	3.9
A_23_P392470	NM_000901	NR3C2	3.46E-03	+	-	3.1
A_24_P330518	NM_001218	CA12	3.96E-03	-	+	6.1
A_23_P155979	NM_001963	EGF	4.19E-03	-	+	3.2
A_24_P416346	NM_001079675	ETV4	4.57E-03	-	+	2.8
A_23_P62953	NM_002585	PBX1	4.59E-03	+	-	3.4
A_23_P148473	NM_000206	IL2RG	5.97E-03	-	+	3.2
A_23_P203558	NM_000518	HBB	7.72E-03	+	-	3.4
A_23_P161698	NM_002422	MMP3	8.09E-03	-	+	3.3
A_23_P118158	NM_006043	HS3ST2	8.43E-03	+	-	3.7
A_33_P3223592	NM_000041	APOE	1.01E-02	-	+	2.8
A_24_P605190	NM_005023	PGGT1B	1.02E-02	+	-	2.6
A_23_P106024	NM_002226	JAG2	1.19E-02	+	-	2.8
A_23_P60146	NM_006207	PDGFRL	1.19E-02	-	+	2.2
A_23_P134454	NM_001753	CAV1	1.20E-02	+	-	2.3
A_23_P83028	NM_021111	RECK	1.20E-02	-	+	4.2
A_23_P63896	NM_000043	FAS	1.21E-02	-	+	2.7
A_23_P70818	NM_005631	SMO	1.25E-02	+	-	3.1
A_23_P207213	NM_000691	ALDH3A1	1.27E-02	+	-	4.1
A_33_P3271350	NM_003640	IKBKAP	1.48E-02	+	-	3
A_33_P3370424	NM_017617	NOTCH1	1.78E-02	+	-	2.1
A_23_P62115	NM_003254	TIMP1	1.81E-02	-	+	3
A_23_P356070	NM_002020	FLT4	1.85E-02	+	-	2.8
A_23_P86599	NM_007329	DMBT1	1.97E-02	-	+	2.9
A_23_P160668	NM_005376	MYCL1	2.01E-02	+	-	3.6
A_33_P3369844	NM_013230	CD24	2.15E-02	+	-	3
A_23_P374695	NM_000459	TEK	2.17E-02	-	+	6
A_33_P3423941	NM_003641	IFITM1	2.25E-02	-	+	4.6
A_33_P3395947	ENST00000380922	IL4R	2.46E-02	+	-	3.2
A_23_P406424	NM_175744	RHOC	2.48E-02	-	+	2.1
A_32_P193322	NM_152756	RICTOR	2.66E-02	-	+	2.5
A_33_P3232692	NM_001185156	IL24	2.71E-02	-	+	4.1
A_23_P167096	NM_005429	VEGFC	2.81E-02	-	+	3.6
A_23_P101131	NM_002091	GRP	2.82E-02	+	-	2.8
A_23_P36397	NM_000785	CYP27B1	2.96E-02	-	+	4.4
A_23_P308058	NM_172367	TUSC5	3.12E-02	+	-	4
A_33_P3407013	ENST00000420258	IL6	3.17E-02	-	+	3.4
A_23_P18447	NM_013261	PPARGC1A	3.79E-02	+	-	3.5
A_23_P204640	NM_024865	NANOG	3.87E-02	+	-	2.9
A_23_P151710	NM_000956	PTGER2	3.94E-02	-	+	3.4
A_23_P42868	NM_000596	IGFBP1	4.09E-02	-	+	2.8
A_24_P365515	NM_021784	FOXA2	4.45E-02	+	-	2.8
A_23_P250251	NM_000605	IFNA2	4.74E-02	+	-	2.5
A_24_P299685	NM_198389	PDPN	4.79E-02	-	+	3.4

Appendix 11: Forty nine EMT-associated genes affected by HOXD10 manipulation.

Probe ID	Accession number	**Gene symbol**	*p* - value (ANOVA)	Effect of HOXD10 over-expression	Effect of HOXD10 silencing	Fold change
A_23_P31721	NM_001951	E2F5	4.88E-11	-	+	2.1
A_33_P3339650	NM_001723	DST	2.65E-08	+	-	2.2
A_33_P3263890	NM_006902	PRRX1	6.74E-08	-	+	2.6
A_23_P399078	NM_000362	TIMP3	7.02E-08	-	+	2.8
A_23_P128574	NM_017993	ENOX1	1.44E-07	+	-	2.1
A_23_P62807	NM_016002	SCCPDH	7.61E-07	-	+	2.2
A_33_P3249534	NM_005382	NEFM	1.10E-06	+	-	3.2
A_23_P8913	NM_000067	CA2	1.99E-06	+	-	2.7
A_23_P433016	NM_001996	FBLN1	3.13E-06	+	-	2.4
A_32_P4018	NM_005012	ROR1	5.32E-05	-	+	2.1
A_23_P46936	NM_000399	EGR2	8.02E-05	+	-	2.3
A_23_P429998	NM_006732	FOSB	8.70E-05	-	+	2.3
A_23_P43810	NM_206943	LTBP1	1.19E-04	-	+	2.8
A_33_P3377364	NM_000213	ITGB4	1.38E-04	-	+	2.2
A_23_P137856	NM_002456	MUC1	2.25E-04	+	-	2.4
A_24_P787897	NM_022166	XYLT1	3.18E-04	-	+	2.8
A_23_P118158	NM_006043	HS3ST2	4.93E-04	+	-	3.7
A_23_P156327	NM_000358	TGFBI	5.91E-04	-	+	2.4
A_23_P106024	NM_002226	JAG2	6.36E-04	+	-	2.8
A_33_P3243887	NM_000641	IL11	7.06E-04	-	+	4.3
A_23_P128323	NM_001038	SCNN1A	7.40E-04	+	-	2.2
A_32_P69368	NM_002166	ID2	7.75E-04	+	-	2.6
A_24_P162173	NM_001149	ANK3	8.64E-04	-	+	2.4
A_33_P3285545	NM_001305	CLDN4	1.08E-03	-	+	2.2
A_23_P6771	NM_014583	LMCD1	1.22E-03	-	+	3.0
A_24_P357169	NM_031308	EPPK1	1.29E-03	+	-	2.3
A_23_P36397	NM_000785	CYP27B1	1.84E-03	-	+	4.3
A_23_P72737	NM_003641	IFITM1	2.31E-03	-	+	3.7
A_33_P3408913	NM_001127380	SAA2	2.67E-03	-	+	2.4
A_23_P151710	NM_000956	PTGER2	2.86E-03	-	+	3.5
A_23_P13772	NM_016569	TBX3	3.21E-03	-	+	2.2
A_33_P3214105	NM_001674	ATF3	4.75E-03	-	+	2.1
A_23_P164284	NM_001307	CLDN7	5.37E-03	+	-	2.6
A_33_P3379456	NM_015677	SH3YL1	6.68E-03	+	-	2.5
A_23_P15394	NM_001251	CD68	7.60E-03	-	+	3.7
A_24_P354689	NM_004598	SPOCK1	9.00E-03	-	+	2.9
A_23_P252306	NM_002165	ID1	1.27E-02	+	-	2.3
A_23_P140830	NM_024712	ELMO3	1.49E-02	+	-	2.4
A_23_P214330	NM_030666	SERPINB1	1.59E-02	+	-	2.0
A_24_P345846	NM_058172	ANTXR2	1.85E-02	-	+	2.2
A_23_P29939	NM_007308	SNCA	1.99E-02	-	+	2.5
A_32_P151544	NM_000224	KRT18	2.24E-02	+	-	2.2
A_23_P108157	NM_014428	TJP3	2.59E-02	+	-	3.1
A_24_P319364	NM_016946	F11R	2.81E-02	+	-	2.5
A_33_P3321642	NM_024915	GRHL2	2.92E-02	+	-	2.2
A_24_P102053	NM_002538	OCLN	3.83E-02	+	-	3.9
A_23_P62901	NM_006763	BTG2	3.83E-02	-	+	2.3
A_23_P150147	NM_017857	SSH3	4.58E-02	+	-	2.1
A_24_P398147	NM_006393	NEBL	4.95E-02	+	-	3.2

Appendix 12: DNA sequences of the cloned promoters including putative HOXD10 binding sites.

>AMOT

AGTCAACTTCATATCCACCCCCAAAATTGTAGGAAAAGAGTTGACTTTTTAAAATAAAGCCACAAAACCTACACGACCTTGGATGTAAATG
CTTAAAACAGATGAGCAAAATATTGATCCTCTAACACATCTAAGACTCTCCATTATGTTTCACACACAAGTAGGAACACATACATGTGTGG
AGTGCTTTATAGTTTTCACACACATTTTATTTTCATACCAGTGCCTTTAAGTTAATAGGGCATGTCTTACAGCATTTACTTACAGGTAGAA
AAATTAAGGCTGAGAAGTTAAGCCACTTAACAGAAGTCATACAATTAGCTAGTAAGTGCCAGGGCCCCAAATGAATTCTTCCTTTGGTTCC
AAATTCCATGTTATTTTGCTACCCTGACACTGTTTCTTTGTGCCTTCGTTTTTCTAGATGTAAAATGGTAGATTTTCCGTGTCCCAGGATG
CTGTTCTCAACCCTCTCGTTTTTAACCAATGTCAGCTTTTTTCACAGTCAGCAGAACACCACAATGACACTTCATGTACTTTCCCCCCTCC
TGATTTAATAAACTAAGGGGGAAGAAAAGTTACCTATTCGCAGAAAGCTGGACAGTGCCATAAACATAAAGGCGTACCAGAGTCGGCCGAC
ACACTGTTGCACTCTTTTATGAGACTGAGAAATAAAGCGTCTCATCTTTCTCTTGCCAGCAGTAGCTGGGGTAGAACATAGAACTACACAA
TAGCCTCTTGTTTAGTTTTATTAATTTTGAGGGCGGGTGGGACAAAGGGATCAGACTGGCGACTCCAGGGCCCAATCGCAAGCGTATACTA
GGCCAGGCTGGGCCAACCGGTAGCTGCAAAAGAATGGTGGGCGGGACTTACGCTTCAGGGACGCGCTCCGCCCAGGCCGGCGCTCCCAGCC
CGCCCAGGGGCGGAGCCTCCGCGGAGCGCCGCCCGCGACCCCTCCTCCCTCGCCCCTCAGCTTCTGCGGCCCGCCAACGGGAGGGGGGCAA
GCGCCCGGGGCGGGGGAGCGGAGGAAGAAAGGAGGGTGTCTCGCAGAGCTGTGACCTGGGTGGGGGCCTGCTGGTAATAGGAAGCAAGAGG
GGAAGGAAAACGAAGGAGCGCTTCTCCTTCCTGCATTATTTTAAGCAGGCTTTGGTGGAAGCTGCTTTTCAGCATTTCTATTATTTGTTTT
TCTTCCTCCCGCAGTGGGGCGGGGAGCTGCACTGGCTCTCCCACGCGTCTATGCCGCGCTGTTGCCGCAGATTATTATTGTAGTCTTTTGC
AAAACCCTTTCCTTCTTCCCAGCCAGTCCCCGCGCAGACTCTCACGGCTACGGCAACCGAGTCGTGCCCCTCTAGTGGGAAGGCGAGAGCG
AGGGTCCCTTGCTTTCGCCCGAGCTCCCGGACCCCGCGGTTCAATTGTTTGAAACTTCCACGCCCCCCGGCCCCGCCTCTCTTCCGCGCTC
CCCTCACTCCTCACTTCCAGGGTCCAGGCAGGCTCCCGCCCTTCTTTCCCCGCCCTCTACTCCGCCTTCCTCCCTCGTCCACCCCTCCCGC
CGCCCTGCTCCGGGGCGGTCCTAGCCGCTCACCGCAAGAAGAAAAGCCACCCGTTCGTCCCCCTCCCCTACCTCCCATCCTTTGCCCAGGA
GCTGCCTTGGCAGTCACGCCCCTTCCTTCCGAGGAGCTTTCTGGCTGCCTAAACTGGTAGACCCCCTGAATTACTCCTCCATCTCCGCTCT
CTTTCGCCTCCTCTTCTCTTAGTTCTCTCCGCCTCCCCCTCAACTACCACCACCTCCAGTCAGTCTCGCCTCCGGCTATCCGCTGCTCCAC
CCTCTGGCCCGGTATCCTGCCTGTCCGCTGCCACCAAGGAGAGCCCGGACGGAGCAGCGAGGAGGGGAGCAGCCGGGAGTTGGGGCTTCCC
CCCTGCCCATCCCTGGCCGCTGGCCCGGGACCGAAGCCACTTGAGCGAGCAGAGAGTCGTCACCTTGTCTTCTTTGCCTTCAGGGGTAGGA
GACGCGTGGTTACTTTTCGCGGGGAACTGGGGCAGGAATCCATGGTGGCATGAACTTGGTCGT

>miR-146a

TGAAAAGCCAACAGGCTCATTGGGCAGCCGATAAAGCTCTCGGGATTTCCCCGCGGGGCTGCGGAGAGTACAGACAGGAAGCCTGGGGACC
CAGCGCCTGACCAGAACTTCCTCGGGGGAGGCTGCAGGGGAGCAGGCGCATCCTGCACAGAACGCTCTAGAGCGCGCAGGCCAAAGCACCA
GGCTGCTCCTGACACGTGCTGCAAGAGGGTCCCCGACCCGGGGGTCCAGACCCTGCACGCATGATGGGGAAGGTGGAGGCTTCCCCCTCAG
CTCCGCGGAGAGAAGCTGACACTGCCAGGCTGGAACCTTCCATTCCGGCCCAGCCTCTTCCTCCCTCGCTGTGCCGAGGAGGGATCTAGAA
GGGACTTTCCAGAGAGGGTTAGCGTGCAGGGTGTGGAAATGGAATAAAAGCATATGCAAATAGGCCTTAGCTGCCTTCCTCTACCCCAGCA
AATAAGAGTCTCTCCAGAAAGATGCTCTTTCTCCAAGACGCTTGACCGCTCTTCCTTTCCTGGATGGCACCAGCAGGGCCGATTGGAGTGG

>CYP2J2

CAGCTGTAGTAAAACAGGAAAATGTGGACTAGAAGGATAAAAGCTAAGCTCTGACTAGTATTTCTCTCTGGACCTTAGAAAGAGAACAAGA
GACTTCAACTCTCATTTCATTTCTTGCTCACACCTGTAATCCTAGCACTTTAGGTGGCTGAGGTGGGCAGATCACCTATGGTCAGGAGATT
GAGACCAGCCTGACCAACACGGCGAAACCCCATCTGTACCAAAAATACAAAAATTAGCCGGGCATGGTGGCATGTGCCTGTAGTCCCAGCT
ACATGGGAGGCTGAAGCAGAAGAATCGCTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAACTCCAGCCTGGGCGACAGAGCAAGACTC
TGTCTCAAAAATAAAGAAATAAAATAAAATATTGTAAGCAGAATATGTGACTCTATGTTGAGTACATGAGTGTGTACTATACTATTCCTGA
GCTATGTTTCATGTTTTCAGTCTCACCATTTGTGAAACCCAAGCAGGCACTCACCAGATGGTATGAAAGGCTGGAACCTGGAGGTCCCACA
GTCTTGGGGCTTTGGAGAGGAGCCTAATAGAGCAAACCAGCAAGGATCGGAACTAAGACATTCCTTCAACAAATGTTAATTAAGCACTTAT
TATGGCTGCTGCTCAGTGCTAGCAGCTTTGGTGTAGTGCCTTTTAATCTTCCAGACACTTTCAAATACAGGACCTCGTTTAGATGTGTTTC
AGAAGGGTGGATTTATTTGTTCACCACATACTTAGTACTTAACTTTTCTGTTTCCTCAACAGCAAGATGAGACTACCGAGAGCTTCTCCTT
CATAATTTTGTCAAAATGAAATAATAATCTATGTAAAGAGGCGGATGTGCTGCCTGGCATATAGTAAGAGCCCAAGAAATGTTAGGTACTG
CTGTTGTTATCGCTGTTTACCAGGCACAGTTCCAATCATAGGAGAGACGGTGATTGAACCGAACAGAGGCTCAAATGTCTGTGGCGCTCTC
AGTAAACCCATGAGCTTTCTAAATAGAAAAGAAACTGCTCCTAGCCTGGCCTTTTCTGAGACCGGTGCGTGTCCTCCCGGGAATCCAGCGC
CTGGCATCTTCGCAGGGTGCTGCGAAGGGGCGGGCTGGGAGGCGGGGCACGGCTGGGAGCGAGGCGGGGCGGGGACCGTCGCCTGCTGGGA
CCGCCGCCTGCTTGGACCGCAGAAGAGCAGGAGGACGTCTGAGCCATGCTCGCGGCGATGGGCTCTCTGGCGGCTGCCCTCTGGGCAGTGG
TCCATCCTCGGACTCTCCTACTGGGCACTGTCGCCTTTCTGCTCGCTGCTGACTTTCTCAAAAGACGGCGCCCAAAGAACTACCCGCCGGG
GCCCTGGCGCCTGCCCTTCCTTGGCAACTTCTTCCTTGTGGACTTCGAGCAGTCGCACCTGGAGGTTCAGCTGGTAGGAGTGGGAGAAGGT
GCCTAGCGTGTCCTGACCCTAACTCTGTTCTGCAGC