FOXA Genes in Lung Cancer In Vitro and In Vivo

Thesis submitted for the degree of

Philosophiae Doctor

by

Audun Trygge Haugen Bersaas

Department of Chemical and Biological Work Environment

National Institute of Occupational Health

Faculty of Medicine

University of Oslo

STAMI
STATENS ARBEIDSMILJØINSTITUTT

NORWEGIAN **CANCER** SOCIETY

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-022-3

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

ACKNOWLEDGEMENTS

The work presented in this thesis was carried out at the National Institute of Occupational Health (STAMI), Oslo. I am grateful to the Norwegian Cancer Society and STAMI for the financial support.

I would like to thank my supervisor, Dr. Steen Mollerup, for all your support throughout this project. Your enthusiasm for the field and excitement over new results have been truly motivating. And when things do not go as planned, you see solutions and possibilities. Thank you for teaching me the art of scientific writing. And thank you for all the nice conversations we have had.

Thank you to my supervisor, Professor Aage Haugen, for invaluable help with the writing of the manuscripts. Thank you for the discussions and suggestions for the project and the thesis. Thank you to Professor Arne Klungland for being my supervisor at the University of Oslo. Thank you to all the co-authors, especially to Mari Sjøberg for sharing long days in the cell lab with me during the transformation study.

Thanks to all my amazing colleagues at Toks! A special thank you to Rita Bæra. Thank you for teaching me everything I know in the lab, and for always being positive and being wonderful to share an office with. Thank you to my fellow PhD-student, Iselin Rynning, for all our scientific and non-scientific discussions. It has been great sharing the PhD-experience with you. Thank you to all the master students at STAMI: Ida, Jenny, Furqan, and Oda.

Thank you to all my friends and everyone in Majorstuveien 35.

To my wonderful parents, Maila and Gunnar Bersaas, thank you for your everlasting support. Thank you for your encouragement and for believing in me. Thank you to my dear sister, Monica Bersaas, for always joking around when I need some cheering up and for being a good friend.

To my boyfriend, Sondre Wittek, you are amazing! Thank you for making me relax when I'm stressed. Thank you for making me dinner while I am working. Thank you for also doing the dishes. Thank you for your encouragement, support, and love; it means everything to me.

INNHOLDSFORTEGNELSE

ABBREVIATIONS

ADCA	Adenocarcinoma
AHR	Aryl hydrocarbon receptor
AKR	Aldo keto reductase
AR	Androgen receptor
ARNT	AHR nuclear transferase
B[*a*]P	Benzo[*a*]pyrene
BPDE	B[*a*]P-7,8-diol-9,10-epoxide
CDH	Cadherin
CDK	Cyclin dependent kinase
ChIP	Chromatin immunoprecipitation
CNV	Copy number variation
CpG	Cytosine-guanine dinucleotides
CSC	Cigarette smoke condensate
CYP	Cytochrome P450
DHD	Dihydrodiol
DNA	Deoxyribonucleic acid
DNMT	DNA methyl transferase
E2	17-b-estradiol
EGFR	Epidermal growth factor receptor
EH	Epoxide hydrolase
EMT	Epithelial to mesenchymal transition
ER	Estrogen receptor
ERE	Estrogen response element
FOXA	Forkhead box A
GST	Glutathione-S-transferase
H1, H2, H3, H4	Histone variant 1, 2, 3, 4
HBEC-KT	Human bronchial epithelial cells, CDK4 and TERT immortalized
HDAC	Histone deacetylase
LC/MS/MS	Liquid chromatography tandem mass spectrometry
LCC	Large cell carcinoma
MeO-E2	Methoxyestradiol
MET	Mesnechymal to epithelial transition

miRNA	MicroRNA
MNU	Methylnitrosourea
mRNA	Messenger RNA
nAChR	Nicotinic acetylcholine receptor
NNK	Nicotine-derived nitrosamine ketone
NSCLC	Non-small cell lung cancer
OH-E2	Hydroxyestradiol
PAH	Polycyclic aromatic hydrocarbon
PM	Particulate matter
RB1	Retinoblastoma 1
RNA	Ribonucleic acid
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
siRNA	Short interferring RNA
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variation
T12KT	Transformed HBEC12
T2KT	Transformed HBEC2
TCDD	2,3,7,8-Tetrachlorodibenzo-*p*-dioxin
TCGA	The cancer genome atlas
TERT	Telomerase reverse transcriptase
TF	Transcription factor
TKI	Tyrosine kinase inhibitor
TP53	Tumor protein 53
TSS	Transcriptional start site
UGT	UDP-glucuronosyltransferase
UTR	Untranslated region
XRE	Xenobiotic response element

SUMMARY

Lung cancer is the leading cause of cancer-related deaths worldwide. Lung cancer survival rates have remained poor, mainly due to late diagnosis at advanced stages in addition to scanty treatment options. The main cause of lung cancer is tobacco smoke, accounting for up to 90% of all cases. Exposure to other environmental carcinogens also contributes to the disease. It is important to understand the underlying mechanisms of carcinogen-induced lung cancer to establish better treatment targets and novel biomarkers for cancer detection.

This thesis aimed at identifying the role of the two proteins FOXA1 and FOXA2 in lung cancer and lung cell transformation. FOXA1 and FOXA2 are pioneer transcription factors that regulate gene expression by binding forkhead DNA-motifs in condensed chromatin resulting in an open chromatin structure. The pioneer factors further guide other transcription factors to their respective binding sites where they exert their regulatory function. FOXA1 and FOXA2 have been implicated in maintaining an epithelial phenotype, and their altered expression has been indicated in different cancer types.

An *in vitro* transformation model using immortalized, non-transformed human bronchial epithelial cells (HBECs), was utilized for studying FOXA1 and FOXA2 during carcinogen-induced cell transformation. Exposure of two HBECs (HBEC2 and HBEC12) to cigarette smoke condensate, benzo[*a*]pyrene (B[*a*]P), and methylnitrosouera resulted in transformed cells with the acquired ability to grow anchorage-independent in soft-agar. All cell colonies isolated from soft-agar showed diminished transcription of both *FOXA1* and *FOXA2*. Also, the transformed cells displayed features of epithelial to mesenchymal transition, evident by gene expression patterns, altered morphology, and increased invasiveness.

The expression levels of *FOXA1* and *FOXA2* was next studied in matched tumor and non-tumor tissue *in vivo* from three geographically independent cohorts from Norway, Italy, and the USA. Lung tumors expressed significantly higher levels of *FOXA1* than matched non-tumor lung tissue. A small subset of the tumors harbored

FOXA1 gene amplifications, accounting for the higher *FOXA1* mRNA levels in this subgroup. The lung tumors expressed significantly lower levels of *FOXA2* compared to non-tumor. CpG-islands located in or near the *FOXA2* gene were more methylated in tumors than in non-tumor tissues. Two publicly available databases (TCGA and CURELUNG) confirmed increased DNA-methylation in *FOXA2*. The changes in DNA methylation suggest that lung tumors suppress *FOXA2* activity through epigenetic regulation.

The transformed cell lines showed altered expression of *CYP1A1* and *CYP1B1*. A study using RNA interference to knock down *FOXA1* was performed to investigate the role of FOXA1 in the regulation of CYP450 enzymes and the implications for B[*a*]P and estradiol metabolism. Knock-down of FOXA1 led to increased expression of *CYP1B1* but had no apparent effect on *CYP1A1* expression. Chromatin immunoprecipitation identified FOXA1 binding in a *CYP1B1* enhancer. No binding of FOXA1 was observed in *CYP1A1* regulatory regions. B[*a*]P and 17β-estradiol metabolites were measured by LC/MS/MS in cell culture media and B[*a*]P-DNA and ^{32}P-postlabeling was used to measure adduct levels. These analyses did not identify any changes in metabolism or DNA-adduct formation in cells with knocked down FOXA1, in spite of increased expression of *CYP1B1*. These findings suggest that FOXA1 can regulate CYP1B1 expression, but this regulation may not be sufficient for affecting metabolic capacity.

Taken together, the three studies conducted for this thesis support a role of the transcription factors FOXA1 and FOXA2 in lung cancer *in vitro* and *in vivo*.

LIST OF PAPERS

Paper I

Audun Bersaas, Yke Jildouw Arnoldussen, Mari Sjøberg, Aage Haugen, Steen Mollerup, Epithelial-mesenchymal transition and FOXA genes during tobacco smoke carcinogen induced transformation of human bronchial epithelial cells, *Toxicology in Vitro, Volume 35, September 2016, Pages 55-65, ISSN 0887-2333, http://dx.doi.org/10.1016/j.tiv.2016.04.012.*

Paper II

Audun Bersaas, Antonella Galvan, Ana I. Robles, Elise D. Bowman, Rita Bæra, Vidar Skaug, Tommasi A. Dragani, Curtis C. Harris, Aage Haugen, and Steen Mollerup, Altered expression of FOXA1 and FOXA2 in non-small cell lung cancer: genetic and epigenetic mechanisms, and diagnostic implications, *Manuscript.*

Paper III

Audun Bersaas, Jiří Neča, Volker Arlt, Miroslav Machala, David H. Phillips, Aage Haugen, Steen Mollerup, The role of FOXA1 in regulation of CYP1A1 and CYP1B1 expression in human bronchial epithelial cells, *Manuscript.*

1 INTRODUCTION

1.1 Lung cancer

1.1.1 Lung cancer epidemiology

Lung cancer is one of the most diagnosed cancer types worldwide with an estimated 1.8 million new cases each year. Lung cancer accounts for nearly 20% of all cancer deaths, the highest number of any cancer type. Although lung cancer has been subject to much research over several decades, the 5-year survival rate (10-20%) remain low, so better diagnostic tools and treatment options are needed [1-3].

Tobacco control is essential to prevent lung cancer. Cigarette smoking habits relate to the distribution of lung cancer throughout the world. The Norwegian cancer registry reported 3019 new cases of lung cancer in Norway in 2014 [4]. As in many other western countries, the incidence rates of lung cancer in Norway has stabilized for men, while the number of women diagnosed with lung cancer has increased 10-fold over the past 60 years and still continues to rise. The gender difference reflects the trends in smoking habits, with an increasing proportion of smokers today being women [4-6]. Epidemiological studies have also identified a possible higher risk of developing lung cancer in female smokers compared to male smokers [7, 8]. More recent studies have not been able to reproduce such findings, indicating that, given the same exposure to tobacco, men and women have a similar risk of developing lung cancer [9-12]. In Norway, lung cancer is the leading cause of cancer deaths with a recorded number of 1198 deaths for men and 960 deaths for women in 2014 [4]. In Europe in general, the lung mortality rate is decreasing in men and increasing in women [13].

There are two main types of lung cancer: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). NSCLC accounts for 85% of total lung cancer cases [14]. NSCLCs comprise squamous cell carcinoma (SCC), adenocarcinoma (ADCA), and large-cell carcinoma (LCC). All lung cancer types correlate with tobacco smoking, but ADCA is also the most common lung cancer type in never-

smokers [15]. Approximately 10-15% of ADCAs are observed in never-smokers [16, 17].

1.1.2 Lung cancer etiology

It has been well-known for several decades that tobacco smoking can cause lung cancer [18]. Both current and former smokers appear to have an increased risk of developing lung cancer, and the risk increases with amount smoked, duration of smoking, and with earlier age of starting to smoke [12]. In Western countries, smoking habits are declining while smoking is increasingly common in economically developing countries in Africa and Asia [19]. More than 70 cigarette smoke chemicals are considered carcinogenic in laboratory animals, and IARC classifies 16 of these as carcinogenic to humans [20, 21]. Polycyclic aromatic hydrocarbons (PAH) and the tobacco-specific N-nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) are considered major etiological factors in lung cancer. DNA adducts from PAH and NNK are found in smokers [21-23].

Besides cigarette smoke, environmental and occupational risk factors contribute to the burden of lung cancer [24, 25]. PAHs form during incomplete combustion of organic material and is present in several sources including indoor combustion and cooking, diesel exhaust, and urban air particulate matter (PM). Other known causes of lung cancer include exposure to second-hand smoke, radon, asbestos, silica, and ionizing radiation [20, 26, 27]. Studies have identified that occupational exposure may contribute to lung carcinogenesis by altering pathways related to signal transduction, immune process and carcinogen metabolism [26].

1.2 Lung carcinogenesis

Cancer occurs when the transcriptional program of a cell deviates from normal cell function to an altered state that promotes growth, invasiveness, and metastasis. Transcriptional reprogramming arises as a result of mutations and epigenetic changes accumulated through the individual's lifetime. The microenvironment in the different tissues of a multicellular organism applies different selection pressure on the genetic and epigenetic variation within the cell population [28]. The selection of somatic variation results in a Darwinian evolution which generates a heterogeneous tumor cell

population. Tumor cells possess traits to overcome the hallmarks of cancer, including continued proliferation, replicative immortality, the resistance of cell death, evasion from immune destruction, induction of angiogenesis, deregulation of cellular energetics and the capability to invade and metastasize [29-31].

1.2.1 Somatic mutations

DNA sequencing has revealed that cancer types associated with mutagen exposure have a higher mutational load than cancer types not related to mutagen exposure. For instance, lung cancer and skin cancer relates to exposure to tobacco smoke and UV-radiation, respectively [32, 33]. Lung tumors from smokers have approximately ten times more mutations than lung tumors from non-smokers [34]. Mutations can alter the function or transcription level of genes that suppress (tumor suppressors) or drive (oncogenes) tumor progression.

Several types of somatic mutations contribute to malignant transformation including single nucleotide variation (SNV), insertion and deletions (indels) of small or large DNA fragments, and copy number variation (CNV). The most prominent mutation type is SNV, and four mutational processes contribute in generating this kind of mutation; inaccurate DNA replication, exposure to exogenous or endogenous mutagens, enzymatic DNA modification and defective DNA repair [35]. Each process leaves a footprint in the DNA mutation. For instance, spontaneous deamination of 5-methyl-cytosine at NpCpG trinucleotides generates C>T transitions. The number of C>T mutations found in tumors correlates with the age of the patient, generating a clock-like mutational signature [36].

Exposure to mutagens produces distinct mutational signatures corresponding to the chemical or physical property of the mutagen. For example, cancers associated with tobacco carcinogens, such as lung cancer, head-and-neck squamous carcinoma and liver cancers, carry a high level of C>A mutations with a transcriptional strand bias [35, 37, 38]. The higher prevalence of C>A transversion is attributed to the effect of PAH metabolites, such as benzo[*a*]pyrene-7,8-diol-9,10-epoxide (BPDE, see section 1.2.4), that preferentially bind guanine and creates bulky DNA adducts. Next generation sequencing results showing C>A mutations in smokers is in line with

earlier reports on C>A mutations being the most common mutation in the tumor suppressor gene *TP53* in lung tumors [39]. Interestingly, different cancer histologies of the same tissue can have different mutational signatures. Lung adenocarcinoma and squamous cell carcinoma have distinct mutation patterns [40]. Mutational signatures can even differ within histologies. C>A transversions is the predominant mutation in smokers and the age-related C>T transitions being predominant in non-smokers [34].

Another frequently observed genetic alteration in cancer is CNV. CNVs are deletions or amplifications of DNA stretches larger than 1 kb [41, 42]. A recent pan-cancer analysis of 4934 tumors across 11 cancer types identified that each cancer on average had 39 CNVs, some of which were whole genome duplications [43]. Among the amplified genes found across different types of cancer were *MDM2*, which targets TP53 for degradation, and the epidermal growth factor receptor (*EGFR*) [44].

1.2.2 Epigenetic Alterations

Mutations can change the gene function if it occurs in coding regions, or the gene expression levels if it takes place in regulatory regions. Variations in the epigenome may also modify the gene expression in a cancer cell. The genetic and epigenetic factors can be subject to selection pressure independently as suggested by twin studies in which identical genotypes acquire different epigenotypes [45]. Combined genetic and epigenetic alterations are likely to contribute to cancer progression.

Epigenetics is defined as somatically heritable changes in gene expression that are not a consequence of changes in the DNA sequence. There are three different types of epigenetic regulation: histone modifications, DNA methylation, and non-coding RNA. Animal studies indicate that prenatal and early post-natal environmental factors, including nutritional supplements, xenobiotic chemicals, behavioral cues, reproductive factors, and low-dose radiation, can result in epigenetic re-programming and altered susceptibility to disease [46]. Epigenetic changes can be inherited mitotically in somatic cells, indicating a potential long-term effect of the environment on the epigenome [47].

Methylation of cytosine at the 5-position (5-methylcytosine, most often referred to as DNA-methylation) is typically considered a suppressive mark [48]. DNA-methylation occurs on cytosine-guanine dinucleotides (CpG). CpG-islands are long stretches of repetitive CpGs and occur regularly throughout the genome in promoters and enhancers. DNA methylation of the 5′-CpG islands of genes plays a major role in gene regulation. Under normal physiological conditions, DNA methylation modulates genome imprinting, X-chromosome inactivation, and inactivation of repetitive sequences [49].

In cancer, tumor suppressor genes are often found to have methylated promoters. The first report on a tumor suppressor gene being inactivated by DNA methylation was in the *RB1* gene, an important cell cycle regulator [50]. The researchers found that while most retinoblastoma patients had mutations in *RB1*, a subgroup did not but instead had hypermethylated RB1 promoters. Interestingly, mutations and DNA hypermethylation were mutually exclusive, indicating that different mechanisms can lead to the same impaired gene function resulting in similar functional implications [50].

DNA methylation in cancer cells is thought to occur in a deterministic rather than random manner, establishing predictable patterns of DNA-methylation [51-53]. DNA methylation patterns can separate non-tumor tissue from NSCLC and can further divide NSCLC tumors into phenotypically distinct subtypes such as SCC and ADCA [54, 55]. A high level of DNA methylation is found in a subset of adenocarcinomas in male smokers and relates to poor prognosis [56]. The patients with high accumulation of DNA methylation are sensitive to DNA demethylating agents such as 5-Aza-deoxycytidine and 5-azacytidine. Demethylating agents could potentially be a treatment option for lung cancer, as it already is for leukemia[56].

MicroRNAs are short, 22 nt long RNA molecules that regulate gene expression post-transcriptionally. MicroRNAs mark mRNA for degradation or inhibits protein translation by binding to the 3' untranslated region (UTR) of mRNA molecules [57-59]. Several microRNAs have been implicated as important in cancer development. As for protein coding genes, microRNA genes are classified as tumor suppressors or oncogenes (often called oncomirs). The expression level of miRNAs is globally down-

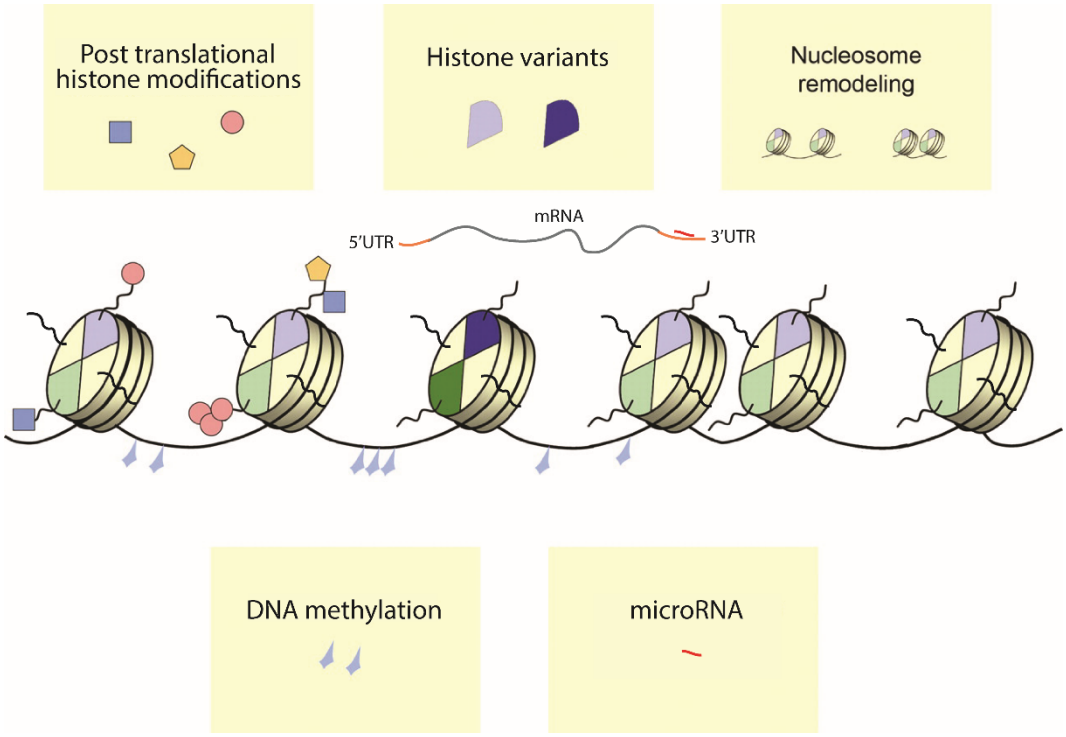

Figure 1: Epigenetic alterations. DNA is wrapped around an octamer of histone proteins. Histones can be made up of different types of histone variants that serve a specific function in regulating chromatin structure. The histone tail can be modified with chemical groups including methylation, acetylation, biotinylation, and phosphorylation. The DNA itself can be modified by the addition of a methyl group at the 5-position of cytosine in CpG dinucleotides. MicroRNAs also plays a pivotal role in regulating gene activity by binding to the 3' UTR of mRNAs which lead to mRNA degradation and inhibition of translation. The figure is modified from Gräff et al., and printed with kind permission from the authors [60].

regulated in tumor cells, associated with the formation of more aggressive tumors. This global down-regulation may function to repress the expression of miRNAs functioning as tumor suppressors [61].

The third epigenetic mechanism is histone modifications. Nucleosomes consist of 147 bp of DNA wrapped around an octamer of histone proteins. Four different types of histone proteins, H2A, H2B, H3, and H4, constitutes the histone complex and is present in two copies of each type. The histone proteins consist of a core region that binds the DNA and a protruding peptide sequence called the histone tail. The histone tail can be modified by the addition of different functional groups such as methylation, acetylation, phosphorylation, ubiquitination, sumoylation and biotinylation (Fig. 1). The functional implication of histone tail modification depends on the chemical group and which amino acid residue of the tail that is modified [62, 63]. For example, methylation of histone H3 at the Lys (K) 4 and Lys 9 residues is

associated with active and silent genes, respectively [64, 65]. The histone modifications establish a histone code that can be translated into transcriptional programs by TFs [62, 66]. The histone code defines functional segments of the chromatin. For instance are H3K79me2/3, H4K20me1, H3K4me1/2 and H3K9me1 present at promoters; H3K4me1, H2BK120ac, H3K27ac and H2BK5ac at enhancers; H3K79me3, H3K79me2, H3K79me1, H3K27me1, H2BK5me1, H4K20me1 and H3K36me3 in gene bodies; and H3K27me3 and H3K9me3 define heterochromatin [67].

In addition to modifications of the histone tails, chromatin state can be altered by inclusion of different histone variants (Fig. 1). For example, the histone H2A family consist of H2A.X, H2A.Z, and macroH2A variants. The histone H2A.X variant is involved in repair of DNA double-strand breaks [68]. Histone H2A.Z enrichment occurs at active intergenic enhancers and both active and suppressed promoters depending on the H2A.Z histone tail modifications [68, 69]. Normal mouse lung express higher levels of *H2afz* (H2A.Z) than lung tumors, and the presence of H2A.Z at transcriptional start sites (TSS) inversely correlates with DNA-methylation levels [70]. The macroH2A histone variant is considered a repressive mark and is involved in generating chromosome X barr-bodies in females [69, 71]. MacroH2A is often up-regulated in senescent cancer cells and is a potential biomarker for NSCLC that predicts good prognosis [72].

1.2.3 Drivers and passengers of carcinogenesis

Not all mutations and epigenetic changes that are present in a tumor cell are involved in establishing a malignant phenotype. Those alterations that contribute to carcinogenesis are called “drivers”. A cancer driver can be defined as a gene, mutation or epigenetic modification that gives the cell in which it occurs a selection advantage over the surrounding cells [73]. The driver mutations drive the cancer cell towards a malignant state by activating oncogenes and inactivating tumor suppressors. Driver mutations or epigenetic changes in regulatory regions may alter the gene expression program and give the cell a growth advantage [74].

Some mutations do not alter the growth advantage of the cell and are referred to as "passenger" mutations. Passenger mutations have become particularly evident through whole genome sequencing, which has revealed that most cancers, including lung cancer in never-smokers, have a relatively small number of mutations while cancers induced by carcinogens, such as lung cancer in smokers and melanomas, have a much higher mutation rate [35]. The lower number of mutations in other cancer types, such as leukemia and pediatric cancers, indicate that the carcinogen-induced mutations are not all necessary for cancer development, and most are likely passenger mutations [73, 75].

1.2.4 Cancer genes

Most mutations found in a cancer cell are passenger mutations and do not contribute to the cells selective advantage. The number of mutations needed for cell transformation is disputed, but findings suggest that the development of lung cancer requires very few mutations [76]. Only three genes are mutated in more than 10% of patients across the range of tumor types: *TP53*, *PIK3CA*, and *BRAF* [75]. *BRAF* is significantly mutated in ADCA but not in SCC. *TP53* and *PIK3CA* are together with *RB1*, *ARID1A, CDKN2A*, and *NF1* the only six genes frequently mutated in both ADCA and SCC [38]. Approximately one-third of all tumors across most cancers contain a mutated *TP53* [77]. In lung cancer, 46% of all adenocarcinoma patients and 81% of all squamous cell carcinomas have *TP53* mutations [78, 79]. TP53 is an essential protein in cell cycle regulation, senescence, apoptosis and maintenance of genome integrity and is considered a "Guardian of the Genome" [80]. When mutations disrupt *TP53* function, the cell will gain the ability to surpass cell cycle checkpoints and avoid senescence and apoptosis, thus acquiring a growth advantage.

Whole exome sequencing has revealed that there are mutational differences between the different lung cancer histologies and between smokers and non-smokers. For instance, in never smokers, *EGFR* mutation, and *ROS1* and *ALK* (anaplastic lymphoma kinase) fusions are observed, while in smokers, the predominant alterations include mutations in *KRAS*, *TP53*, *BRAF*, *JAK2* and *JAK3* [34]. *KRAS* and *EGFR* are frequently mutated in adenocarcinomas, but mutations in these genes

are rare in squamous cell carcinoma [78, 79, 81, 82]. The RTK/RAS/RAF signaling pathway, a downstream target of EGFR, is activated by mutations in 76% of all lung adenocarcinomas [79]. Activation of this oncogenic pathway leads to the continuous production of signals promoting cell proliferation and survival. Interestingly, mutations in *TP53* and *KRAS* tends to be mutually exclusive, suggesting that disruption of these genes leads to similar functional consequences [34]. Some cases of adenocarcinoma have no apparent oncogenic activation [79].

Several tyrosine kinase inhibitors (TKI) targeting two different proteins, EGFR and ALK, have been approved for treating NSCLC. *ALK* is translocated in 1-7% of all NSCLCs and is often fused to echinoderm microtubule associated protein like 4 (*EML4*) [83, 84]. Approximately 10-20% of North-American and European lung cancer patients and up to 60% of Asian patients have activating mutations in EGFR [81, 85]. Unfortunately, the development of resistance to EGFR-TKI is widespread and is associated with tumor progression [86]. Resistance to EGFR-TKI may evolve by an acquisition of SNVs (T790M, found in 50–65% of EGFR-mutant resistance cases) prior to or after TKI treatment [85, 87]. Notwithstanding the resistance, TKI is a first line treatment option that increases the median overall survival compared to conventional chemotherapy, from 23.6 to 30.5 months, respectively [88].

Squamous cell carcinoma genomes comprise a high rate of CNVs compared to other cancer types [78]. Lung SCC and ADCA have many of the same genetic and epigenetic alterations, but there are also some differences. For instance, chromosome 3q, containing *SOX2*, is frequently amplified in squamous cell carcinoma but not in adenocarcinoma [78, 89]. One of the functions of SOX2 in the non-cancerous cell is in the differentiation of squamous cells, and other genes involved in squamous differentiation (*TP63* and *NOTCH1*) are also frequently altered in SCC [78, 89]. The gene coding for the catalytic subunit of telomerase *TERT* has been found to be amplified across different types of cancers, including lung cancer [44, 90]. The two cyclin-dependent kinase (CDK) genes, *CDK4* and *CDK6*, have been shown to be amplified in lung cancer. Inhibitors of these CDKs, *CDKN2A*/*CDKN2B*, have been found to be deleted in 3% of lung adenocarcinomas, indicating disrupted cell cycle regulation as an important mechanism in carcinogenesis [90]. The *RB1* gene, which

as previously mentioned is frequently mutated or methylated in lung cancer, has also been shown to be deleted in lung cancer [90].

A few miRNAs are up-regulated in tumor cells, functioning as oncomirs. For example, breast cancer cells have increased expression of the miR-103/107 family (*MIR103A1, MIR103A2, MIR103B1, MIR103B1, MIR107*) [91]. These two microRNAs target and down-regulate Dicer, a protein involved in microRNA maturation. Silencing of *MIR-103/107* inhibits metastasis, suggesting that a global suppression of microRNA maturation is necessary for the tumor cells to gain metastatic potential [91]. Down-regulation of Dicer predicts poor prognosis in lung cancer [92]. The microRNA miR-21 (*MIR21*) is amplified in ADCAs and is related to poor survival and metastasis of lung cancer [38, 93, 94], breast cancer [95], prostate cancer [96], colorectal cancer [97], and melanoma [98]. Serum from smokers contains higher levels of miR-21 than serum from non-smokers. Moreover, serum from heavy smokers contains higher levels of miR-21 than serum from light smokers [99]. Human bronchial epithelial cells transformed by exposure to cigarette smoke extract *in vitro* expressed higher levels of miR-21 than non-transformed cells and excreted more miR-21 to the cell culture medium [99]. Transformed HBE cells also showed lower levels of the epithelial microRNA miR-200c (*MIR200C*) [100].

1.2.5 Carcinogen Metabolism

As mentioned in section 1.1.2, two significant contributors to lung carcinogenesis are the PAH and the tobacco-specific NNKs, leading to mutations through adduct formation [101]. PAHs are pro-carcinogens formed during incomplete combustion of organic material and NNKs are produced naturally in tobacco leaves. Both NNKs and PAHs are pro-carcinogens that are activated by CYPs to their mutagenic metabolite [101-103].

In phase I metabolism, PAHs are metabolized primarily by CYP1A1 and CYP1B1 but also other enzymes including epoxide hydrolases (EH) and aldo-keto reductases (AKR) (Fig. 2) [104]. B[*a*]P is metabolized into B[*a*]P-epoxides, -diols and diolepoxides in a multistep pathway [105]. B[*a*]P in itself is not able to induce DNA

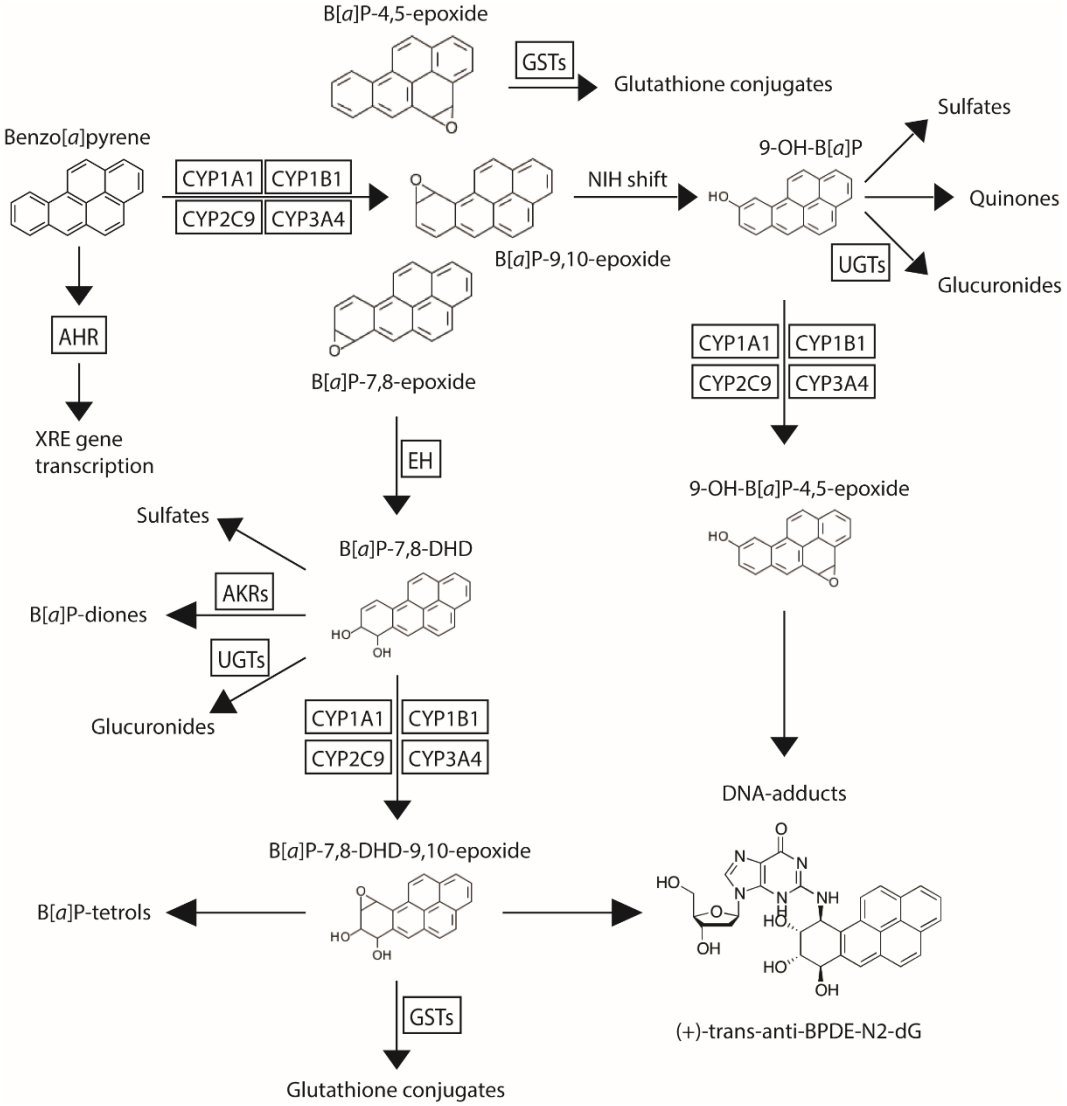

Figure 2: Benzo[a]pyrene metabolism. Benzo[*a*]pyrene is metabolized by CYP enzymes to B[*a*]P-epoxides. The epoxides are further converted to phenols by an NIH-shift, or to dihydrodiols by epoxide hydrolase (EH). The CYP enzymes again convert the dihydrodiols and phenols to diolepoxides and phenolepoxides, respectively, which can create bulky DNA-adducts. Phase II metabolizing enzymes conjugates the B[*a*]P metabolites with glutathione and glucuronic acid. GST, glutathione-S-transferase; UGT, UDP-glucuronosyltransferase; AKR, aldo-keto reductase.

damage, but the carcinogenic metabolite BPDE can form bulky DNA adducts by binding DNA at the N2 position of deoxyguanosine (BPDE-N2-dG, Fig. 2). [106].

B[*a*]P and other PAHs are lipophilic compounds that can cross the cell membrane through passive diffusion. In the cytosol, B[*a*]P binds to the ligand-activated aryl hydrocarbon receptor (AHR), which then relocates to the nucleus where it dimerizes with ARNT. The heterodimer can recognize xenobiotic response elements (XRE) in

the DNA and induce gene expression of targeted genes including *CYP1A1* and *CYP1B1* [107, 108]. In addition to the metabolism of exogenous compounds such as B[*a*]P, the CYP enzymes are involved in the metabolism of endogenous steroids including 17-β-estradiol (E2) [101, 109].

Growing evidence support a role of steroid receptors in lung cancer development. Steroid receptor may establish sex differences in molecular mechanisms such as carcinogen metabolism. Both lung tissue and lung cell lines from women have higher expression levels of *CYP1A1* than male lung tissue and lung cells [110, 111]. This increase in *CYP1A1* gene expression was associated with increased adduct formation in women [112]. Interestingly, there is evidence of estrogen receptor alpha being important in *CYP1B1* regulation through direct binding to estrogen responsive elements (ERE) in the *CYP1B1* enhancer in breast cancer cells [113]. However, studies have not found any gender differences in expression of *CYP1B1* [111, 112].

A cross-talk between metabolism of xenobiotics and steroids have been demonstrated. CYP1A1 and CYP1B1 hydroxylate E2 at the 2- and 4-position, respectively, creating 2-OH-E2 and 4-OH-E2. The 4-OH-E2 metabolite is genotoxic and able to create DNA-adducts [114]. The catechol-O-methyl transferase (COMT) further detoxifies these hydroxyl-metabolites to 2-methoxy-E2 (2-MeO-E2) and 4-methoxy-E2 (4-MeO-E2) [109, 115]. Co-exposure of B[*a*]P and E2 has previously been shown to increase the levels of the hydroxyl metabolites while co-exposure of E2 with 2,3,7,8-Tetrachlorodibenzo-p-dioxin (dioxin, TCDD) leads to increased levels of the estradiol methoxides [116-118].

1.2.6 Lung cancer susceptibility

The risk of developing lung cancer varies among individuals. An individual's susceptibility to developing cancer is affected by the genetic composition. If cancer occurs as an accumulation of somatic mutations, any pre-existing germ-line mutation that can drive cancer progression can increase an individual's susceptibility. Single nucleotide polymorphisms (SNPs) are mutations that are present in more than 1% of the population. SNPs in any cellular process associated with the cancer hallmarks increase the possibility to develop cancer. A pre-existing mutation may not confer

any cancer risk in itself, but reduces the number of additional mutations required for malignant transformation. An individual's ability to metabolize and detoxify carcinogens affects the susceptibility towards lung cancer and other cancer types associated with carcinogen exposure. Hence, individual lung cancer susceptibility is a product of genetic and epigenetic predisposition, the innate capacity to detoxify xenobiotics, capacity for metabolic activation of PAHs and other pro-carcinogens, and DNA adduct formation and efficiency of DNA repair systems.

Persons from families with a history of lung cancer have an increased risk of developing lung cancer, and the disease can occur at a younger age in these families than in the general population [119, 120]. The link between family history and risk of lung cancer imply the presence of a genetic factor in disease susceptibility. A risk factor for developing lung cancer is an addiction to nicotine [121]. The genes *CHRNA2*, *CHRNA5*, and *CHRNB4* is situated in the 15q25 chromosome region. Two SNPs in this region, rs1051730 and rs8034191, are associated with increased susceptibility to lung cancer [122, 123]. These three genes encode nicotinic acetylcholine receptors (nAChRs) that confer nicotine sensitivity. SNPs in nAChRs might alter receptor functions and increase addiction to tobacco. Increased nicotine addiction can result in increased exposure to tobacco carcinogens. The SNP rs2736100 at 5p15 is associated with increased expression of *TERT* and increased lung cancer susceptibility [124-126]. SNPs can also associate with resistance to treatment. The T790M mutation in *EGFR* is associated with TKI drug–resistance and is found in patients with familial history of NSCLC [127].

Homozygous gene deletion (null genotype) of the phase II metabolizing enzyme glutathione S-transferase M1 (*GSTM1*) is thought to infer an increased risk of developing lung cancer and chronic obstructive pulmonary disease (COPD) [128-130]. The *GSTM1* null genotype leads to decreased ability to detoxify PAHs, and patients with this genotype have increased levels PAH-derived DNA-adducts [131]. Another SNP found to be associated with an increased risk of developing lung cancer is the I157T variant of *CHEK2*, a key cell-cycle control gene that activates cell cycle checkpoints in response to DNA damage. CHEK2 plays a main part in maintaining genetic integrity [128]. The I157T variant is present at a relatively high frequency in

populations from northern and central Europe (5-7%). Dysfunctional CHEK2 is thought to hinder the capacity of cells with damaged DNA to undergo cell-cycle arrest and DNA repair at defined checkpoints of cells containing heavily damaged DNA.

There is an increasing acknowledgment that the biology of the sexes is different also when it comes to the risk of cancer development. Over the past few years, there has been an increased focus on uncovering sexual differences in disease development. Studies have indicated that both women who smoke and female never-smokers are at higher risk of developing the disease than smoking and never-smoking men, respectively. The data are, however, conflicting, and recent studies have shown that men and women have a similar risk of developing lung cancer [6, 12, 132-134]. There may not be a sexual dimorphism in lung cancer susceptibility, but a sex difference in lung cancer biology possibly play a role in lung cancer development, diagnosis, and treatment selection.

Both lung tumors and normal human lung have been shown to express the estrogen receptors ERα (*ESR1*) and ERβ (*ESR2*) [135-138]. ERs seem to play a biological role in lung and is associated with induced proliferation, induced migration, activated transcription and estrogen-stimulated secretion of growth factors [136] [138, 139]. *ESR1* is significantly higher expressed in malignant lung epithelial cells compared to non-malignant cells [140]. The expression levels of *ESR2* is often up-regulated in tumor tissue compared to normal tissue from the same patient, and expression of *ESR2* in metastatic lung cancer is associated with poor prognosis in both male and female patients [141, 142]. Female ADCAs are enriched in *EGFR* mutations, while male ADCAs are enriched with mutations in the tumor suppressor *SMARCA4* [38, 79].

1.3 Epithelial to mesenchymal transition

Epithelial to mesenchymal transition (EMT) is a cellular reprogramming of the epithelial cell that may be important in cell invasion, metastasis and resistance to different types of treatment in lung cancers. Epithelial cells are held together by various kinds of junctions, forming a tightly connected sheet that lines the exterior

and interior surfaces of the body. Epithelial cells are anchored to a base membrane (basal lamina) by integrins and other transmembrane proteins. Cell-cell anchoring junctions are formed by transmembrane proteins of the cadherin superfamily. Extracellularly, cadherins bind each other in a homophilic manner, while they in cytosol bind actin and intermediate filaments through beta-catenin and other catenins and anchor proteins. The cell-cell junctions provide the cells with mechanical strength, actin-based motility and cell-cell signaling [143]. Carcinomas originate in the epithelial tissue and are the most common cancer types. Loss of cell polarity, loss of anchorage-dependent growth and loss of cell-cell junctions characterizes carcinomas [144].

1.3.1 Transcriptional reprogramming during EMT

Epithelial cells possess some level of plasticity and can disassemble and migrate away from the epithelium. EMT normally occurs during development and tissue repair. Branching morphogenesis of the lung requires explicit signaling from both epithelial and mesenchymal compartments to regulate cell proliferation, migration, and subsequent lung-specific gene expression [145]. During EMT, the epithelial cells lose their cell-cell junctions and apical-basal polarity and gain an elongated, fibroblast-like morphology. One of the main event during EMT is a loss of the epithelial E-cadherin (*CDH1*) and gain of the mesenchymal N-cadherin (*CDH2*), an event called the cadherin switch. The mesenchymal cells have an increased ability to migrate and invade [146]. One important feature of EMT is its transient nature and the possibility of mesenchymal cells to revert to epithelial cells through mesenchymal-to-epithelial transition (MET). During development, several rounds of EMT/MET occurs to give rise to the different organs [147].

Regulation of EMT is complex and multilayered and involves several different transcription factors, DNA-methylation, histone modifications, microRNAs and post-translational modifications (Fig. 3) [148, 149]. Much research on EMT has focused on the EMT master transcription factors. The two zinc finger transcription families Snail (*SNAI1*, *SNAI2*) and ZEB (*ZEB1*, *ZEB2*), and the basic helix-loop-helix family member TWIST are transcriptionally activated early during EMT [143]. All of these genes can bind E-boxes in the DNA and act as both repressors and

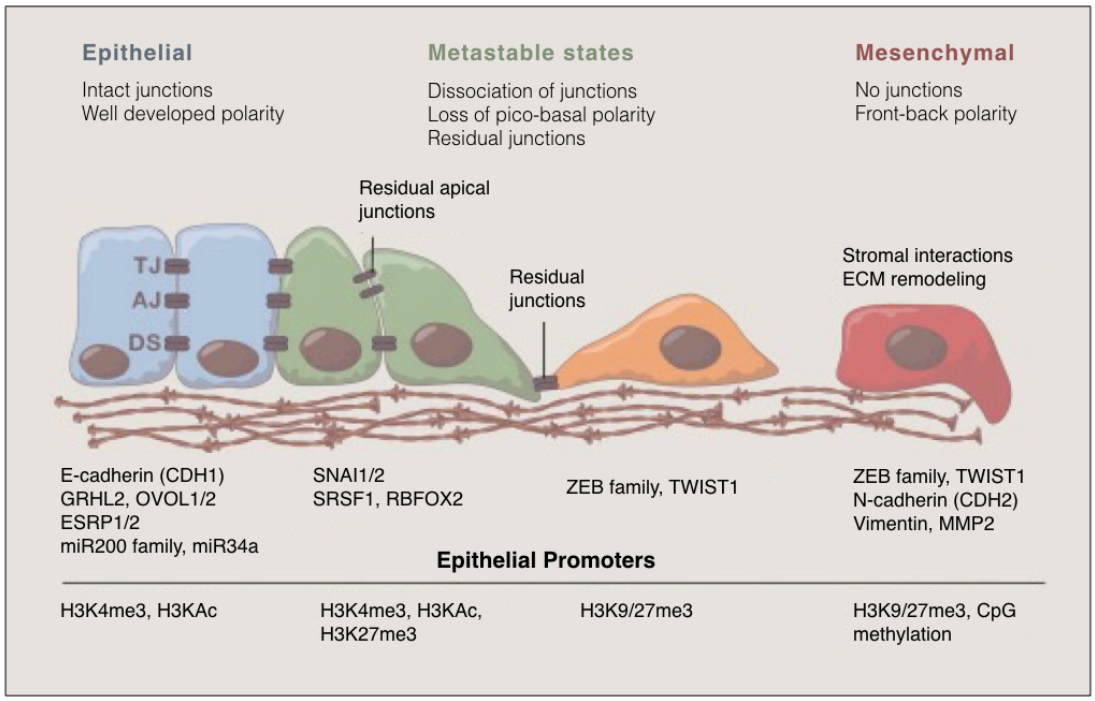

Figure 3: Epithelial to mesenchymal transition. Epithelial cells are polar cells that form a rigid tissue layer interconnecting through junctions (Tight Junction, Adherence Junctions, Desmosomes). During EMT epithelial cells lose expression of epithelial genes such as E-cadherin (CDH1). A range of transcription factors including the SNAI and ZEB families are involved in activating EMT. EMT also involves epigenetic reprogramming of the histone code, altered DNA methylation and loss of expression of epithelial-specific microRNAs. The transition from epithelial to mesenchymal state occurs over a gradient, and there may be several metastable states where the cells possess partially mesenchymal and epithelial traits. Mesenchymal cells do not have any junctions, are more motile, and can invade surrounding tissues. Figure modified from Nieto et al. 2016 [148]. Printed with permission from Elsevier.

activators of gene transcription [143]. TWIST actively induces transcription of *CDH2* by binding to an E-box element in the first intron of *CDH2*, making it an essential transcription factor (TF) in establishing the cadherin switch [150].

During EMT, the gene expression landscape is modulated through an epigenetic reprogramming of the histone landscape with a global reduction in the heterochromatin mark H3K9Me2 and an increase in the euchromatin mark H3K4Me3 and the transcription mark H3K36Me3 [151]. The EMT driver SNAI1 recruit chromatin modifiers including SIN3A, HDAC1, HDAC2, LSD-1, PRC2, and the G9a and Suv39H1 histone methyltransferases. SNAI1 binds transiently to its target promoters, triggering both transient and long-lasting chromatin changes [148, 152]. ZEB1 represses its target genes by recruiting the LSD1-containing co-repressor complex, as well as HDAC1 and HDAC2 [148, 153, 154]). Although

alterations in *CDH1* expression can occur through multiple mechanisms, including loss of heterozygosity and mutational inactivation, *CDH1* is frequently silenced through DNA-methylation of its promoter which has been observed in several types of cancer. Hypermethylation of the *CDH1* promoter is a potential prognostic factor for NSCLC [55, 155, 156].

Several microRNAs have also been implicated in EMT, some of them directly regulated by TP53. In the absence of wild-type TP53 function, SNAI1-dependent EMT is activated in the colon, breast and lung carcinoma cells as a consequence of a decrease in miR-34 (*MIR34*) levels. miR-34 suppresses SNAI1 activity by binding to highly conserved 3'UTR in *SNAI1* mRNA [157]. The two microRNAs miR-21 and miR-31 enhances motility and invasiveness of colon carcinoma cell lines in TGF-b induced EMT [158]. Members of the microRNA miR-200 family (*MIR141*, *MIR200A*, *MIR200B*, *MIR200C* and *MIR429*) are important in maintaining an epithelial phenotype [159]. For example, miR-200c regulates EMT by repressing *ZEB1* and *ZEB2*, thus ensuring normal expression of *CDH1* [160]. The miR-200 family is directly activated by TP53, which serves as a guardian of the genome and guardian of the epithelial phenotype [148, 160]. TP53 transactivates miR-200c through direct binding to the miR-200c promoter. Loss of TP53 in mammary epithelial cells leads to decreased expression of miR-200c and activates the EMT program [160].

1.3.2 EMT in lung cancer

Tumor cells are not purely "epithelial" or "mesenchymal", but transition through a spectrum of intermediate phases, and cells carrying characteristics of both epithelial and mesenchymal phenotype has been observed (Fig. 3) [149]. These intermediate states are referred to as "metastable" and cells in these states can induce or reverse the EMT process [161]. In a metastable state, the cell can have lost E-cadherin without induction of N-cadherin. Cells that are in different metastable states might have different phenotypes, such as distinct migration and invasion abilities due to differences in adhesion behavior [148]. EMT cells can also reach an epigenetically fixed state that is reversible only through chromatin remodeling [161].

EMT is readily observed *in vitro* after lung cancer cells are treated with the EMT-inducing TGF-β, and most research on EMT has been done on cell lines [162]. Whether EMT occurs in cancer *in vivo* has been controversial, but EMT has been reported in the invading front of colorectal tumors [163, 164]. In NSCLC, DNA-methylation patterns separate tumors and cancer cell lines into epithelial-like and mesenchymal-like, suggesting that cells with different degrees of EMT could be present in tumors *in vivo* [55]. Due to its role in enabling cells to migrate and invade, EMT has been suggested as an important mechanism during metastasis. Recent studies suggest that EMT may be important in initiating invasion and intravasation, but may be dispensable for establishing the metastatic tumor at secondary sites [165-167]. In mouse models, motile tumor cells had undergone EMT while the epithelial tumor cells that expressed high levels of E-cadherin were not migratory. Mesenchymal cells that arrived at the secondary site adapted an epithelial state (MET) after a few cell divisions [167].

An interesting question is whether EMT occurs as a consequence of random alterations during carcinogenesis, or if cellular signaling deliberately activates it. Intriguingly, the AHR has been found to regulate more than just transcriptional activation of xenobiotic metabolizing genes. AHR is found to govern a vast range of cellular functions including cell cycle progression, cell adhesion, cell morphology, and cell migration [168-170]. AHR directly binds to an XRE in the *SNAI2* gene and induces its transcription. Knock-down of AHR in HaCaT human keratinocyte cells resulted in silencing of *SNAI2* transcription which inhibited EMT in keratinocytes and palate cell lines [171]. Another study found opposite effect of Ahr on *Snai2* in mouse keratinocytes [172]. Ahr repressed *Snai2* expression and inhibited EMT. Even if the findings in human and mice cell lines differ they indicate a possible cross-talk between PAH exposure and EMT, and this could play an important role in lung carcinogenesis.

When selecting treatment options for lung cancer patients, research suggests that EMT status of the tumor should be taken into account. As previously mentioned, EGFR TKIs have become important in treating lung cancer patients with EGFR mutations, but patients often acquire resistance to these inhibitors. Cancer cell lines

undergoing EMT have intrinsic resistance to EGFR TKIs, and EMT has been negatively associated with EGFR TKI sensitivity in NSCLC [173]. *ZEB1* is upregulated in NSCLC cell lines with resistance to EGFR tyrosine kinase inhibitors while miR-200c expression was markedly reduced in NSCLC cell lines [174]. In mammary cells *in vivo*, overexpression of miR-200c, which targets ZEB1, increased the susceptibility for chemotherapy, showing that blocking EMT may be an approach for treating chemoresistant cancers [165].

An exciting new treatment option for lung cancers is immunotherapy. This type of therapy is, however, also troubled by the development of treatment-resistant cells. Resistance is often conferred by the tumor cells expression of the PD-L1 (CD274) surface protein which signals "self" to the PD-1 (*PDCD1*) receptor on the T-cells. miR-200 is found to repress *PD-L1* in NSCLC. When ZEB1 represses miR-200 during EMT, PD-L1 protein is generated at a higher level, thus providing the tumor cell with a mechanism for evading immune destruction [175].

1.4 Forkhead Box Transcription Factors

Increasing evidence supports a role of steroid receptors and steroid signaling pathways in lung cancer. Steroid receptors act as ligand-activated transcription factors that are stimulated by steroid hormones.

1.4.1 Pioneer transcription factors

Transcription factors regulate gene activity by binding to regulatory elements throughout the genome. These proteins can roughly be divided into three main groups; lineage- determining transcription factors (LDTFs), collaborating transcription factors (CTFs), and signal-dependent transcription factors (SDTFs). Common for all these groups of TFs are their ability to recognize short DNA sequences, called transcription factor binding sites (TFBSs), and initiate transcription upon binding. LDTFs poises enhancers for binding of other TFs by opening up the chromatin structure. The forkhead box A (FOXA) family belongs to a group of TFs called pioneer factors due to their ability to bind TFBS that resides in condensed, inactive chromatin [176, 177]. Binding of FOXA factors initiates chromatin restructuring resulting in the availability of TFBSs of other transcription factors,

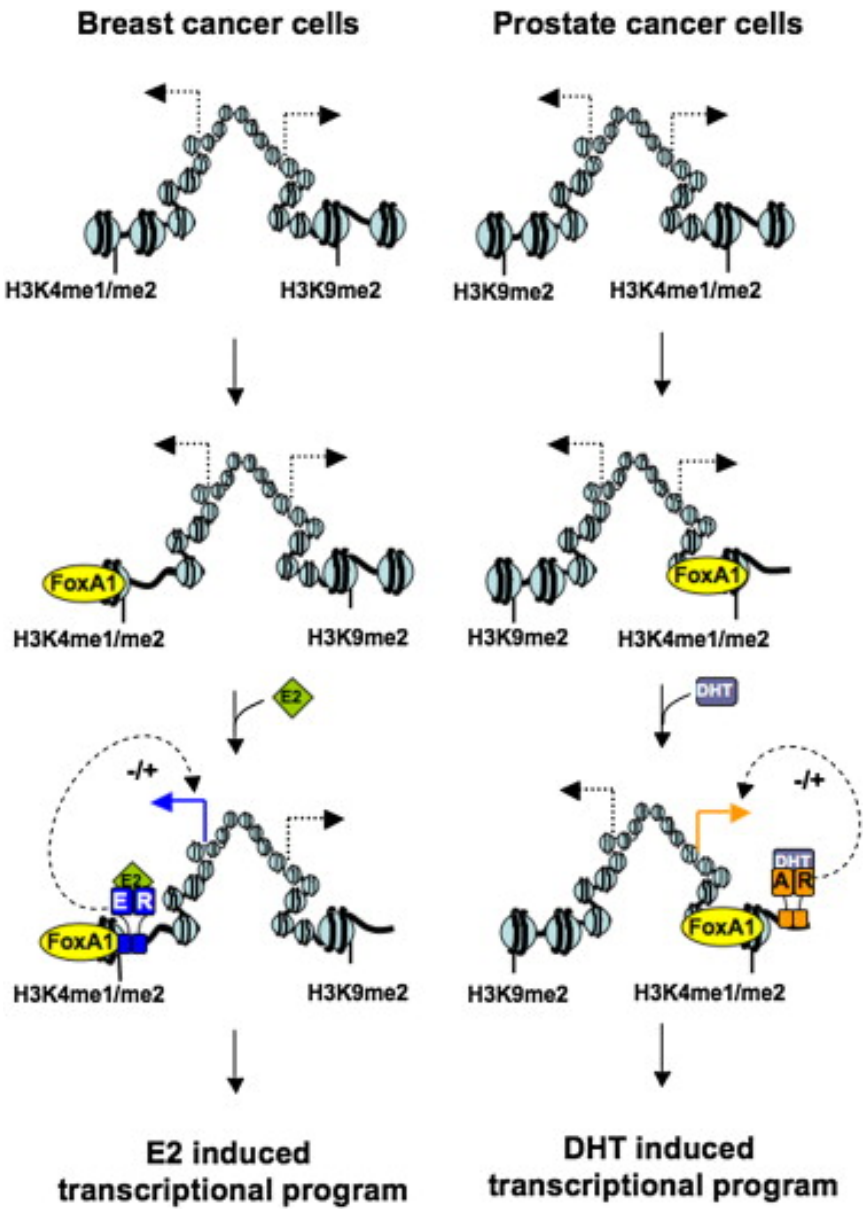

Figure 4: FOXA1 binding to chromatin. FOXA1 can identify and bind TFBSs in condensed chromatin. FOXA1 binding occurs in a sequence specific and lineage-dependent way predefined by the histone code. FOXA1 binding enables binding of signal-dependent transcription factors, such as ER or AR. TFBS binding activates transcription through long-range chromatin interactions. The figure is adapted from Lupien et al., 2008 [66]. Printed with permission from Elsevier.

including estrogen receptors and androgen receptor (AR) [178, 179]. In liver, FOXA1 and FOXA2 binding occur mainly in enhancers distal to gene transcription start sites (TSS) [180].

FOXAs contain a winged helix domain, a structure also found in linker histones H1 and H5 [181, 182]. In contrast to most transcription factors, which are not able to bind condensed chromatin, the winged-helix domain makes FOXAs able to

recognize and bind partially available binding motifs on nucleosome cores [176]. The C-terminal domain of FOXAs recognizes nucleosome core structures by binding histone H3 and H4 complexes [177]. The winged-helix and the C-terminal domains cooperate in the binding of nucleosome DNA. FOXAs displace linker histone H1 on the nucleosome core particle. The displacement leads to the release of compacted linker DNA from the nucleosome and decompaction of chromatin [176]. FOXA1 binds to its TFBS in a lineage-specific way, guided by H3K4me1/2 [66]. FOXA1 is, however, unable to bind forkhead motifs marked with H3K9me2 (Fig. 4). Thus the cell type-specific epigenetic landscape dictates active and repressed FOXA binding sites.

The open chromatin configuration in FOXA occupied enhancers makes it possible for other transcription factors to bind their respective binding sites. In the prostate, FOXA1 and FOXA2 are essential in the binding of the androgen receptor (AR) to its response element [183, 184]. In mammary cells, FOXA1 is essential for binding of ERα to DNA and is thus necessary for estrogen-mediated gene regulation [185]. FOXA1 and ERα are involved in long-range gene regulation and chromatin looping of distal regulatory sites (enhancers) to gene proximal sites (promoters) [179, 186].

1.4.2 FOXAs in lung cancer

There are three known FOXA factors, FOXA1, FOXA2, and FOXA3. The lung only expresses *FOXA1* and *FOXA2* while the liver, stomach, and pancreas expresses *FOXA3* [187]. *FOXA1* and *FOXA2* are critical for normal branching of the lung during morphogenesis and is also expressed in adult lung tissue [145, 188]. Previous studies have suggested a possible increase of *FOXA1* in lung adenocarcinoma [189], lung squamous cell carcinoma [190], pancreatic cancer [191], prostate cancer [192] and thyroid cancer [193]. These findings indicate that *FOXA1* may be oncogenic in these cancer types. Contradictory to this, *FOXA1* may act as a tumor suppressor in the breast and expression of *FOXA1* is associated with good prognosis in breast cancer [194, 195]. *FOXA2*, on the other hand, may be suppressed in lung cancer cell lines [196, 197] and gastric cancer [198], and is involved in suppressing metastasis in lung adenocarcinomas [199]. In breast cancer, expression of *FOXA2* is associated with poor prognosis [200]. The function of FOXA1 and FOXA2 thus appear to be tissue

dependent. *FOXA1* is potentially overexpressed due to gene amplification in esophageal and lung ADCA, while *FOXA2* is found to be down-regulated by DNA methylation in lung cancer cell lines [189, 197].

FOXA1 and FOXA2 share homologous protein sequence and a degree of overlapping function has been found [145]. In liver, FOXA1 and FOXA2 have been found to have many overlapping binding sites, but the two transcription factors also bind and activate different sets of genes. FOXA1 binding sites are enriched for TP53 binding, where FOXA1 and TP53 seem to have opposite effect on transcription of the target gene [201, 202]. FOXA1 is involved in regulating expression of genes involved in cell cycle regulation, while FOXA2 binding occurs in genes involved in steroid and lipid metabolism [201].

In lung cells FOXA2 regulate the xenobiotic metabolizing genes *CYP2A13* and *ABCB1* through directly binding to enhancer elements in these genes, suggesting a role for FOXA in regulating the cell's response to environmental chemicals [203]. FOXA2, together with Nkx2-1, which is significantly amplified in ADCAs, and Cdx2 has previously been reported to cooperate in inhibition of a metastatic program in lung adenocarcinoma [38].

1.4.3 FOXAs in EMT

FOXA1 and FOXA2 are considered to play an essential role in maintaining the epithelial phenotype, and loss of *FOXA1* and *FOXA2* expression has been reported during EMT in some tissues, including lung [161, 204]. FOXA binds forkhead binding sites ~800 bp upstream of the *CDH1* transcription start site, and this binding coincides with transcriptional activation of *CDH1* [205]. Interestingly, increased expression of FOXA has been reported to reduce the migratory capacity of tumor cells regardless of *CDH1* expression status [205]. This uncoupling suggests that FOXAs are involved in regulating EMT not only by inducing *CDH1* but possibly also by regulating other genes. Down-regulation of *FOXA2* is observed in cells entering EMT [196, 206]. Both FOXA1 and FOXA2 has been found to repress expression of *SNAI2* directly by binding its promoter [196, 207]. FOXA1 recruits

HDAC2 to the SNAI2 promoter, resulting in deacetylation of histone H3K9, initiating the repressive mark H3K9me3 [207].

In breast cancer, exposure to the endocrine disruptor bisphenol A (BPA) results in downregulation of *FOXA1* which in turn leads to EMT [208]. *FOXA2*, although associated with poor prognosis in breast cancer, has been found to prevent EMT in breast cancer cells were *FOXA2* and epithelial phenotype strongly correlated [206]. Zhang and colleagues found that FOXA2 directly stimulated the promoter of *CDH1* and repressed the expression of EMT-related transcription factor *ZEB2* by recruiting the repressor TLE3 to the *ZEB2* promoter. In addition to maintaining the epithelial phenotype, the overexpression of FOXA2 abolished metastasis of breast cancer *in vivo* [206]. Similar findings of FOXA2 maintaining an epithelial phenotype has been done in lung cancer cell lines. High expression levels of *FOXA2* were associated with elevated levels of *CDH1* and low levels of vimentin (*VIM)*, whereas low expression of *FOXA2* is linked to low levels of *CDH1* and high *VIM* levels. *FOXA2* knockdown induces EMT and promotes invasion in lung cancer cells. FOXA2 binds to the promoter of *SNAI2* suppress its transcription [196].

FOXA1 and FOXA2 are important regulators of steroid signaling and are also involved in regulating the transcriptional reprogramming of EMT, establishing a possible link between ER activity and EMT. Moreover, EMT is often observed during cell transformation *in vitro*, as well as in treatment resistance [165, 166]. Considering the close connection between FOXA1 and FOXA2 and EMT, it is thus plausible that abnormal expression of these two transcription factors is important in cancer progression.

1.5 Lung cell transformation *in vitro*

Cellular transformation is a stepwise process, and several genetic, and epigenetic changes in oncogenes and tumor suppressor genes are required for malignant transformation to occur. For decades, cancer research has used *in vitro* experimental models to study the biology of cancer and carcinogenesis.

Human primary cells have limited ability to proliferate *in vitro*. After a few population doublings, they will enter replicative senescence, or they enter cellular crisis with

slowed cellular growth and apoptosis [209, 210]. Tumor cells, on the other hand, are capable of escaping senescence and cellular crisis and can be cultured as cell lines for extended periods of time. The most cited lung epithelial cell line, A549, is derived from bronchial alveolar carcinoma [211]. Despite its origin in malignant tissue, it has been widely used to study lung biology. A concern regarding tumor cell lines is that their biology is altered from that of normal cells and that they contain many undefined molecular changes [212].

Efforts have been made to establish cell lines that are not derived from tumors. In the literature, immortalized cells are often distinguished from transformed cells. Immortalized cells have the ability to divide indefinitely while they do not grow in soft-agar or in xenografts in immunodeficient (nude) mice. Transformed cells are immortal but also, they have the ability to grow in soft-agar and/or as tumors xenografts in nude mice. Immortalization of cells by infection with SV40 virus has been a common approach [213, 214]. Infected cells are capable of avoiding senescence, but most cells enter crisis after a few rounds of subculturing. A few clones, however, can be resistant to the crisis and will continue to grow and can be cultured more or less indefinitely. SV40 make cells able to surpass the senescence threshold by binding TP53 and RB proteins. Inhibition of TP53 substantially extends the lifespan of several cell types *in vitro* [215]. TP53 induces transcription of the CDK inhibitor p21CIP1, which inhibits cell cycle progression. Thus activation of TP53 initiates senescence in cultured cells [216].

Lundberg and colleagues reported in 2002 the generation of an immortalized human airway epithelial cell line by making the cells express both the large T oncoprotein of the SV40 (LT) virus and *TERT* [217]. These cells were able to avoid both senescence and crisis, showing that the shortening of telomeres is important during the crisis and that this is avoided by maintaining the telomeres through *TERT* expression. The concern of using viral oncogenes for the immortalization of cells it that there might be unknown, uncontrollable effects that contribute to the transformation process.

Other approaches in immortalization of bronchial epithelial cells without viral oncoprotein is for instance by overexpressing *TERT*, *CDK4* and a dominant negative form of *TP53* [218]. The expression of *CDK4* avoids a p16INK4a –mediated stress

response that occurs in cultured epithelial cells in defined medium. The nonfunctional form of TP53 mimics the effect of the SV40 oncoprotein. Another group reported the immortalization of human bronchial epithelial cells (HBECs) using vectors expressing *CDK4* and *TERT* without creating a non-functional TP53 [219]. None of the cells immortalized in this manner generates colonies in soft-agar or tumors in nude mice. They retain an intact TP53 signaling pathway and the gene expression patterns of immortalized HBECs clusters with normal non-immortalized bronchial epithelial cells and not with lung tumors [219]. This makes HBECs well suited for studying molecular pathways in normal lung cells. Cells that are immortalized can further be transformed to soft-agar growing cells. This can be achieved by making the cells express medium levels of the oncogenes *KRAS* and *MYC* [218]. Human small airway epithelial cells immortalized by *TERT*, *CDK4* and a double negative form of TP53 were able to form colonies in soft agar and tumors in nude mice only if they also expressed both *KRAS* and *MYC* [218].

Malignant transformation of lung cells has been shown *in vitro* for different types of carcinogenic compounds including BPDE, B[*a*]P, hexavalent chromium, MNU and cigarette smoke condensate (CSC) [220-222]. Transformation of immortalized human bronchial epithelial cells by BPDE or MNU exposure resulted in increased protein levels of DNA-methyl transferase 1 (*DNMT1*) [220]. The increased DNMT1 levels led to hypermethylation and suppressed gene expression of *CDH1* and *FOXA2* [220]. Exposure of bronchial epithelial cells to B[*a*]P has been associated with transformation involving a higher expression of *TP53* in transformed versus non-transformed cells [221].

The first study showing that epithelial cells suspended in a collagen gel obtained a mesenchymal-like phenotype was published in 1982 [223]. Since then, several studies have shown that cell transformation *in vitro*, involves EMT. Nicotine is usually not considered a carcinogen but is a considerable risk factor in lung cancer due to its addictiveness. In cancer cell lines, nicotine has been shown to induce EMT, evident by reduction of *CDH1* and induction of *VIM* and fibronectin (*FN1*) [224]. The nicotine-treated cells had disrupted cell-cell contacts and increased invasiveness mediated by nAChRs, according to the EMT phenotype. Their findings show that

nicotine could aid the transformation of lung cells and induce invasive and migratory phenotype. EMT has also been observed in lung cells after exposure to other environmental carcinogens, such as ultrafine particulate matter [225], multi-walled carbon nanotubes [226], or as a spontaneous event without carcinogenic exposure in prostate cell lines [227]. Human bronchial epithelial cells exposed to CSC had repressed expression of miR-200c which leads to EMT and cell transformation [100]. Similarities between the different approaches to cell transformation *in vitro* is the observation that transformed cells often have increased motility, gain the ability to grow in soft agar or nude mice xenografts, loss of expression of *CDH1*, and increased expression of *CDH2* and *VIM* [227, 228]. Together, these studies indicate that EMT may be a major mechanism during cell transformation.

2 AIMS OF STUDY

The main goal was to explore the role of FOXA1 and FOXA2 in lung carcinogenesis with focus on cell transformation, EMT, carcinogen metabolism, and clinical implications *in vivo*. Specifically, the goals were to:

- Investigate the role of FOXA1, FOXA2, and EMT during carcinogen-induced lung cell transformation.
- Investigate the expression and gene regulation of FOXA1 and FOXA2 in lung cancer patients and identify possible clinical implications.
- Characterize the role of FOXA1 in the regulation of CYP1A1 and CYP1B1 and the implications on carcinogen metabolism.

3 SUMMARY OF PAPERS

Paper I: Epithelial-mesenchymal transition and FOXA genes during tobacco smoke carcinogen-induced transformation of human bronchial epithelial cells.

An *in vitro* cell transformation study was performed on two different immortalized human bronchial epithelial cell lines (HBEC2, HBEC12). Cells were exposed to B[*a*]P, cigarette smoke condensate (CSC) or methylnitrosourea (MNU) for 12 weeks. Cells that were able to grow in soft-agar after carcinogen exposure was defined as transformed. Gene expression was analyzed by RT-qPCR and protein levels were analyzed by immunofluorescent microscopy. Cell migration was analyzed by a wound healing assay and invasiveness was analyzed by a Matrigel wound healing assay. Time-lapse images were acquired every hour for 48 and 72 hours for migration and invasion, respectively. Chromatin immunoprecipitation (ChIP) was used for identification of histone variants in regulatory regions of *FOXA1* and *FOXA2*. For HBEC2, exposure to CSC leads to the formation of a subpopulation of cells that were able to grow anchorage-independent in soft-agar. For HBEC12, exposure to B[*a*]P, CSC, and MNU all lead to colony formation in soft-agar. Cell colonies were isolated and cultured in monolayer. These transformed cells displayed altered, elongated morphology, increased invasiveness and gene expression patterns evident of EMT. Specifically, *CDH1* (E-cadherin) was down-regulated in all transformed cells. *CDH2* (N-cadherin) was up-regulated in both HBEC2 and one of the HBEC12 transformed cell lines. The EMT-markers *SNAI1*, *ZEB1*, *VIM*, and *MMP2* were up-regulated, and *SNAI2* was down-regulated in HBEC2 transformed cell lines. The expression levels of these EMT markers were more inconsistent in the HBEC12 transformed cell lines. Apart from upregulation of *ZEB1* in B[*a*]P and MNU transformed cells and reduced expression of *SNAI2* in all but one of the B[*a*]P transformed cell lines, the gene expression of the EMT markers was unchanged from the non-transformed HBEC12. The switch in cadherin expression from *CDH1* to *CDH2* appeared in HBEC2 prior to full transformation and was observed after six weeks of exposure. The expression of the two transcription factors *FOXA1* and *FOXA2* was down-regulated in all transformed cell lines. This down-regulation occurred before full transformation and was present after six weeks of carcinogen

exposure. Reduction of *FOXA1* and *FOXA2* was also observed by immunofluorescent microscopy. Increased binding of histones H3, macroH2A, and H2A.Z was seen in transformed cell lines compared to non-transformed cells for all assays analyzed, except those assays spanning TSS of both the *FOXA1* and *FOXA2* genes. Together the results show that tobacco smoke carcinogens can transform human bronchial epithelial cells *in vitro*. Cell transformation may involve FOXA1 and FOXA2, and EMT appears to play a major role in the transformation.

Paper II: Altered expression of FOXA1 and FOXA2 in non-small cell lung cancer: genetic and epigenetic mechanisms, and diagnostic implications

A case-control study including three geographically independent lung cancer cohorts was performed. *FOXA1* and *FOXA2* expression was measured by RT-qPCR in matched pairs of adenocarcinomas (ADCA) in cohorts from Norway and Italy (n=155), squamous cell carcinomas (SCC) from Norway and USA (n=94), and large cell carcinomas (LCC) from Norway (n=19). Gene expression of FOXA1 and FOXA2 was analyzed by RT-qPCR on matched pairs of tumor and non-tumor lung tissue from the three cohorts. Gene amplification of *FOXA1* was studied in the Norwegian cohort by qPCR. DNA methylation levels of *FOXA2* was analyzed by bisulfite pyrosequencing in the Norwegian cohort and by Illumina Infinium HumanMethylation27 BeadChip. Additionally, two public datasets, TCGA and CURELUNG, containing DNA methylation data from Illumina Infinium HumanMethylation450 BeadChip were mined. The gene expression of *FOXA1* was increased while that of *FOXA2* was decreased in tumor compared to non-tumor tissue. In non-tumor lung tissue the expression levels of FOXA1 and FOXA2 correlated. Lung tumors appeared to express low levels of both factors or high levels of one factor, but never both. While the expression levels of both FOXAs in non-tumor tissue showed little interindividual variation, greater variation was observed in tumors indicating a destabilization of regulation of these genes during carcinogenesis. A small sub-population of patients showed gene amplification of *FOXA1*. Patients with gene amplification expressed more *FOXA1* mRNA than patients without gene amplification. Increased DNA methylation was found in the *FOXA2* gene in tumors, providing a possible mechanism for the suppressed gene expression. Increased DNA-

methylation was observed in a CpG-island (CpG1) located 1.3 kb upstream of *FOXA2* TSS, while no increase was observed in a CpG-island (CpG2) located 700 bp upstream of the TSS in the Norwegian cohort. DNA-methylation analysis in the US cohort only included two probes associated with *FOXA2.* One of the probes were located between CpG1 and CpG2, which showed increased methylation in tumor compared to non-tumor. This could correspond to the increased methylation in CpG1 of the Norwegian cohort. Contrary to the unchanged DNA methylation levels in CpG2 in the Norwegian cohort, increased DNA methylation levels were found in CpG2 in tumors compared to matched non-tumor tissue in TCGA. The most pronounced difference in DNA methylation between tumor and non-tumor in the Norwegian cohort was observed in an assay located at the 5' end of a CpG island (CpG3, End assay) within the *FOXA2* gene body. DNA methylation in this area was confirmed in the TCGA database. Data from the CURELUNG database contained DNA methylation data from lung tumors only. The highest level of DNA methylation was observed in and around the same region as the End assay from the Norwegian cohort. Interestingly, unsupervised clustering of DNA methylation patterns in the TCGA dataset successfully separated lung tumors from non-tumor tissue indicating distinct DNA methylation patterns. Taken together the results show that deregulation of *FOXA1* and *FOXA2* may be important in lung cancer and that DNA-methylation patterns of *FOXA2* may serve as a potential clinical lung cancer biomarker.

Paper III: The role of FOXA1 in regulation of *CYP1A1* and *CYP1B1* expression in human bronchial epithelial cells.

Steroid receptors have increasingly been implicated in lung cancer, and sex differences in *CYP1A1* expression and B[*a*]P-DNA-adduct formation has been reported. We hypothesized that FOXAs may be involved in regulating gene expression of cytochrome P450s and that this may affect the metabolism of B[*a*]P and consequent DNA-adduct formation. The study focused on FOXA1, which was knocked down by RNA interference in five different human bronchial epithelial cell lines, including HBEC2. Gene expression of *FOXA1, CYP1A1* and *CYP1B1* was analyzed by RT-qPCR in siRNA transfected cells and transformed HBEC2 cells. Binding of FOXA1

to *CYP1A1* and *CYP1B1* regulatory regions was analyzed by chromatin immunoprecipitation (ChIP) and qPCR. B[*a*]P-metabolism was analyzed by LC/MS/MS. B[*a*]P-DNA-adduct formation was analyzed by ^{32}P post-labeling. In transformed HBEC2 cells (from paper I) which show suppressed levels of *FOXA1*, *CYP1A1* was decreased and *CYP1B1* was increased compared to non-transformed cells. The CSC-transformed cell line (T2KT-CSC-H) generated more B[*a*]P metabolites than the non-transformed counterpart. FOXA1 binding was observed in a distal regulatory site of *CYP1B1*. No indication of FOXA1 binding to *CYP1A1* was found. Knockdown of *FOXA1* in five different bronchial epithelial cell lines (HBEC2, HBEC3, HBEC12, BEP2D, BEAS-2B) lead to increased expression of *CYP1B1*. FOXA1 siRNA transfected cells were exposed to E2, TCDD, and B[*a*]P. There was no difference in generation of B[*a*]P metabolites between *FOXA1* siRNA transfected cells and scrambled control siRNA transfected cells. Moreover, the knockdown of *FOXA1* did not show any effect on the formation of B[*a*]P-DNA-adducts. The formation of E2 metabolites was similar between FOXA1 knockdown and control siRNA transfected cells. The only difference in E2 metabolites was observed for 2-MeO-E2 metabolite between cells treated with TCDD and those exposed to E2 only. The *in vitro* data show that while FOXA1 may be involved in regulating gene expression of *CYP1B1*, its contribution is not sufficient for altered metabolism of E2 or B[*a*]P and does not contribute to changes in B[*a*]P-DNA adduct formation.

4 DISCUSSION

Cancer arises as a combined result of genetic and epigenetic alterations triggered by extrinsic (exposure) and intrinsic mechanisms, and individual predisposition. Some cancer types associate more strongly with carcinogen exposure than others. Scrotal cancer was the first cancer type to be connected with a particular occupational risk among London chimney sweepers in 1775 by Sir Percivall Pott [229]. Cook, Hewett, and Hieger identified Benzo[a]pyrene as a carcinogenic constituent of coal tar for the first time in 1932 [230]. Since then B[*a*]P and other PAHs has been found to cause several types of cancers including lung cancer, colorectal adenoma [231] and bladder cancer [232]. B[*a*]P is generated during incomplete combustion of organic material and is found in cigarette smoke, urban air pollution, diesel exhaust, and food products such as cereals, potato chips and fried meat [233].

Transformation of cells *in vitro* has been used to model the carcinogenic effect of diverse chemical compounds. Studies have demonstrated that papillomavirus-immortalized BEP2D can be transformed to grow in soft-agar by exposure to NNK [234] and cigarette smoke condensate [235]. The activated metabolic mutagen BPDE can transform HBECs (HBEC1 and HBEC2) to grow in soft agar *in vitro* [220]. The results from paper I show that long-term exposure to the procarcinogenic B[*a*]P can transform HBECs *in vitro*. B[*a*]P is not a mutagen *per se* but is metabolically activated to a DNA-binding form through the cell's metabolic system. Moreover, the results in paper I show that CSC and MNU are also able to transform HBECs *in vitro* after 12 weeks of exposure. The transformed cells from all exposure scenarios were able to grow anchorage-independent in soft-agar, displayed altered morphology and changes in gene expression programs. These results show that *in vitro* transformation may serve as a useful model for observing the transformational capacity of a range of carcinogenic compounds.

Whether there is a sex difference in lung cancer susceptibility is disputed, but steroid receptors could present molecular differences that are relevant for diagnosis or treatment options [236]. Estrogen receptors are expressed in lung and accumulating evidence suggest a role of estrogen receptors in lung carcinogenesis [137, 138].

Estrogen signaling in lung cells could be involved in the sex differences observed in metabolic capacity [237]. Female lung cells contain more DNA-adducts of B[*a*]P metabolites than male cells which suggest a possible role of steroid signaling in carcinogen metabolism [111]. The two pioneer factors FOXA1 and FOXA2 are crucial in estrogen-regulated gene transcription. In paper I, *FOXA1* and *FOXA2* was down-regulated in all transformed cell lines. Down-regulation of *FOXA2* in transformed lung cells was previously observed after exposure to BPDE [220]. We also found reduced protein levels of FOXA1/2, as observed by immunofluorescent microscopy in transformed HBEC2 cells (T2KT). For transformed HBEC12 cells (T12KT) the difference in fluorescent signal was weak, but there was an indication of a difference in the protein levels. The down-regulation of *FOXA1* and *FOXA2* suggest that they may be critical in maintaining normal cell function and that their loss can lead to cell transformation.

In paper III, the transformed HBEC2 cell lines (T2KT) were found to express lower levels of *CYP1A1*. Moreover, the transformed cell line T2KT-CSC-H (HBEC2, CSC High dose, 3 μg/ml) had higher expression levels of *CYP1B1*. Cell culture media from T2KT-CSC-H was analyzed by LC/MS/MS showing that the transformed cell line had a greater ability to metabolize B[*a*]P than the non-transformed HBEC2 cell line. Greater metabolic activation of B[*a*]P may result in increased mutation rate. It was hypothesized that the altered CYP1 gene expressions and increased B[*a*]P-metabolite levels could be a result of ablated FOXA gene expression in the transformed cells. Given the role of FOXA1 in ERα signaling, the involvement of ERα in regulating *CYP1B1* expression, and the CYP1 enzymes participation in steroid metabolism, a direct regulatory link between FOXA1 and CYP1 is plausible. FOXA1 has recently been found to regulate two other xenobiotic metabolizing genes, *CYP2A13* and *ABCB1* [203], but its role in CYP1 regulation was unknown.

Experiments with siRNAs targeting *FOXA1* showed that down-regulation of *FOXA1* resulted in increased *CYP1B1* expression in HBEC2 cells and four other bronchial epithelial cell lines (HBEC3, HBEC12, BEAS-2B, and BEP2D). Overall, the *CYP1A1* expression was not affected by *FOXA1* siRNA. ChIP against FOXA1 was performed in *CYP1A1* and *CYP1B1* regulatory regions. No binding of FOXA1

was observed in *CYP1A1* in agreement with no general effect of FOXA1 knockdown on *CYP1A1*. In the *CYP1B1* gene, however, FOXA1 binding was found at a distal regulatory site approximately 20 kb upstream of TSS. FOXA1 and ER are known to be able to induce chromatin interaction between distal sites [179], so the distal region in *CYP1B1* is a putative enhancer. The RNA interference and ChIP experiments suggest that *CYP1B1*, but not *CYP1A1*, is directly regulated by FOXA1, and that binding of FOXA1 to a distal regulatory region repress the expression of *CYP1B1*.

The findings of FOXA1's regulation of *CYP1B1* in paper III is limited by the lack of functional investigation of the distal site. The observations suggest a link between FOXA1 binding to the region 20 kb upstream of *CYP1B1* TSS, but whether this region interacts with the *CYP1B1* promoter in an FOXA1 dependent manner and has any effect on *CYP1B1* gene activity remains unknown. However, chromatin interaction data available in ENCODE show that the investigated distal region interacts with *CYP1B1* TSS through ERα, suggesting a functional role of this site. The mechanisms underlying inhibition of *CYP1B1* should be subject to further studies involving chromatin conformation capture and reporter assays to confirm functional enhancer activity.

Cell culture media from *FOXA1* siRNA transfected cells were analyzed by LC/MS/MS to investigate whether the FOXA1 mediated induction of *CYP1B1* had any effect on the generation of B[*a*]P and E2 metabolites. Increased formation of the mutagenic metabolites BPDE and 4-OH-E2 could result in increased mutation rate in cells with low *FOXA1* levels. No effect of *FOXA1* status on the metabolism of B[*a*]P and E2 was observed. Taken together the results in paper III indicate that FOXA1 may repress *CYP1B1* expression, but this has no observed effect on the metabolism of E2 or B[*a*]P.

The transformation of a normal cell to a cancer cell is a multistage process where the cell must overcome several hallmarks of cancer before reaching tumorigenicity. The transcriptional reprogramming of EMT may be pivotal in establishing invasiveness, metastasis, evasion of immunosuppression and chemoresistance of cancer cells [147, 148, 165, 166, 175]. One of the main events in EMT is a switch from *CDH1* expression to *CDH2* expression (the cadherin switch). *CDH1* has been reported to

be down-regulated in cells transformed by BPDE exposure [220]. In paper I, we show that *CDH1* was down-regulated in all transformed cell lines. Loss of E-cadherin is associated with increased local invasiveness and metastasis [238, 239]. The transformed T2KT showed increased ability to invade a matrigel in wound healing assay. In contrast, migration was decreased in the T2KT cells compared to non-transformed HBEC2. Reduced cell migration in EMT cells *in vitro* has previously been reported by others, suggesting that migration and invasion are uncoupled events [240]. This finding suggests that invasiveness may be a better than migration as a measure of EMT and cell transformation in the transformation model in paper I.

Many transcription factors are thought to be important in EMT. Among the most established EMT driver TFs are SNAI1, SNAI2, ZEB1, and TWIST1. All of these TFs repress the expression of *CDH1* [241-243]. *SNAI1* and *ZEB1* were up-regulated in transformed versus non-transformed HBEC2 cell lines, while the transformed HBEC12 cell lines were found to express EMT-TFs to the same degree as non-transformed cells. The unclear patterns of EMT-TF transcription in HBEC12 suggest that other mechanisms than TF mediated repression are involved in suppressing *CDH1*. Epigenetic reprogramming has been implicated in EMT. Histone modifications, DNA methylation and microRNAs (miR-9) have all been found to repress *CDH1* [244, 245]. Epigenetic regulation of *CDH1* was not studied in paper I, so the exact mechanisms regulating *CDH1* in this cell transformation model remains to be investigated. SNAI2 is, as SNAI1 and ZEB1, involved in repression of *CDH1*. Therefore, it was interesting to note that *SNAI2* is down-regulated in the transformed HBECs. Moreover, *SNAI2* is directly targeted and repressed by FOXA1 which binds in the *SNAI2* enhancer [246, 247], meaning that a decrease in *FOXA1* level, as observed in the transformed cell lines, would result in induced expression of *SNAI2* [246, 247]. In conclusion, the results from paper I suggest that carcinogen-induced cell transformation involves loss of epithelial markers and progression of an EMT.

As models for carcinogenesis studies, cell culture experiments have several limitations. For example, cells *in vivo* grow in a three-dimensional environment, while they in cell culture grow in a two-dimensional monolayer, which may not be a good

representation of physiological conditions. Also, during prolonged cell culturing, natural selection does occur, and evolutionary drift could be a concern. Authentication of cell lines is an important measure in avoiding such drift as well as cross contamination by other cell lines. The HBECs used for the transformation-study in paper I was recently verified by an external laboratory (Leibniz Institut DSMZ). The HBECs are immortalized by insertion of the *TERT* gene which provides the cell with immortality, and the *CDK4* gene which confers the cell prolonged proliferation [219]. With these two genes, the HBECs are given a kickstart in the transformation process. The *TERT* and *CDK4* are both often found to be altered in cell transformation *in vitro* as well as in tumors [90, 218, 248-250]. Nevertheless, the HBECs do not grow in soft agar and are not tumorigenic in immunodeficient mice, like tumor-derived cell lines can be. Furthermore, HBECs do not enter senescence, as primary cell lines do. The capability for prolonged culturing of non-transformed cell make the HBECs a useful model for studying carcinogenesis; they are not transformed but still keep on proliferating.

In the transformation study conducted for paper I, human lung cells were exposed to three different carcinogenic compounds, MNU, B[*a*]P and CSC. It is interesting to note that all three exposures lead to transformed cells with similar characteristics. Other studies have found similar results: exposure of lung cell lines to BPDE, CSC or even spontaneous transformation of prostate cells without exposure, lead to cells with mesenchymal characteristics that could grow in soft agar [220, 227]. These observations raise the intriguing question of whether exposure type specifies the features of the transformed cells or if the mechanisms underlying cell transformation are similar regardless of the type of exposure. One possibility is that the transformed phenotype is selected for under the applied pressure of the soft-agar assay. In our study, *FOXA1/2* and *CDH1* were down-regulated in HBEC2 after six weeks of carcinogen exposure and before full transformation. *CDH2* was up-regulated at the same time point, indicating that the cadherin switch and *FOXA1/2* ablation occurs as a result of exposure and cellular stress, rather than an effect of selection in soft-agar. Alterations of FOXAs and CDHs may be early events in carcinogenesis, but they are apparently not sufficient alone for full transformation. Other, later events are necessary for anchorage-independent growth. The *in vitro* transformation model is

useful in establishing an understanding of the mechanisms regulating EMT and can potentially uncover biomarkers for detection of cancer or targets for treatment.

In paper II, the expression of *FOXA1* and *FOXA2* was studied in matched pairs of tumor and non-tumor from a total of 268 patients from three geographically independent cohorts (the USA, Italy, and Norway). A small variation in expression of *FOXA1* and *FOXA2* was observed in non-tumor, indicating that a relatively stable expression of FOXAs is necessary for normal lung cell function. The variable *FOXA1* and *FOXA2* expression in NSCLC could disrupt the normal lung cell function. *FOXA2* was found to be down-regulated in the tumor, as it was in the transformed cells *in vitro*. In contrast to the *in vitro* transformed cells, *FOXA1* was increased in lung tumor compared to matched non-tumor. The discrepancy of *FOXA1* expression *in vitro* and *in vivo* is unknown. Similar to the results in paper II, a recent study showed that expression of *FOXA1* was increased, and expression of *FOXA2* decreased in NSCLC [251]. *FOXA1* is for lung cancer and several other cancer types considered an oncogene, while it is associated with good prognosis in breast cancer. Increased *FOXA1* expression has previously been reported as promoting EMT in lung cell lines [204]. *FOXA2* on the other hand, show opposite patterns with tumor suppressor function in tissues such as lung and prostate, and oncogene function in the breast. Thus, the role of *FOXA1* and *FOXA2* in cancer may depend on tissue and steroid sensitivity of the tumor cells.

The observation that increased *FOXA1* and decreased *FOXA2* gene expression occurs in three different histologies of NSCLC suggest that the deregulation of *FOXA1* and *FOXA2* may serve as a common mechanism of lung carcinogenesis. Previous studies reported increased *FOXA1* in lung adenocarcinoma [189], lung squamous cell carcinoma [190], pancreatic cancer [191], prostate cancer [192] and thyroid cancer [193] and *FOXA1* is considered an oncogene in these cancer types. The expression of *FOXA2* is suppressed in lung cancer cell lines [196, 197] and gastric cancer [198], and was suggested to be important in suppressing metastasis in lung adenocarcinomas [199].

A study in prostate cells found that the expressions of more than 900 genes were altered by *FOXA1* knockdown [247] and another study found by ChIP-seq that

FOXA1 bind to 40,000 – 80,000 sites in mammary gland cancer cell lines [185]. If chemical exposure deregulates *FOXA1* transcription, it may have a great impact on the chromatin landscape and the transcriptional program of the cell due to the high number of FOXA1 targets and FOXA1's ability to bind compact chromatin and remodel chromatin structure [66].

Significantly increased expression of *FOXA1* was found in tumor tissue compared to matched non-tumor tissues from all three cohorts studied in paper II. Moreover, the increased *FOXA1* expression occurred independently of tumor histology. Mutations and CNV have been reported for *FOXA1* in lung adenocarcinoma, resulting in higher expression of *FOXA1* in tumors [252]. It is apparent that regulation of *FOXA1* in lung cancer may be complex. In paper II, lung cancer tissue samples were analyzed by qPCR of genomic DNA to identify whether the observed higher expression levels in a tumor could be due to gene amplification. The results showed that amplification rate was low in ADCA and SCC with 4.9% and 3.1%, respectively. The percentage of cases with amplification of FOXA1 was higher in LCC with 26.3%, although the total number of LCC cases were quite small and gave a greater degree of uncertainty to the data. The cases with amplification of *FOXA1* showed higher average expression level of *FOXA1* than cases with no amplification of *FOXA1*. Recently published data from the TCGA found that *FOXA1* amplifications in 6.4% (73/1144: 9.1% ADCA, 2.5% SCC) of NSCLCs (ADCA and SCC), similar to our findings [38]. The results from the *FOXA1* gene amplification analysis indicate that amplification of *FOXA1* accounts for the highest expression levels of *FOXA1*, but apart from a minor fraction of tumors showing *FOXA1* gene amplification, the mechanisms behind increased expression of *FOXA1* in NSCLC tissue of all major histology's remains unknown.

No difference in expression of *FOXA1* or *FOXA2* was observed between men and women (paper II). *FOXA1* has been associated with poor prognosis in prostate cancer and good prognosis in female breast cancer. This discrepancy indicates that the function of *FOXA1* may be dependent on the cooperating steroid factor and that ER and AR may have opposite effect. The female and male lung express similar levels of ER, but female ADCAs may be more sensitive to E2 [253, 254]. It is possible that even though the expression of FOXAs in men and women are equal, the functional

result on lung cell or lung tumor physiology may differ between the sexes. Future studies should look into the gender-specific effect of FOXAs in lung tumors. Experiments could be done *in vitro* in male and female cell lines or by generating transgenic cell lines with manipulated (overexpression and knockdown) expression of ER or AR, along with treatment of steroids.

Tumors are heterogenic and consist of cells with different genotype and gene expression patterns. The same is the case for non-tumor cells and a recent study show that normal (eyelid) skin consist of heterogeneous cell populations that are not malignant but still contain many of the same alterations as in cancer cells [255]. The consequences of heterogeneity in a matched pair study as the one conducted in paper II is that changes that are drivers but not sufficient for tumorigenicity can be present in the non-tumor tissue. The presence of drivers in non-tumor tissue can confound actual drivers in the tumor. Other issues with analyzing tissue samples are that i) the sampled tissue may contain a small subpopulation with genetic, epigenetic and gene expression patterns not present in the majority of the tumor cells, ii) the tumor may contain surrounding normal tissue distorting the identification of drivers iii) tumors are infiltrated by stromal cells that can occlude downstream analysis. Particular care was taken when isolating tumors analyzed in paper II. The tumors were pathologically classified by experienced pathologists and only tumor tissues containing at least 80% of tumor cells were analyzed in the study.

The regulation of FOXAs may be due to epigenetic changes, as shown by DNA methylation in paper II and histone modifications in paper I. MacroH2A bind *FOXA2* in cell lines that do not express *FOXA2* but is absent in cell lines that do express the gene [256]. In paper I, macroH2A binding is increased in transformed cell lines compared to the non-transformed cells. MacroH2A is associated with transcriptional repression [71] and has previously been associated with EMT in breast cancer cells [257] and positive prognosis in lung cancer [72]. An interesting observation in the ChIP-qPCR experiments is that all investigated chromatin locations had higher levels of histone binding in the T2KT transformed cell lines (T2KT-CSC-L and T2KT-CSC-H). These findings suggest that suppression of *FOXA1* and *FOXA2* in these cell lines may be due to increased nucleosome density

and macroH2A enrichment, resulting in compaction of chromatin and gene silencing. Histone H2A.Z was also found to be enriched at *FOXA1* and *FOXA2* in transformed T2KTs. Histone H2A.Z serves a different function in regulating transcription depending on location. When bound in gene bodies H2A.Z is associated with repressed gene activity [258], consistent with the findings in the paper I where H2A.Z binding was enriched in the gene body assays. In promoters, on the other hand, H2A.Z is associated with nucleosome depletion and active transcription [259, 260]. The enrichment of H2A.Z in promoters in the transformed cell lines is, therefore, unexpected.

In paper II, DNA hypermethylation was found by bisulfite pyrosequencing in several regions of *FOXA2* in tumors from the Norwegian cohort. Specifically, a CpG-island located upstream of *FOXA2* (CpG1 assay) and an assay located within the gene body at the end of a long CpG-island spanning the length of the *FOXA2* gene (End assay), were found to contain the highest DNA methylation levels in both tumor and non-tumor tissue. The DNA-methylation level was elevated even further in tumors compared to non-tumor for both CpG1 (14.8% increase) and End (20.4% increase). In the US cohort, two Illumina Infinium 27K DNA methylation probes, one located between two CpG-islands (CpG1 and CpG2) upstream of *FOXA2* and within the *FOXA2* gene body (CpG3), both showed increased DNA methylation in tumors compared to matched non-tumors. To confirm these findings, we mined two different databases, TCGA and CURELUNG, both of which contain Illumina Infinium 450K DNA methylation data. The TCGA included matched pair tumor and non-tumor and showed increased DNA methylation patterns in CpG2 and around the End assay in CpG3.

An interesting observation is that in unsupervised clustering, cases from the TCGA clusters into the tumor and non-tumor depending on the *FOXA2* DNA methylation levels. Moreover, cases from the CURELUNG project show distinct DNA methylation patterns in CpG3 in adenocarcinoma and squamous cell carcinoma. DNA methylation patterns have previously been shown to separate different subtypes of NSCLC, including ADC and SCC, but also NSCLC with epithelial or mesenchymal characteristics or adenocarcinomas with or without EGFR mutations

[54-56]. These earlier findings suggest that distinct DNA methylation patterns could represent potential biomarkers for detection of cancer as well as sub-classification of cancer types. The result in paper II indicates that DNA methylation patterns in *FOXA2* may be used for diagnosing and classifying NSCLC. The analysis of DNA methylation in a single gene compared to the whole genome of a patient could provide a quick method for correct diagnosis.

The three papers conducted for this thesis together support a role of *FOXA1* and *FOXA2* in lung cancer and lung carcinogenesis *in vitro* and *in vivo*. Paper I note that the expression of the FOXA genes is down-regulated by long-term tobacco smoke carcinogen exposure and that this coincides with EMT. Paper II shows deregulation of both *FOXA1* and *FOXA2* expression in cancer, partly due to DNA amplification and DNA methylation, respectively. The finding of increased *FOXA1* and decreased *FOXA2* gene expression occurs in three different histologies of NSCLC suggest that their deregulation may serve as a common mechanism of lung carcinogenesis. It is becoming increasingly clear that steroid receptors and steroid signaling pathways are involved in lung cancer. In paper II no effect of sex was observed in the gene expression patterns of *FOXA1* or *FOXA2*. In paper III, *FOXA1* expression affected CYP1B1 transcription, but this had no consequence for B[*a*]P and E2 metabolism. Together these results do not support a functional role of FOXAs in metabolic differences between the sexes during carcinogenesis. FOXA1/2 may, however, affect lung cancer in other steroid dependent ways not studied here.

EMT has been invoked in chemoresistance, resistance to EGFR tyrosine kinase inhibitors and immune evasion. Since *FOXA1* and *FOXA2* are associated with EMT in paper I and changes in *FOXA1/2* expression *in vivo* in paper II, it is intriguing to suggest the possibility that manipulating these transcription factors could sensitize resistant tumors to conventional treatment of lung cancer. This thesis lay the grounds for further studies to uncover the potential of *FOXA1* and *FOXA2* as biomarkers in diagnosis, histology classification, and possible treatment options in lung cancer.

5 FURTHER STUDIES

Some interesting findings were made in the three papers conducted for this thesis that can lay the basis for further research. One limitation of the transformation study was that the role of FOXA1/2 in EMT was not investigated mechanistically. Further studies should examine the direct effect of FOXA1/2 in regulating EMT TFs through ChIP-seq to unravel the full contribution of these two pioneer factors in activating EMT in lung cells. Furthermore, the effect of FOXA1/2 in establishing and maintaining an epithelial epigenotype, and the effect of FOXA loss should be studied. The HBEC and their transformed derivatives provide an excellent platform for such research. Loss of FOXA function in HBEC and overexpression of FOXAs in transformed cells in parallel with ChIP of key histone mark at EMT genes (or ChIP-seq) or whole genome bisulfite methyl-DNA-seq could provide a thorough understanding of FOXA function in lung cells.

Several studies have found that DNA-methylation, CNV, and SNV are mutually exclusive mechanisms in gene silencing of tumor suppressors or activating oncogenes. In paper II, CNV was investigated as an underlying mechanism of increased *FOXA1* expression in tumors. Only a small subset of tumors (6.7%) were found to have gene amplifications of *FOXA1*. For the cases that did not harbor FOXA1 amplifications, but still expressed a higher level of *FOXA1* in tumor than in matched non-tumor tissue, another mechanism for gene regulation must be present. Further studies should look into resolving the mechanisms underlying increased gene expression, such as mutations in response elements in promoter and enhancer or changes in gene expression of TFs that regulate *FOXA1*. The same goes for *FOXA2*, where DNA methylation is found to possibly down-regulate gene expression in lung tumors. Functional studies of the identified methylated regions should be conducted to elucidate which regions have the greatest impact on *FOXA2* transcription.

Both *FOXA1* and *FOXA2* was down-regulated during EMT in the transformation assay, while the FOXA1 expression was increased and FOXA2 expression was decreased in lung tumor. The dissimilarity between the *in vivo* and *in vitro* studies raises the question whether both *FOXA1* and *FOXA2* is necessary for inducing EMT

or if down-regulation of *FOXA2* is sufficient. One way to investigate this is to simulate the observed expression patterns in lung cancer in cells *in vitro*. *FOXA1* and *FOXA2* could individually be overexpressed by expression vectors in the transformed cells to study the effect on invasiveness and expression of EMT markers. Reversely, *FOXA1* and *FOXA2* could individually be knocked down (or out) in non-transformed HBEC, as an extension of the siRNA experiments performed for paper III.

Considering the results from papers I and II, it would be interesting to study how *FOXA1/2* expression correlate with EMT *in vivo* by measuring gene expression of EMT markers in matched tumor and non-tumor tissue. The finding that the cadherin switch and *FOXA1/2* suppression occurs before transformation suggest that this is a result of exposure rather than selection in soft agar, indicating that EMT and FOXA TFs are important in tobacco-related lung carcinogenesis. As a future study, it would be interesting to examine if the suppression of *FOXA1* and *FOXA2* is a direct result of exposure and if the AHR mediates this repression.

A large-scale study could be performed on the transformed cells. Whole genome DNA-seq on transformed and non-transformed HBEC could provide a map of the mutational burden induced by carcinogen exposure. Exome sequencing could be added to unveil the effect that mutations have on the transcriptional program. Then add whole genome bisulfite-seq and histone ChIP-seq to map the epigenetic landscape of the transformed and non-transformed cell lines. The goal of such a study would be to identify which alterations are due to mutations and DNA damage, and which alterations are due to signal disruptive properties of B[*a*]P, CSC, and MNU. This would provide a thorough understanding of how the environment affects our genome, and how carcinogens affect the transcriptional program of human lung cells and induce transcriptional reprogramming to a promote a transformed, EMT-like phenotype.

REFERENCES

1. Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang XS, et al. Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). Lancet. 2015;385(9972):977-1010. doi:10.1016/s0140-6736(14)62038-9

2. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, et al. SEER Cancer Statistics Review, 1975-2013 SEER web site: National Cancer Institute; 2016 [updated 04.15.16. Available from: http://seer.cancer.gov/csr/1975_2013/.

3. Robles AI, Harris CC. Integration of multiple "OMIC" biomarkers: A precision medicine strategy for lung cancer. Lung Cancer. 2016;Article in Press. doi:10.1016/j.lungcan.2016.06.003

4. Cancer Registry of Norway. Cancer in Norway 2014 - Cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway; 2015. Report No.: 978-52-90343-91-4.

5. Thun M, Peto R, Boreham J, Lopez AD. Stages of the cigarette epidemic on entering its second century. Tob Control. 2012;21(2):96-101. doi:10.1136/tobaccocontrol-2011-050294

6. Torre LA, Siegel RL, Ward EM, Jemal A. International variation in lung cancer mortality rates and trends among women. Cancer Epidemiol Biomarkers Prev. 2014;23(6):1025-1036. doi:10.1158/1055-9965.epi-13-1220

7. Risch HA, Howe GR, Jain M, Burch JD, Holowaty EJ, Miller AB. Are female smokers at higher risk for lung cancer than male smokers? A case-control analysis by histologic type. Am J Epidemiol. 1993;138(5):281-293. Available from:

8. Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. J Natl Cancer Inst. 1996;88(3-4):183-192. Available from:

9. Kreuzer M, Boffetta P, Whitley E, Ahrens W, Gaborieau V, Heinrich J, et al. Gender differences in lung cancer risk by smoking: a multicentre case-control study in Germany and Italy. Br J Cancer. 2000;82(1):227-233. doi:10.1054/bjoc.1999.0904

10. Bain C, Feskanich D, Speizer FE, Thun M, Hertzmark E, Rosner BA, et al. Lung cancer rates in men and women with comparable histories of smoking. J Natl Cancer Inst. 2004;96(11):826-834. Available from:

11. Flanders WD, Lally CA, Zhu BP, Henley SJ, Thun MJ. Lung cancer mortality in relation to age, duration of smoking, and daily cigarette consumption: results from Cancer Prevention Study II. Cancer Res. 2003;63(19):6556-6562. Available from:

12. Lee PN, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. BMC Cancer. 2012;12:90. doi:10.1186/1471-2407-12-385

13. Malvezzi M, Bertuccio P, Rosso T, Rota M, Levi F, La Vecchia C, et al. European cancer mortality predictions for the year 2015: does lung cancer have the highest death rate in EU women? Ann Oncol. 2015;26(4):779-786. doi:10.1093/annonc/mdv001

14. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016;893:1-19. doi:10.1007/978-3-319-24223-1_1

15. Motadi LR, Misso NL, Dlamini Z, Bhoola KD. Molecular genetics and mechanisms of apoptosis in carcinomas of the lung and pleura: therapeutic targets. Int Immunopharmacol. 2007;7(14):1934-1947. doi:10.1016/j.intimp.2007.07.013

16. Yokota J, Kohno T. Molecular footprints of human lung cancer progression. Cancer Sci. 2004;95(3):197-204. Available from:

17. Samet JM, Avila-Tang E, Boffetta P, Hannan LM, Olivo-Marston S, Thun MJ, et al. Lung cancer in never smokers: clinical epidemiology and environmental risk factors. Clin Cancer Res. 2009;15(18):5626-5645. doi:10.1158/1078-0432.ccr-09-0376

18. Proctor RN. The history of the discovery of the cigarette-lung cancer link: evidentiary traditions, corporate denial, global toll. Tob Control. 2012;21(2):87-91. doi:10.1136/tobaccocontrol-2011-050338

19. Youlden DR, Cramb SM, Baade PD. The International Epidemiology of Lung Cancer: geographical distribution and secular trends. J Thorac Oncol. 2008;3(8):819-831. doi:10.1097/JTO.0b013e31818020eb

20. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012;100(Pt E):1-538. Available from:

21. Phillips DH, Venitt S. DNA and protein adducts in human tissues resulting from exposure to tobacco smoke. Int J Cancer. 2012;131(12):2733-2753. doi:10.1002/ijc.27827

22. Hecht SS, Stepanov I, Carmella SG. Exposure and Metabolic Activation Biomarkers of Carcinogenic Tobacco-Specific Nitrosamines. Acc Chem Res. 2016;49(1):106-114. doi:10.1021/acs.accounts.5b00472

23. Foiles PG, Akerkar SA, Carmella SG, Kagan M, Stoner GD, Resau JH, et al. Mass spectrometric analysis of tobacco-specific nitrosamine-DNA adducts in smokers and nonsmokers. Chem Res Toxicol. 1991;4(3):364-368. Available from:

24. Nafstad P, Haheim LL, Oftedal B, Gram F, Holme I, Hjermann I, et al. Lung cancer and air pollution: a 27 year follow up of 16 209 Norwegian men. Thorax. 2003;58(12):1071-1076. Available from:

25. Boffetta P. Human cancer from environmental pollutants: the epidemiological evidence. Mutat Res. 2006;608(2):157-162. doi:10.1016/j.mrgentox.2006.02.015

26. Malhotra J, Sartori S, Brennan P, Zaridze D, Szeszenia-Dabrowska N, Swiatkowska B, et al. Effect of occupational exposures on lung cancer susceptibility: a study of gene-environment interaction analysis. Cancer Epidemiol Biomarkers Prev. 2015;24(3):570-579. doi:10.1158/1055-9965.epi-14-1143-t

27. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. Eur Respir J. 2016;48(3):889-902. doi:10.1183/13993003.00359-2016

28. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719-724. doi:10.1038/nature07943

29. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013

30. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70. Available from:

31. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. N Engl J Med. 2012;366(10):883-892. Available from: <Go to ISI>://WOS:000301172500005

32. Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature. 2010;463(7278):184-190. doi:10.1038/nature08629

33. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. 2010;463(7278):191-196. doi:10.1038/nature08658

34. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell. 2012;150(6):1121-1134. doi:10.1016/j.cell.2012.08.024

35. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-421. doi:10.1038/nature12477

36. Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S, et al. Clock-like mutational processes in human somatic cells. Nat Genet. 2015;47(12):1402-1407. doi:10.1038/ng.3441

37. Alexandrov LB, Stratton MR. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. Curr Opin Genet Dev. 2014;24:52-60. doi:10.1016/j.gde.2013.11.014

38. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet. 2016;48(6):607-616. doi:10.1038/ng.3564

39. Hainaut P, Pfeifer GP. Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. Carcinogenesis. 2001;22(3):367-374. Available from:

40. Jia P, Pao W, Zhao Z. Patterns and processes of somatic mutations in nine major cancers. BMC Med Genomics. 2014;7:11. doi:10.1186/1755-8794-7-11

41. Langie SA, Koppen G, Desaulniers D, Al-Mulla F, Al-Temaimi R, Amedei A, et al. Causes of genome instability: the effect of low dose chemical exposures in modern society. Carcinogenesis. 2015;36 Suppl 1:S61-88. doi:10.1093/carcin/bgv031

42. Yang Z, Zhuan B, Yan Y, Jiang S, Wang T. Integrated analyses of copy number variations and gene differential expression in lung squamous-cell carcinoma. Biol Res. 2015;48(1):47. doi:10.1186/s40659-015-0038-3

43. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet. 2013;45(10):1134-1140. doi:10.1038/ng.2760

44. Ocak S, Yamashita H, Udyavar AR, Miller AN, Gonzalez AL, Zou Y, et al. DNA copy number aberrations in small-cell lung cancer reveal activation of the focal adhesion pathway. Oncogene. 2010;29(48):6331-6342. doi:10.1038/onc.2010.362

45. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A. 2005;102(30):10604-10609. doi:10.1073/pnas.0500398102

46. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. Nature Reviews Genetics. 2012;13(2):97-109. doi:10.1038/nrg3142

47. Probst AV, Dunleavy E, Almouzni G. Epigenetic inheritance during the cell cycle. Nature Reviews Molecular Cell Biology. 2009;10(3):192-206. doi:10.1038/nrm2640

48. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001;293(5532):1068-1070. doi:10.1126/science.1063852

49. Toyota M, Suzuki H, Yamashita T, Hirata K, Imai K, Tokino T, et al. Cancer epigenomics: implications of DNA methylation in personalized cancer therapy. Cancer Sci. 2009;100(5):787-791. doi:10.1111/j.1349-7006.2009.01095.x

50. Toyota M, Issa JP. CpG island methylator phenotypes in aging and cancer. Semin Cancer Biol. 1999;9(5):349-357. doi:10.1006/scbi.1999.0135

51. Reynolds PA, Sigaroudinia M, Zardo G, Wilson MB, Benton GM, Miller CJ, et al. Tumor suppressor p16INK4A regulates polycomb-mediated DNA hypermethylation in human mammary epithelial cells. J Biol Chem. 2006;281(34):24790-24802. doi:10.1074/jbc.M604175200

52. Dumont N, Wilson MB, Crawford YG, Reynolds PA, Sigaroudinia M, Tlsty TD. Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers. Proc Natl Acad Sci U S A. 2008;105(39):14867-14872. doi:10.1073/pnas.0807146105

53. Costello JF, Frühwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet. 2000;24(2):132-138. doi:10.1038/72785

54. Lin SH, Wang J, Saintigny P, Wu C-C, Giri U, Zhang J, et al. Genes suppressed by DNA methylation in non-small cell lung cancer reveal the epigenetics of epithelial–mesenchymal transition. BMC Genomics. 2014;15(1):1-15. doi:10.1186/1471-2164-15-1079

55. Walter K, Holcomb T, Januario T, Du P, Evangelista M, Kartha N, et al. DNA methylation profiling defines clinically relevant biological subsets of non-small cell lung cancer. Clin Cancer Res. 2012;18(8):2360-2373. doi:10.1158/1078-0432.ccr-11-2635-t

56. Shinjo K, Okamoto Y, An B, Yokoyama T, Takeuchi I, Fujii M, et al. Integrated analysis of genetic and epigenetic alterations reveals CpG island methylator phenotype associated with distinct clinical characters of lung adenocarcinoma. Carcinogenesis. 2012;33(7):1277-1285. doi:10.1093/carcin/bgs154

57. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993;75(5):843-854. Available from:

58. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature. 2010;466(7308):835-840. doi:10.1038/nature09267

59. Larsson O, Nadon R. Re-analysis of genome wide data on mammalian microRNA-mediated suppression of gene expression. Translation. 2013;1(1):e24557. doi:10.4161/trla.24557

60. Graff J, Kim D, Dobbin MM, Tsai LH. Epigenetic regulation of gene expression in physiological and pathological brain processes. Physiol Rev. 2011;91(2):603-649. doi:10.1152/physrev.00012.2010

61. Lynam-Lennon N, Maher SG, Reynolds JV. The roles of microRNA in cancer and apoptosis. Biol Rev Camb Philos Soc. 2009;84(1):55-71. doi:10.1111/j.1469-185X.2008.00061.x

62. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol. 2015;16(3):144-154. doi:10.1038/nrm3949

63. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007;39(3):311-318. doi:10.1038/ng1966

64. Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science. 2001;292(5514):110-113. doi:10.1126/science.1060118

65. Strahl BD, Ohba R, Cook RG, Allis CD. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proc Natl Acad Sci U S A. 1999;96(26):14967-14972. Available from:

66. Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, et al. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell. 2008;132(6):958-970. doi:10.1016/j.cell.2008.01.018

67. Ernst J, Kellis M. Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat Biotechnol. 2010;28(8):817-825. doi:10.1038/nbt.1662

68. Thambirajah AA, Li A, Ishibashi T, Ausio J. New developments in post-translational modifications and functions of histone H2A variants. Biochem Cell Biol. 2009;87(1):7-17. doi:10.1139/o08-103

69. Bernstein E, Hake SB. The nucleosome: a little variation goes a long way. Biochem Cell Biol. 2006;84(4):505-517. doi:10.1139/o06-085

70. Reamon-Buettner SM, Borlak J. Dissecting epigenetic silencing complexity in the mouse lung cancer suppressor gene Cadm1. PLoS One. 2012;7(6):e38531. doi:10.1371/journal.pone.0038531

71. Costanzi C, Pehrson JR. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature. 1998;393(6685):599-601. doi:10.1038/31275

72. Sporn JC, Kustatscher G, Hothorn T, Collado M, Serrano M, Muley T, et al. Histone macroH2A isoforms predict the risk of lung cancer recurrence. Oncogene. 2009;28(38):3423-3428. doi:10.1038/onc.2009.26

73. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546-1558. doi:10.1126/science.1235122

74. Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. Nat Rev Cancer. 2013;13(7):497-510. doi:10.1038/nrc3486

75. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. Science. 2015;349(6255):1483-1489. doi:10.1126/science.aab4082

76. Tomasetti C, Marchionni L, Nowak MA, Parmigiani G, Vogelstein B. Only three driver gene mutations are required for the development of lung and colorectal cancers. Proc Natl Acad Sci U S A. 2015;112(1):118-123. doi:10.1073/pnas.1421839112

77. Bouaoun L, Sonkin D, Ardin M, Hollstein M, Byrnes G, Zavadil J, et al. TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. Hum Mutat. 2016;37(9):865-876. doi:10.1002/humu.23035

78. Hammerman PS, Lawrence MS, Voet D, Jing R, Cibulskis K, Sivachenko A, et al. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489(7417):519-525. doi:10.1038/nature11404

79. Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, et al. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):543-550. doi:10.1038/nature13385

80. Lane DP. Cancer. p53, guardian of the genome. Nature. 1992;358(6381):15-16. doi:10.1038/358015a0

81. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004;350(21):2129-2139. doi:10.1056/NEJMoa040938

82. Bergethon K, Shaw AT, Ou SHI, Katayama R, Lovly CM, McDonald NT, et al. ROS1 Rearrangements Define a Unique Molecular Class of Lung Cancers. J Clin Oncol. 2012;30(8):863-870. doi:10.1200/jco.2011.35.6345

83. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. 2007;448(7153):561-566. doi:10.1038/nature05945

84. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, et al. Anaplastic Lymphoma Kinase Inhibition in Non-Small-Cell Lung Cancer. N Engl J Med. 2010;363(18):1693-1703. doi:10.1056/NEJMoa1006448

85. Camidge DR, Pao W, Sequist LV. Acquired resistance to TKIs in solid tumours: learning from lung cancer. Nature Reviews Clinical Oncology. 2014;11(8):473-481. doi:10.1038/nrclinonc.2014.104

86. Gazdar AF. Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. Oncogene. 2009;28:S24-S31. doi:10.1038/onc.2009.198

87. Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, Mulvey HE, et al. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat Med. 2016;22(3):262-269. doi:10.1038/nm.4040

88. Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, et al. Gefitinib or Chemotherapy for Non-Small-Cell Lung Cancer with Mutated EGFR. N Engl J Med. 2010;362(25):2380-2388. doi:10.1056/NEJMoa0909530

89. Bass AJ, Watanabe H, Mermel CH, Yu S, Perner S, Verhaak RG, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet. 2009;41(11):1238-1242. doi:10.1038/ng.465

90. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature. 2007;450(7171):893-898. doi:10.1038/nature06358

91. Martello G, Rosato A, Ferrari F, Manfrin A, Cordenonsi M, Dupont S, et al. A MicroRNA targeting dicer for metastasis control. Cell. 2010;141(7):1195-1207. doi:10.1016/j.cell.2010.05.017

92. Karube Y, Tanaka H, Osada H, Tomida S, Tatematsu Y, Yanagisawa K, et al. Reduced expression of Dicer associated with poor prognosis in lung cancer patients. Cancer Sci. 2005;96(2):111-115. doi:10.1111/j.1349-7006.2005.00015.x

93. Li B, Ren SX, Li XF, Wang YS, Garfield D, Zhou SW, et al. MiR-21 overexpression is associated with acquired resistance of EGFR-TKI in non-small cell lung cancer. Lung Cancer. 2014;83(2):146-153. doi:10.1016/j.lungcan.2013.11.003

94. Singh M, Garg N, Venugopal C, Hallett R, Tokar T, McFarlane N, et al. STAT3 pathway regulates lung-derived brain metastasis initiating cell capacity through miR-21 activation. Oncotarget. 2015;6(29):27461-27477. doi:10.18632/oncotarget.4742

95. Yan LX, Huang XF, Shao Q, Huang MY, Deng L, Wu QL, et al. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. Rna-a Publication of the Rna Society. 2008;14(11):2348-2360. doi:10.1261/rna.1034808

96. Schramedei K, Morbt N, Pfeifer G, Lauter J, Rosolowski M, Tomm JM, et al. MicroRNA-21 targets tumor suppressor genes ANP32A and SMARCA4. Oncogene. 2011;30(26):2975-2985. doi:10.1038/onc.2011.15

97. Asangani IA, Rasheed SAK, Nikolova DA, Leupold JH, Colburn NH, Post S, et al. MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. Oncogene. 2008;27(15):2128-2136. doi:10.1038/sj.onc.1210856

98. Yang CH, Yue JM, Pfeffer SR, Handorf CR, Pfeffer LM. MicroRNA miR-21 Regulates the Metastatic Behavior of B16 Melanoma Cells. J Biol Chem. 2011;286(45):39172-39178. doi:10.1074/jbc.M111.285098

99. Liu Y, Luo F, Wang BR, Li HQ, Xu Y, Liu XL, et al. STAT3-regulated exosomal miR-21 promotes angiogenesis and is involved in neoplastic processes of transformed human bronchial epithelial cells. Cancer Lett. 2016;370(1):125-135. doi:10.1016/j.canlet.2015.10.011

100. Zhao Y, Xu Y, Li Y, Xu WC, Luo F, Wang BR, et al. NF-B-Mediated Inflammation Leading to EMT via miR-200c Is Involved in Cell Transformation Induced By Cigarette Smoke Extract. Toxicol Sci. 2013;135(2):265-276. doi:10.1093/toxsci/kft150

101. Hecht SS. Tobacco smoke carcinogens and lung cancer. J Natl Cancer Inst. 1999;91(14):1194-1210. doi:10.1093/jnci/91.14.1194

102. Phillips DH, Hewer A, Martin CN, Garner RC, King MM. Correlation of DNA adduct levels in human lung with cigarette smoking. Nature. 1988;336(6201):790-792. doi:10.1038/336790a0

103. Phillips DH, Schoket B, Hewer A, Bailey E, Kostic S, Vincze I. Influence of cigarette smoking on the levels of DNA adducts in human bronchial epithelium and white blood cells. Int J Cancer. 1990;46(4):569-575. Available from:

104. Moorthy B, Chu C, Carlin DJ. Polycyclic aromatic hydrocarbons: from metabolism to lung cancer. Toxicol Sci. 2015;145(1):5-15. doi:10.1093/toxsci/kfv040

105. Wislocki PG, Wood AW, Chang RL, Levin W, Yagi H, Hernandez O, et al. High mutagenicity and toxicity of a diol epoxide derived from benzo(a)pyrene. Biochem Biophys Res Commun. 1976;68(3):1006-1012. Available from:

106. Osborne MR, Beland FA, Harvey RG, Brookes P. The reaction of (+/-)-7alpha, 8beta-dihydroxy-9beta, 10beta-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene with DNA. Int J Cancer. 1976;18(3):362-368. Available from: http://onlinelibrary.wiley.com/store/10.1002/ijc.2910180315/asset/2910180315_ftp.pdf?v=1&t=isknbfih&s=3d8a9c8837aa2c8cc7e7b168bdd236fa98635959

107. Poland A, Glover E, Kende AS. Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that the binding species is

receptor for induction of aryl hydrocarbon hydroxylase. J Biol Chem. 1976;251(16):4936-4946. Available from:

108. Swanson HI, Chan WK, Bradfield CA. DNA Binding Specificities and Pairing Rules of the Ah Receptor, ARNT, and SIM Proteins. J Biol Chem. 1995;270(44):26292-26302. Available from: <Go to ISI>://WOS:A1995TC97800046

109. Tsuchiya Y, Nakajima M, Yokoi T. Cytochrome P450-mediated metabolism of estrogens and its regulation in human. Cancer Lett. 2005;227(2):115-124. doi:10.1016/j.canlet.2004.10.007

110. Berge G, Mollerup S, S OV, Hewer A, Phillips DH, Eilertsen E, et al. Role of estrogen receptor in regulation of polycyclic aromatic hydrocarbon metabolic activation in lung. Lung Cancer. 2004;45(3):289-297. doi:10.1016/j.lungcan.2004.02.014

111. Uppstad H, Osnes GH, Cole KJ, Phillips DH, Haugen A, Mollerup S. Sex differences in susceptibility to PAHs is an intrinsic property of human lung adenocarcinoma cells. Lung Cancer. 2011;71(3):264-270. doi:10.1016/j.lungcan.2010.09.006

112. Mollerup S, Berge G, Baera R, Skaug V, Hewer A, Phillips DH, et al. Sex differences in risk of lung cancer: Expression of genes in the PAH bioactivation pathway in relation to smoking and bulky DNA adducts. Int J Cancer. 2006;119(4):741-744. doi:10.1002/ijc.21891

113. Tsuchiya Y, Nakajima N, Kyo S, Kanaya T, Inoue M, Yokoi T. Human CYP1B1 is regulated by estradiol via estrogen receptor. Cancer Res. 2004;64(9):3119-3125. doi:10.1158/0008-5472.can-04-0166

114. Zhao ZL, Kosinska W, Khmelnitsky M, Cavalieri EL, Rogan EG, Chakravarti D, et al. Mutagenic activity of 4-hydroxyestradiol, but not 2-hydroxyestradiol, in BB rat2 embryonic cells, and the mutational spectrum of 4-hydroxyestradiol. Chem Res Toxicol. 2006;19(3):475-479. doi:10.1021/tx0502645

115. Dawling S, Roodi N, Mernaugh RL, Wang X, Parl FF. Catechol-O-methyltransferase (COMT)-mediated metabolism of catechol estrogens: comparison of wild-type and variant COMT isoforms. Cancer Res. 2001;61(18):6716-6722. Available from: http://cancerres.aacrjournals.org/content/61/18/6716

116. Hayes CL, Spink DC, Spink BC, Cao JQ, Walker NJ, Sutter TR. 17 beta-estradiol hydroxylation catalyzed by human cytochrome P450 1B1. Proc Natl Acad Sci U S A. 1996;93(18):9776-9781. Available from:

117. Lin P, Chang YC, Chen CH, Yang WJ, Cheng YH, Chang LW. A comparative study on the effects of 2,3,7,8,-tetrachlorodibenzo-p-dioxin polychlorinated biphenyl126 and estrogen in human bronchial epithelial cells. Toxicol Appl Pharmacol. 2004;195(1):83-91. doi:10.1016/j.taap.2003.11.001

118. Chang LW, Chang YC, Ho CC, Tsai MH, Lin P. Increase of carcinogenic risk via enhancement of cyclooxygenase-2 expression and hydroxyestradiol accumulation in human lung cells as a result of interaction between BaP and 17-beta estradiol. Carcinogenesis. 2007;28(7):1606-1612. doi:10.1093/carcin/bgm013

119. Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. Br J Cancer. 2005;93(7):825-833. doi:10.1038/sj.bjc.6602769

120. Lin H, Huang YS, Yan HH, Yang XN, Zhong WZ, Ye HW, et al. A family history of cancer and lung cancer risk in never-smokers: A clinic-based case-control study. Lung Cancer. 2015;89(2):94-98. doi:10.1016/j.lungcan.2015.05.017

121. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature. 2008;452(7187):638-642. doi:10.1038/nature06846

122. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet. 2008;40(5):616-622. doi:10.1038/ng.109

123. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature. 2008;452(7187):633-637. doi:10.1038/nature06885

124. Zou P, Gu A, Ji G, Zhao L, Zhao P, Lu A. The TERT rs2736100 polymorphism and cancer risk: a meta-analysis based on 25 case-control studies. BMC Cancer. 2012;12:7. doi:10.1186/1471-2407-12-7

125. Wei RR, Cao L, Pu HY, Wang HW, Zheng YL, Niu XM, et al. TERT Polymorphism rs2736100-C Is Associated with EGFR Mutation-Positive Non-Small Cell Lung Cancer. Clin Cancer Res. 2015;21(22):5173-5180. doi:10.1158/1078-0432.ccr-15-0009

126. McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet. 2008;40(12):1404-1406. doi:10.1038/ng.254

127. Bell DW, Gore I, Okimoto RA, Godin-Heymann N, Sordella R, Mulloy R, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. Nat Genet. 2005;37(12):1315-1316. doi:10.1038/ng1671

128. Brennan P, Hainaut P, Boffetta P. Genetics of lung-cancer susceptibility. The Lancet Oncology. 2011;12(4):399-408. doi:10.1016/S1470- 2045(10)70126-1

129. Benhamou S, Lee WJ, Alexandrie AK, Boffetta P, Bouchardy C, Butkiewicz D, et al. Meta- and pooled analyses of the effects of glutathione S-transferase M1 polymorphisms and smoking on lung cancer risk. Carcinogenesis. 2002;23(8):1343-1350. doi:10.1093/carcin/23.8.1343

130. Young RP, Hopkins RJ, Hay BA, Gamble GD. GSTM1 null genotype in COPD and lung cancer: evidence of a modifier or confounding effect? Appl Clin Genet. 2011;4:137-144. doi:10.2147/tacg.s21517

131. Lee MS, Su L, Christiani DC. Synergistic effects of NAT2 slow and GSTM1 null genotypes on carcinogen DNA damage in the lung. Cancer Epidemiol Biomarkers Prev. 2010;19(6):1492-1497. doi:10.1158/1055-9965.epi-09-1195

132. De Matteis S, Consonni D, Pesatori AC, Bergen AW, Bertazzi PA, Caporaso NE, et al. Are women who smoke at higher risk for lung cancer than men who smoke? Am J Epidemiol. 2013;177(7):601-612. doi:10.1093/aje/kws445

133. Papadopoulos A, Guida F, Leffondre K, Cenee S, Cyr D, Schmaus A, et al. Heavy smoking and lung cancer: are women at higher risk? Result of the ICARE study. Br J Cancer. 2014;110(5):1385-1391. doi:10.1038/bjc.2013.821

134. Wakelee HA, Chang ET, Gomez SL, Keegan TH, Feskanich D, Clarke CA, et al. Lung cancer incidence in never smokers. J Clin Oncol. 2007;25(5):472-478. doi:10.1200/jco.2006.07.2983

135. Brandenberger AW, Tee MK, Lee JY, Chao V, Jaffe RB. Tissue distribution of estrogen receptors alpha (ER-alpha) and beta (ER-beta) mRNA in the midgestational human fetus. J Clin Endocrinol Metab. 1997;82(10):3509-3512. doi:10.1210/jc.82.10.3509

136. Omoto Y, Kobayashi Y, Nishida K, Tsuchiya E, Eguchi H, Nakagawa K, et al. Expression, function, and clinical implications of the estrogen receptor beta in human lung cancers. Biochem Biophys Res Commun. 2001;285(2):340-347. doi:10.1006/bbrc.2001.5158

137. Mollerup S, Jorgensen K, Berge G, Haugen A. Expression of estrogen receptors alpha and beta in human lung tissue and cell lines. Lung Cancer. 2002;37(2):153-159. doi:10.1016/S0169-5002(02)00039-9

138. Stabile LP, Davis AL, Gubish CT, Hopkins TM, Luketich JD, Christie N, et al. Human non-small cell lung tumors and cells derived from normal lung express both estrogen receptor alpha and beta and show biological responses to estrogen. Cancer Res. 2002;62(7):2141-2150. Available from: http://cancerres.aacrjournals.org/content/62/7/2141

139. Hsu LH, Liu KJ, Tsai MF, Wu CR, Feng AC, Chu NM, et al. Estrogen adversely affects the prognosis of patients with lung adenocarcinoma. Cancer Sci. 2015;106(1):51-59. doi:10.1111/cas.12558

140. Skjefstad K, Grindstad T, Khanehkenari MR, Richardsen E, Donnem T, Kilvaer T, et al. Prognostic relevance of estrogen receptor alpha, beta and aromatase expression in non-small cell lung cancer. Steroids. 2016;113:5-13. doi:10.1016/j.steroids.2016.05.008

141. Navaratnam, Skliris G, Qing GF, Banerji S, Badiani K, Tu DS, et al. Differential Role of Estrogen Receptor Beta in Early Versus Metastatic Non-small Cell Lung Cancer. Horm Cancer. 2012;3(3):93-100. doi:10.1007/s12672-012-0105-y

142. Stabile LP, Dacic S, Land SR, Lenzner DE, Dhir R, Acquafondata M, et al. Combined analysis of estrogen receptor beta-1 and progesterone receptor expression identifies lung cancer patients with poor outcome. Clin Cancer Res. 2011;17(1):154-164. doi:10.1158/1078-0432.ccr-10-0992

143. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2014;15(3):178-196. doi:10.1038/nrm3758

144. Halaoui R, McCaffrey L. Rewiring cell polarity signaling in cancer. Oncogene. 2015;34(8):939-950. doi:10.1038/onc.2014.59

145. Wan H, Dingle S, Xu Y, Besnard V, Kaestner KH, Ang SL, et al. Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis. J Biol Chem. 2005;280(14):13809-13816. doi:10.1074/jbc.M414122200

146. Hazan RB, Qiao R, Keren R, Badano I, Suyama K. Cadherin switch in tumor progression. Ann N Y Acad Sci. 2004;1014:155-163. Available from:

147. Thiery JP, Acloque H, Huang RYJ, Nieto MA. Epithelial-Mesenchymal Transitions in Development and Disease. Cell. 2009;139(5):871-890. doi:10.1016/j.cell.2009.11.007

148. Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. Cell. 2016;166(1):21-45. doi:10.1016/j.cell.2016.06.028

149. Schliekelman MJ, Taguchi A, Zhu J, Dai XD, Rodriguez J, Celiktas M, et al. Molecular Portraits of Epithelial, Mesenchymal, and Hybrid States in Lung Adenocarcinoma and Their Relevance to Survival. Cancer Res. 2015;75(9):1789-1800. doi:10.1158/0008-5472.can-14-2535

150. Alexander NR, Tran NL, Rekapally H, Summers CE, Glackin C, Heimark RL. N-cadherin gene expression in prostate carcinoma is modulated by integrin-dependent nuclear translocation of Twist1. Cancer Res. 2006;66(7):3365-3369. doi:10.1158/0008-5472.can-05-3401

151. McDonald OG, Wu H, Timp W, Doi A, Feinberg AP. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nat Struct Mol Biol. 2011;18(8):867-874. doi:10.1038/nsmb.2084

152. Javaid S, Zhang JM, Anderssen E, Black JC, Wittner BS, Tajima K, et al. Dynamic Chromatin Modification Sustains Epithelial-Mesenchymal Transition following Inducible Expression of Snail-1. Cell Reports. 2013;5(6):1679-1689. doi:10.1016/j.celrep.2013.11.034

153. Wang JX, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, et al. Opposing LSD1 complexes function in developmental gene activation and repression programmes. Nature. 2007;446(7138):882-887. doi:10.1038/nature05671

154. Aghdassi A, Sendler M, Guenther A, Mayerle J, Behn CO, Heidecke CD, et al. Recruitment of histone deacetylases HDAC1 and HDAC2 by the transcriptional repressor ZEB1 downregulates E-cadherin expression in pancreatic cancer. Gut. 2012;61(3):439-448. doi:10.1136/gutjnl-2011-300060

155. Droufakou S, Deshmane V, Roylance R, Hanby A, Tomlinson I, Hart IR. Multiple ways of silencing E-cadherin gene expression in lobular carcinoma of the breast. Int J Cancer. 2001;92(3):404-408. doi:10.1002/ijc.1208

156. Kanai Y, Ushijima S, Hui AM, Ochiai A, Tsuda H, Sakamoto M, et al. The E-cadherin gene is silenced by CpG methylation in human hepatocellular carcinomas. Int J Cancer. 1997;71(3):355-359. Available from: <Go to ISI>://WOS:A1997WW91100008

157. Kim NH, Kim HS, Li XY, Lee I, Choi HS, Kang SE, et al. A p53/miRNA-34 axis regulates Snail1-dependent cancer cell epithelial-mesenchymal transition. J Cell Biol. 2011;195(3):417-433. doi:10.1083/jcb.201103097

158. Cottonham CL, Kaneko S, Xu L. miR-21 and miR-31 Converge on TIAM1 to Regulate Migration and Invasion of Colon Carcinoma Cells. J Biol Chem. 2010;285(46):35293-35302. doi:10.1074/jbc.M110.160069

159. Bracken CP, Gregory PA, Kolesnikoff N, Bert AG, Wang J, Shannon MF, et al. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. Cancer Res. 2008;68(19):7846-7854. doi:10.1158/0008-5472.can-08-1942

160. Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, Li CW, et al. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. Nat Cell Biol. 2011;13(3):317-323. doi:10.1038/ncb2173

161. Thomson S, Petti F, Sujka-Kwok I, Mercado P, Bean J, Monaghan M, et al. A systems view of epithelial-mesenchymal transition signaling states. Clin Exp Metastasis. 2011;28(2):137-155. doi:10.1007/s10585-010-9367-3

162. Kasai H, Allen JT, Mason RM, Kamimura T, Zhang Z. TGF-beta 1 induces human alveolar epithelial to mesenchymal cell transition (EMT). Respir Res. 2005;6:15. doi:10.1186/1465-9921-6-56

163. Brabletz T, Hlubek F, Spaderna S, Schmalhofer O, Hiendlmeyer E, Jung A, et al. Invasion and metastasis in colorectal cancer: Epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. Cells Tissues Organs. 2005;179(1-2):56-65. doi:10.1159/000084509

164. Brabletz T, Jung A, Reu S, Porzner M, Hlubek F, Kunz-Schughart LA, et al. Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. Proc Natl Acad Sci U S A. 2001;98(18):10356-10361. doi:10.1073/pnas.171610498

165. Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST, et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature. 2015;527(7579):472-476. doi:10.1038/nature15748

166. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature. 2015;527(7579):525-530. doi:10.1038/nature16064

167. Beerling E, Seinstra D, de Wit E, Kester L, van der Velden D, Maynard C, et al. Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity. Cell Reports. 2016;14(10):2281-2288. doi:10.1016/j.celrep.2016.02.034

168. Dietrich C, Kaina B. The aryl hydrocarbon receptor (AhR) in the regulation of cell-cell contact and tumor growth. Carcinogenesis. 2010;31(8):1319-1328. doi:10.1093/carcin/bgq028

169. Barhoover MA, Hall JM, Greenlee WF, Thomas RS. Aryl hydrocarbon receptor regulates cell cycle progression in human breast cancer cells via a functional interaction with cyclin-dependent kinase 4. Mol Pharmacol. 2010;77(2):195-201. doi:10.1124/mol.109.059675

170. Gao Z, Bu Y, Zhang G, Liu X, Wang X, Ding S, et al. Effect of TCDD on the fate of epithelial cells isolated from human fetal palatal shelves (hFPECs). Toxicol Appl Pharmacol. 2016;305:186-193. doi:10.1016/j.taap.2016.06.016

171. Ikuta T, Kawajiri K. Zinc finger transcription factor Slug is a novel target gene of aryl hydrocarbon receptor. Exp Cell Res. 2006;312(18):3585-3594. doi:10.1016/j.yexcr.2006.08.002

172. Rico-Leo EM, Alvarez-Barrientos A, Fernandez-Salguero PM. Dioxin Receptor Expression Inhibits Basal and Transforming Growth Factor beta-induced Epithelial-to-mesenchymal Transition. J Biol Chem. 2013;288(11):7841-7856. doi:10.1074/jbc.M112.425009

173. Suda K, Tomizawa K, Fujii M, Murakami H, Osada H, Maehara Y, et al. Epithelial to Mesenchymal Transition in an Epidermal Growth Factor Receptor-Mutant Lung Cancer Cell Line with Acquired Resistance to Erlotinib. J Thorac Oncol. 2011;6(7):1152-1161. doi:10.1097/JTO.0b013e318216ee52

174. Yoshida T, Song LX, Bai Y, Kinose F, Li JN, Ohaegbulam KC, et al. ZEB1 Mediates Acquired Resistance to the Epidermal Growth Factor Receptor-Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. PLoS One. 2016;11(1):22. doi:10.1371/journal.pone.0147344

175. Chen LM, Gibbons DL, Goswami S, Cortez MA, Ahn YH, Byers LA, et al. Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. Nature Communications. 2014;5:12. doi:10.1038/ncomms6241

176. Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, et al. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. EMBO J. 1998;17(1):244-254. doi:10.1093/emboj/17.1.244

177. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell. 2002;9(2):279-289. doi:10.1016/s1097-2765(02)00459-8

178. Gao N, Zhang J, Rao MA, Case TC, Mirosevich J, Wang Y, et al. The role of hepatocyte nuclear factor-3 alpha (Forkhead Box A1) and androgen receptor in transcriptional regulation of prostatic genes. Mol Endocrinol. 2003;17(8):1484-1507. doi:10.1210/me.2003-0020

179. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, et al. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell. 2005;122(1):33-43. doi:10.1016/j.cell.2005.05.008

180. Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, et al. The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Mol Cell. 2016;62(1):79-91. doi:10.1016/j.molcel.2016.03.001

181. Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature. 1993;362(6417):219-223. doi:10.1038/362219a0

182. Cerf C, Lippens G, Ramakrishnan V, Muyldermans S, Segers A, Wyns L, et al. Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: full assignment, tertiary structure, and comparison with the globular domain of histone H5. Biochemistry. 1994;33(37):11079-11086. Available from:

183. Yu X, Gupta A, Wang Y, Suzuki K, Mirosevich J, Orgebin-Crist MC, et al. Foxa1 and Foxa2 interact with the androgen receptor to regulate prostate and epididymal genes differentially. Ann N Y Acad Sci. 2005;1061:77-93. doi:10.1196/annals.1336.009

184. Li Z, Tuteja G, Schug J, Kaestner KH. Foxa1 and Foxa2 are essential for sexual dimorphism in liver cancer. Cell. 2012;148(1-2):72-83. doi:10.1016/j.cell.2011.11.026

185. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat Genet. 2011;43(1):27-33. doi:10.1038/ng.730

186. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature. 2009;462(7269):58-64. doi:10.1038/nature08497

187. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome 2015 [updated Jan 23; cited 2016 12/09/16]. 2015/01/24:[1260419]. Available from: http://www.proteinatlas.org/.

188. Besnard V, Wert SE, Hull WM, Whitsett JA. Immunohistochemical localization of Foxa1 and Foxa2 in mouse embryos and adult tissues. Gene Expr Patterns. 2004;5(2):193-208. doi:10.1016/j.modgep.2004.08.006

189. Lin L, Miller CT, Contreras JI, Prescott MS, Dagenais SL, Wu R, et al. The hepatocyte nuclear factor 3 alpha gene, HNF3alpha (FOXA1), on chromosome band 14q13 is amplified and overexpressed in esophageal and lung adenocarcinomas. Cancer Res. 2002;62(18):5273-5279. Available from: http://cancerres.aacrjournals.org/content/62/18/5273.long

190. Deutsch L, Wrage M, Koops S, Glatzel M, Uzunoglu FG, Kutup A, et al. Opposite roles of FOXA1 and NKX2-1 in lung cancer progression. Genes Chromosomes Cancer. 2012;51(6):618-629. doi:10.1002/gcc.21950

191. Song Y, Washington MK, Crawford HC. Loss of FOXA1/2 is essential for the epithelial-to-mesenchymal transition in pancreatic cancer. Cancer Res. 2010;70(5):2115-2125. doi:10.1158/0008-5472.can-09-2979

192. Mirosevich J, Gao N, Gupta A, Shappell SB, Jove R, Matusik RJ. Expression and role of Foxa proteins in prostate cancer. Prostate. 2006;66(10):1013-1028. doi:10.1002/pros.20299

193. Nucera C, Eeckhoute J, Finn S, Carroll JS, Ligon AH, Priolo C, et al. FOXA1 is a potential oncogene in anaplastic thyroid carcinoma. Clin Cancer Res. 2009;15(11):3680-3689. doi:10.1158/1078-0432.ccr-08-3155

194. Badve S, Turbin D, Thorat MA, Morimiya A, Nielsen TO, Perou CM, et al. FOXA1 expression in breast cancer--correlation with luminal subtype A and survival. Clin Cancer Res. 2007;13(15 Pt 1):4415-4421. doi:10.1158/1078-0432.ccr-07-0122

195. Mehta RJ, Jain RK, Leung S, Choo J, Nielsen T, Huntsman D, et al. FOXA1 is an independent prognostic marker for ER-positive breast cancer. Breast Cancer Res Treat. 2012;131(3):881-890. doi:10.1007/s10549-011-1482-6

196. Tang Y, Shu G, Yuan X, Jing N, Song J. FOXA2 functions as a suppressor of tumor metastasis by inhibition of epithelial-to-mesenchymal transition in human lung cancers. Cell Res. 2011;21(2):316-326. doi:10.1038/cr.2010.126

197. Basseres DS, D'Alo F, Yeap BY, Lowenberg EC, Gonzalez DA, Yasuda H, et al. Frequent downregulation of the transcription factor Foxa2 in lung cancer through epigenetic silencing. Lung Cancer. 2012;77(1):31-37. doi:10.1016/j.lungcan.2012.01.011

198. Zhu CP, Wang J, Shi B, Hu PF, Ning BF, Zhang Q, et al. The transcription factor FOXA2 suppresses gastric tumorigenesis in vitro and in vivo. Dig Dis Sci. 2015;60(1):109-117. doi:10.1007/s10620-014-3290-4

199. Li CM, Gocheva V, Oudin MJ, Bhutkar A, Wang SY, Date SR, et al. Foxa2 and Cdx2 cooperate with Nkx2-1 to inhibit lung adenocarcinoma metastasis. Genes Dev. 2015;29(17):1850-1862. doi:10.1101/gad.267393.115

200. Perez-Balaguer A, Ortiz-Martinez F, Garcia-Martinez A, Pomares-Navarro C, Lerma E, Peiro G. FOXA2 mRNA expression is associated with relapse in patients with Triple-Negative/Basal-like breast carcinoma. Breast Cancer Res Treat. 2015;153(2):465-474. doi:10.1007/s10549-015-3553-6

201. Bochkis IM, Schug J, Ye DZ, Kurinna S, Stratton SA, Barton MC, et al. Genome-wide location analysis reveals distinct transcriptional circuitry by paralogous regulators Foxa1 and Foxa2. PLoS Genetics. 2012;8(6):e1002770. doi:10.1371/journal.pgen.1002770

202. Nguyen TT, Cho K, Stratton SA, Barton MC. Transcription factor interactions and chromatin modifications associated with p53-mediated, developmental repression of the alpha-fetoprotein gene. Mol Cell Biol. 2005;25(6):2147-2157. doi:10.1128/mcb.25.6.2147-2157.2005

203. Xiang C, Wang J, Kou X, Chen X, Qin Z, Jiang Y, et al. Pulmonary expression of CYP2A13 and ABCB1 is regulated by FOXA2, and their genetic interaction is associated with lung cancer. The FASEB Journal. 2015;29(5):1986-1998. doi:10.1096/fj.14-264580

204. Wang H, Meyer CA, Fei T, Wang G, Zhang F, Liu XS. A systematic approach identifies FOXA1 as a key factor in the loss of epithelial traits during the epithelial-to-mesenchymal transition in lung cancer. BMC Genomics. 2013;14:680. doi:10.1186/1471-2164-14-680

205. Liu Y-N, Lee W-W, Wang C-Y, Chao T-H, Chen Y, Chen JH. Regulatory mechanisms controlling human E-cadherin gene expression. Oncogene. 2005;24(56):8277-8290. doi:10.1038/sj.onc.120899

206. Zhang Z, Yang C, Gao W, Chen T, Qian T, Hu J, et al. FOXA2 attenuates the epithelial to mesenchymal transition by regulating the transcription of E-cadherin and ZEB2 in human breast cancer. Cancer Lett. 2015;361(2):240-250. doi:10.1016/j.canlet.2015.03.008

207. Wang W, Yi M, Chen S, Li J, Li G, Yang J, et al. Significance of the NOR1-FOXA1/HDAC2-Slug regulatory network in epithelial-mesenchymal transition of tumor cells. Oncotarget. 2016;7(13):16745-16759. doi:10.18632/oncotarget.7778

208. Zhang X-L, Wang H-S, Liu N, Ge L-C. Bisphenol A stimulates the epithelial mesenchymal transition of estrogen negative breast cancer cells via FOXA1 signals. Arch Biochem Biophys. 2015;585:10-16. doi:10.1016/j.abb.2015.09.006

209. Lustig AJ. Crisis intervention: the role of telomerase. Proc Natl Acad Sci U S A. 1999;96(7):3339-3341. Available from:

210. Maser RS, DePinho RA. Connecting chromosomes, crisis, and cancer. Science. 2002;297(5581):565-569. doi:10.1126/science.297.5581.565

211. Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, et al. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J Natl Cancer Inst. 1973;51(5):1417-1423. Available from:

212. Geraghty RJ, Capes-Davis A, Davis JM, Downward J, Freshney RI, Knezevic I, et al. Guidelines for the use of cell lines in biomedical research. Br J Cancer. 2014;111(6):1021-1046. doi:10.1038/bjc.2014.166

213. Girardi AJ, Jensen FC, Koprowski H. SV40-induced transformation of human diploid cells: Crisis and recovery. J Cell Comp Physiol. 1965;65(1):69-83. doi:10.1002/jcp.1030650110

214. Reddel RR, Ke Y, Gerwin BI, McMenamin MG, Lechner JF, Su RT, et al. Transformation of Human Bronchial Epithelial Cells by Infection with SV40 or Adenovirus-12 SV40 Hybrid Virus, or Transfection via Strontium Phosphate Coprecipitation with a Plasmid Containing SV40 Early Region Genes. Cancer Res. 1988;48(7):1904-1909. Available from: http://cancerres.aacrjournals.org/content/48/7/1904.abstract

215. Wynford-Thomas D. Telomeres, p53 and cellular senescence. Oncol Res. 1996;8(10-11):387-398. Available from:

216. Carnero A, Blanco-Aparicio C, Kondoh H, Lleonart ME, Martinez-Leal JF, Mondello C, et al. Disruptive chemicals, senescence and immortality. Carcinogenesis. 2015;36 Suppl 1:S19-37. doi:10.1093/carcin/bgv029

217. Lundberg AS, Randell SH, Stewart SA, Elenbaas B, Hartwell KA, Brooks MW, et al. Immortalization and transformation of primary human airway epithelial cells by gene transfer. Oncogene. 2002;21(29):4577-4586. doi:10.1038/sj.onc.1205550

218. Sasai K, Sukezane T, Yanagita E, Nakagawa H, Hotta A, Itoh T, et al. Oncogene-Mediated Human Lung Epithelial Cell Transformation Produces Adenocarcinoma Phenotypes In Vivo. Cancer Res. 2011;71(7):2541-2549. doi:10.1158/0008-5472.can-10-2221

219. Ramirez RD, Sheridan S, Girard L, Sato M, Kim Y, Pollack J, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. Cancer Res. 2004;64(24):9027-9034. doi:10.1158/0008-5472.can-04-3703

220. Damiani LA, Yingling CM, Leng S, Romo PE, Nakamura J, Belinsky SA. Carcinogen-induced gene promoter hypermethylation is mediated by DNMT1 and causal for transformation of immortalized bronchial epithelial cells. Cancer Res. 2008;68(21):9005-9014. doi:10.1158/0008-5472.CAN-08-1276

221. Zhao P, Fu J, Yao B, Song Y, Mi L, Li Z, et al. In vitro malignant transformation of human bronchial epithelial cells induced by benzo(a)pyrene. Toxicology in vitro : an international journal published in association with BIBRA. 2012;26(2):362-368. doi:10.1016/j.tiv.2011.12.013

222. Rodrigues CF, Urbano AM, Matoso E, Carreira I, Almeida A, Santos P, et al. Human bronchial epithelial cells malignantly transformed by hexavalent chromium exhibit an aneuploid phenotype but no microsatellite instability. Mutat Res. 2009;670(1-2):42-52. doi:10.1016/j.mrfmmm.2009.07.004

223. Greenburg G, Hay ED. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. J Cell Biol. 1982;95(1):333-339. Available from:

224. Dasgupta P, Rizwani W, Pillai S, Kinkade R, Kovacs M, Rastogi S, et al. Nicotine induces cell proliferation, invasion and epithelial-mesenchymal transition in a variety of human cancer cell lines. Int J Cancer. 2009;124(1):36-45. doi:10.1002/ijc.23894

225. Thevenot PT, Saravia J, Jin N, Giaimo JD, Chustz RE, Mahne S, et al. Radical-containing ultrafine particulate matter initiates epithelial-to-mesenchymal transitions in airway epithelial cells. Am J Respir Cell Mol Biol. 2013;48(2):188-197. doi:10.1165/rcmb.2012-0052OC

226. Polimeni M, Gulino GR, Gazzano E, Kopecka J, Marucco A, Fenoglio I, et al. Multi-walled carbon nanotubes directly induce epithelial-mesenchymal transition in human bronchial epithelial cells via the TGF-beta-mediated Akt/GSK-3beta/SNAIL-1 signalling pathway. Part Fibre Toxicol. 2016;13(1):27. doi:10.1186/s12989-016-0138-4

227. Ke XS, Qu Y, Goldfinger N, Rostad K, Hovland R, Akslen LA, et al. Epithelial to mesenchymal transition of a primary prostate cell line with switches of cell adhesion modules but without malignant transformation. PLoS One. 2008;3(10):e3368. doi:10.1371/journal.pone.0003368

228. Ke XS, Li WC, Hovland R, Qu Y, Liu RH, McCormack E, et al. Reprogramming of cell junction modules during stepwise epithelial to mesenchymal transition and accumulation of malignant features in vitro in a prostate cell model. Exp Cell Res. 2011;317(2):234-247. doi:10.1016/j.yexcr.2010.10.009

229. Waldron HA. A brief history of scrotal cancer. Br J Ind Med. 1983;40(4):390-401. Available from: http://oem.bmj.com/content/40/4/390.full.pdf

230. Cook JW, Hewett C, Hieger I. Coal Tar Constituents and Cancer. Nature. 1932;130:926-926. doi:10.1038/130926a0

231. Sinha R, Kulldorff M, Gunter MJ, Strickland P, Rothman N. Dietary benzo[a]pyrene intake and risk of colorectal adenoma. Cancer Epidemiol Biomarkers Prev. 2005;14(8):2030-2034. doi:10.1158/1055-9965.epi-04-0854

232. Theriault G, Tremblay C, Cordier S, Gingras S. Bladder cancer in the aluminium industry. Lancet. 1984;1(8383):947-950. Available from:

233. Lee BM, Shim GA. Dietary exposure estimation of benzo[a]pyrene and cancer risk assessment. J Toxicol Environ Health A. 2007;70(15-16):1391-1394. doi:10.1080/15287390701434182

234. Zhou H, Calaf GM, Hei TK. Malignant transformation of human bronchial epithelial cells with the tobacco-specific nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. Int J Cancer. 2003;106(6):821-826. doi:10.1002/ijc.11319

235. Hu YC, Yang ZH, Zhong KJ, Niu LJ, Pan XJ, Wu DC, et al. Alteration of transcriptional profile in human bronchial epithelial cells induced by cigarette smoke condensate. Toxicol Lett. 2009;190(1):23-31. doi:10.1016/j.toxlet.2009.06.860

236. Siegfried JM, Gubish CT, Rothstein ME, Henry C, Stabile LP. Combining the multitargeted tyrosine kinase inhibitor vandetanib with the antiestrogen fulvestrant enhances its antitumor effect in non-small cell lung cancer. J Thorac Oncol. 2012;7(3):485-495. doi:10.1097/JTO.0b013e31824177ea

237. Han W, Pentecost BT, Pietropaolo RL, Fasco MJ, Spivack SD. Estrogen receptor alpha increases basal and cigarette smoke extract-induced expression of CYP1A1 and CYP1B1, but not GSTP1, in normal human bronchial epithelial cells. Mol Carcinog. 2005;44(3):202-211. doi:10.1002/mc.20128

238. Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, Weinberg RA. Loss of E-Cadherin Promotes Metastasis via Multiple Downstream Transcriptional Pathways. Cancer Res. 2008;68(10):3645-3654. doi:10.1158/0008-5472.can-07-2938

239. Vleminckx K, Vakaet L, Jr., Mareel M, Fiers W, van Roy F. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell. 1991;66(1):107-119. Available from:

240. Schaeffer D, Somarelli JA, Hanna G, Palmer GM, Garcia-Blanco MA. Cellular Migration and Invasion Uncoupled: Increased Migration Is Not an Inexorable Consequence of Epithelial-to-Mesenchymal Transition. Mol Cell Biol. 2014;34(18):3486-3499. doi:10.1128/mcb.00694-14

241. Batlle E, Sancho E, Francí C, Domínguez D, Monfar M, Baulida J, et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol. 2000;2(2):84-89. doi:10.1038/35000034

242. Eger A, Aigner K, Sonderegger S, Dampier B, Oehler S, Schreiber M, et al. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene. 2005;24(14):2375-2385. doi:10.1038/sj.onc.1208429

243. Conacci-Sorrell M, Simcha I, Ben-Yedidia T, Blechman J, Savagner P, Ben-Ze'ev A. Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of beta-catenin signaling, Slug, and MAPK. J Cell Biol. 2003;163(4):847-857. doi:10.1083/jcb.200308162

244. Lombaerts M, van Wezel T, Philippo K, Dierssen JW, Zimmerman RM, Oosting J, et al. E-cadherin transcriptional downregulation by promoter methylation but not

mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br J Cancer. 2006;94(5):661-671. doi:10.1038/sj.bjc.6602996

245. Ma L, Young J, Prabhala H, Pan E, Mestdagh P, Muth D, et al. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nat Cell Biol. 2010;12(3):247-256. doi:10.1038/ncb2024

246. Liu YN, Abou-Kheir W, Yin JJ, Fang L, Hynes P, Casey O, et al. Critical and reciprocal regulation of KLF4 and SLUG in transforming growth factor beta-initiated prostate cancer epithelial-mesenchymal transition. Mol Cell Biol. 2012;32(5):941-953. doi:10.1128/mcb.06306-11

247. Jin HJ, Zhao JC, Ogden I, Bergan RC, Yu J. Androgen receptor-independent function of FoxA1 in prostate cancer metastasis. Cancer Res. 2013;73(12):3725-3736. doi:10.1158/0008-5472.can-12-3468

248. Zhu CQ, Cutz JC, Liu N, Lau D, Shepherd FA, Squire JA, et al. Amplification of telomerase (hTERT) gene is a poor prognostic marker in non-small-cell lung cancer. Br J Cancer. 2006;94(10):1452-1459. doi:10.1038/sj.bjc.6603110

249. Cao Y, Bryan TM, Reddel RR. Increased copy number of the TERT and TERC telomerase subunit genes in cancer cells. Cancer Sci. 2008;99(6):1092-1099. doi:10.1111/j.1349-7006.2008.00815.x

250. Wu A, Wu B, Guo J, Luo W, Wu D, Yang H, et al. Elevated expression of CDK4 in lung cancer. J Transl Med. 2011;9:38. doi:10.1186/1479-5876-9-38

251. Wang R, Shi Y, Chen L, Jiang Y, Mao C, Yan B, et al. The ratio of FoxA1 to FoxA2 in lung adenocarcinoma is regulated by LncRNA HOTAIR and chromatin remodeling factor LSH. Sci Rep. 2015;5:17826. doi:10.1038/srep17826

252. Li X, Chen S, Sun T, Xu Y, Chen Y, Liu Y, et al. The transcriptional regulation of SOX2 on FOXA1 gene and its application in diagnosis of human breast and lung cancers. Clin Lab. 2014;60(6):909-918. Available from:

253. Ivanova MM, Mazhawidza W, Dougherty SM, Klinge CM. Sex differences in estrogen receptor subcellular location and activity in lung adenocarcinoma cells. Am J Respir Cell Mol Biol. 2010;42(3):320-330. doi:10.1165/rcmb.2009-0059OC

254. Dougherty SM, Mazhawidza W, Bohn AR, Robinson KA, Mattingly KA, Blankenship KA, et al. Gender difference in the activity but not expression of estrogen receptors alpha and beta in human lung adenocarcinoma cells. Endocr Relat Cancer. 2006;13(1):113-134. doi:10.1677/erc.1.01118

255. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science. 2015;348(6237):880-886. doi:10.1126/science.aaa6806

256. Barrero MJ, Sese B, Kuebler B, Bilic J, Boue S, Marti M, et al. Macrohistone variants preserve cell identity by preventing the gain of H3K4me2 during reprogramming to pluripotency. Cell Rep. 2013;3(4):1005-1011. doi:10.1016/j.celrep.2013.02.029

257. Lavigne AC, Castells M, Mermet J, Kocanova S, Dalvai M, Bystricky K. Increased macroH2A1.1 expression correlates with poor survival of triple-negative breast cancer patients. PLoS One. 2014;9(6):e98930. doi:10.1371/journal.pone.0098930

258. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. 2010;328(5980):916-919. doi:10.1126/science.1186366

259. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007;129(4):823-837. doi:10.1016/j.cell.2007.05.009

260. Li Z, Gadue P, Chen K, Jiao Y, Tuteja G, Schug J, et al. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell. 2012;151(7):1608-1616. doi:10.1016/j.cell.2012.11.018

I

Toxicology in Vitro 35 (2016) 55–65

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

Epithelial-mesenchymal transition and FOXA genes during tobacco smoke carcinogen induced transformation of human bronchial epithelial cells☆

Audun Bersaas, Yke Jildouw Arnoldussen, Mari Sjøberg, Aage Haugen, Steen Mollerup *

Section for Toxicology and Biological Working Environment, Department of Biological and Chemical Working Environment, National Institute of Occupational Health, N-0033 Oslo, Norway

ARTICLE INFO

Article history:
Received 23 February 2016
Received in revised form 18 April 2016
Accepted 19 April 2016
Available online 21 May 2016

Keywords:
In vitro transformation
Lung carcinogenesis
EMT
Gene expression
Cigarette smoke condensate
Benzo[*a*]pyrene

ABSTRACT

Lung cancer is largely an environmentally caused disease with poor prognosis. An *in vitro* transformation model of human bronchial epithelial cells (HBEC) was used to study long-term effects of tobacco smoke carcinogens on epithelial-mesenchymal transition (EMT) and the forkhead box transcription factors FOXA1 and FOXA2. *CDK4* and *hTERT* immortalized HBEC2 and HBEC12 cell lines were exposed weekly to either cigarette smoke condensate (CSC), benzo[*a*]pyrene, or methylnitrosourea. Transformed cell lines were established from soft-agar colonies after 12 weeks of exposure. HBEC12 was transformed by all exposures while HBEC2 was only transformed by CSC. Untransformed HBEC2 showed little invasive capacity, whereas transformed cell lines completely closed the gap in a matrigel scratch wound assay. *CDH1* was down-regulated in all of the transformed cell lines. In contrast, *CDH2* was up-regulated in both HBEC2 and one of the HBEC12 transformed cell lines. Furthermore, transformed cells showed activation of EMT markers including *SNAI1*, *ZEB1*, *VIM*, and *MMP2*. All transformed cell lines had significant down-regulation of *FOXA1* and *FOXA2*, indicating a possible role in cell transformation and EMT. ChIP analysis showed increased binding of Histone-H3 and macroH2A in *FOXA1* and *FOXA2* in the transformed HBEC2 cell lines, indicating a compact chromatin. In conclusion, long-term carcinogen exposure lead to down-regulation of *FOXA1* and *FOXA2* concomitantly with the occurrence of EMT and *in vitro* transformation in HBEC cells.

1. Introduction

Lung cancer is the most common cause of cancer death worldwide (Ferlay et al., 2015). The advances in the treatment of lung cancer have been slow with current 5-year relative survival rates below 20% (Siegel et al., 2015). The single most important etiological factor for lung cancer is tobacco smoke but also occupational and environmental carcinogens contribute significantly to the disease, making it largely an environmentally induced cancer (Dela Cruz et al., 2011). In Norway, it is estimated that approximately 20% of all lung cancer cases may be a result of occupational exposure (Haldorsen et al., 2004).

The development of human lung carcinoma is a long-term process typically spanning several decades, resulting from chronic exposure to environmental carcinogens. Cigarette smoke is composed of a complex chemical mixture containing around 5700 different compounds of which 76 are classified as carcinogenic in laboratory animals or humans (Hecht, 2012; Rodgman and Perfetti, 2013). Based on their potency and levels, polycyclic aromatic hydrocarbons (PAH), *N*-nitrosamines, aromatic amines, aldehydes, ethylene epoxides, 1,3-butadiene, and benzene are considered the most important carcinogens in cigarette smoke (Hecht, 2006).

Epithelial-to-mesenchymal transition (EMT) is a complex reprogramming process providing epithelial cells with a mesenchymal phenotype. EMT takes place *via* several different pathways and plays an essential role during organogenesis, tissue repair and fibrosis, in addition to tumor invasion and metastasis (Thiery et al., 2009). In cancer development, transcriptional changes occurring during EMT may involve the loss of epithelial markers (E-cadherin, α- and ϒ-catenin) and the gain of mesenchymal cell markers (N-cadherin, vimentin, and

Abbreviations: B[*a*]P, benzo[*a*]pyrene; ChIP, chromatin immunoprecipitation; CSC, cigarette smoke condensate; EMT, epithelial-mesenchymal transition; HBEC, human bronchial epithelial cells; MNU, *N*-Nitroso-*N*-methylurea; NSCLC, non-small cell lung carcinoma; PAH, polycyclic aromatic hydrocarbon; qPCR, quantitative PCR; TF, transcription factor; TSS, transcription start site.

☆ Grant support: This study was supported by grants from the Norwegian Cancer Society.

* Corresponding author at: Section for Toxicology and Biological Working Environment, Department of Biological and Chemical Working Environment, National Institute of Occupational Health, PO box 8149, Dep., Gydas vei 8, N-0033 Oslo, Norway.
E-mail address: steen.mollerup@stami.no (S. Mollerup).

http://dx.doi.org/10.1016/j.tiv.2016.04.012

fibronectin) (Thiery, 2002). The "cadherin switch", which involves down-regulation of E-cadherin (*CDH1*) and up-regulation of N-cadherin (*CDH2*), is a typical EMT marker in cancer. The resulting mesenchymal cells display a more stem cell like phenotype with increased migratory capacity and increased invasiveness (Morel et al., 2012).

Several studies have reported reduced levels of E-cadherin in lung tumors (Bremnes et al., 2002; Nagathihalli et al., 2012). This may be caused by direct or indirect repression of transcription by several transcription factors (TFs). Two zinc finger TF families, Snail (SNAI1 and SNAI2) and ZEB (ZEB1 and ZEB2), repress *CDH1* gene expression by directly binding to E-box elements in the *CDH1* promoter. Other transcription factors, such as the basic helix-loop-helix TF TWIST, down-regulate E-cadherin levels by binding E-boxes in *CDH1* or in other repressors of *CDH1* (Pozharskaya et al., 2009; Vesuna et al., 2008). Epigenetic regulation of the *CDH1* gene may also be important and *CDH1* promoter hypermethylation represents a prognostic factor for non-small cell lung cancer (NSCLC) (Yoshiura et al., 1995). TWIST actively induces transcription of *CDH2* (Alexander et al., 2006), making it a critical player in establishing the cadherin switch. The expression of certain matrix metalloproteinases (including MMP2 and MMP9) are increased during EMT and cell invasion. These proteins proteolytically degrade components in the extracellular matrix, rendering the cell able to migrate and invade (Ura et al., 1989).

FOXA1 and FOXA2 are pioneer TFs that bind to and open condensed chromatin, making response elements in regulatory regions available for binding of other TFs. FOXA1 and FOXA2 are considered to play an essential role in maintaining the epithelial phenotype, and loss of *FOXA1* and *FOXA2* expression has been reported as essential for EMT in some tissues, including lung (Song et al., 2010; Tang et al., 2011; Wang et al., 2013). Moreover, FOXA1 expression has been associated with metastasis and low survival in lung squamous cell carcinoma (Deutsch et al., 2012). The chromatin "opening" action of FOXA proteins provides the foundation for hormone receptor interactions with DNA, and their deregulation may result in reprogramming of gene regulatory actions (Carroll et al., 2005). Interestingly, whereas *FOXA1* may be up-regulated by gene amplification, *FOXA2* has been found to be down-regulated in association with promoter hypermethylation in lung cancers (Basseres et al., 2012; Lin et al., 2002). The role of FOXA1/2 in cancer is not completely understood and may be tissue dependent. For instance, opposing roles of FOXA1 are observed during cancer development in the prostate and breast (Robinson and Carroll, 2012).

A large number of studies have characterized mechanisms involved in lung carcinogenesis *in vitro* (Gazdar et al., 2010). In the past, many of these have relied on already tumorigenic lung cell lines or bronchial epithelial cell lines immortalized with viral oncogenes (*e.g.* A549, BEAS-2B, and others). Such cell lines often show mutated TP53 and/or impaired TP53 signaling, which makes them less suited for studies of early steps during lung carcinogenesis. In the present study we used two human bronchial epithelial cell lines immortalized in the absence of viral oncoproteins (HBEC2-KT and HBEC12-KT). The HBECs cluster with primary lung epithelial cells in global gene expression analysis, show UV-inducible, wild type TP53, a high degree of chromosomal stability, have retained the ability to differentiate and may represent a relevant model for *in vitro* lung epithelial carcinogenesis studies (Delgado et al., 2011; Ramirez et al., 2004).

Little is known about the role of FOXA1 and FOXA2 during chemical carcinogenesis in human lung cells. Here, *in vitro* transformation of human bronchial epithelial cell lines (HBECs) was studied after long-term exposure to a cigarette smoke condensate (CSC), the model PAH benzo[*a*]pyrene (B[*a*]P), and methylnitrosourea (MNU). Transformed cell lines show characteristics of loss of epithelial and gain of mesenchymal morphology and phenotype. Concomitant down-regulation of the FOXA1 and FOXA2 factors in both pre-transformed and transformed cells was demonstrated by gene expression analysis and confocal immunofluorescence microscopy. Reduced expression of *FOXA* genes may be due to epigenetic regulation indicated by altered histone marks in regulatory regions of the genes.

2. Materials and methods

2.1. Cell culture

The *hTERT* and *Cdk4* immortalized human bronchial epithelial cell lines HBEC2-KT (male donor, 68 y.o., NSCLC, smoker) and HBEC12-KT (female donor, 55 y.o., NSCLC, ex-smoker) were a kind gift from Dr. John D. Minna (Ramirez et al., 2004). The HBEC lines have recently been authenticated by the Leibniz-Institut DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). Cells were grown in LHC-9 media (Thermo Fisher Sci.) supplemented with FBS (10%). For transformation cells were seeded in 6-well multidishes with a seeding density of 1×10^4 cells per well (HBEC2) and 2×10^4 cells per well (HBEC12). Seeding density was determined in order to get confluent wells after 1 week of incubation. The cells were incubated in a humidified atmosphere at 37 °C, 5% CO_2.

2.2. Cell transformation

Cytotoxicity was assessed using CellTiter-Blue cell viability assay (Promega). Cells were passaged (day 1) and seeded in quadruplicates in 6-well multidishes. During the whole experiment, cells from each of these wells were kept separate. Twenty-four hours after passaging, the cells were exposed to B[*a*]P (high dose: 1.0 μM, low dose: 0.3 μM) or CSC (high dose: 3.0 μg/μl, low dose: 1.0 μg/μl). Due to different toxicity of MNU between the cell lines 2KT and 12KT, the doses used were 1.0 μM and 0.5 μM, respectively. Vehicle control was DMSO (0.1%). After 72 h of exposure the culture media were removed, the cells were washed with PBS, and fresh media were added. The cells were then incubated for 72 h before being passaged again. This seven days exposure routine was repeated for a total duration of 12 weeks (outline of experimental design in Supplementary Fig. S1). After the 6th, 9th, and 12th exposure, a fraction of cells from each treatment was assayed for growth in soft-agar. In addition, cells were seeded for gene expression analysis of pre-transformed cells.

2.3. Selection in soft-agar

The soft-agar assay was performed in a 6-well multidish. The bottom and top layer had an agarose concentration of 0.7% (1.5 ml per well) and 0.35% (1.0 ml per well), respectively. Seeding density was 1×10^4 cells per well. Quadruplicates were seeded for each of the four replicates for each treatment (total of 16 wells soft-agar for each treatment). Soft-agar cultures were incubated for 3–4 weeks before colonies were detected. Colonies were then isolated using a micropipette and transferred to a 24-well multidish. Cells were incubated until confluency and then passaged to 100 mm petri dishes. Soft agar assay was carried out twice to ensure stable transformation and true clonality of the isolated cell lines. The cell lines established from soft agar colonies were cryopreserved. Transformed cell lines established from soft-agar colonies after 12 weeks of carcinogen exposure were used for further studies.

2.4. Cell migration and invasion assay

Migration and invasion studies were performed using IncuCyte Zoom Live Cell Imaging microscope and software (Essen BioScience). Cells were seeded on matrigel (BD Biosciences, 356230) coated ImageLock 96-well plates (2.5×10^4 cells/well). After 24 h, a scratch wound was made in the confluent cell layer using the WoundMaker toll (Essen BioScience). For migration, the plates were readily placed within an IncuCyte Zoom microscope inside the incubator and images were acquired every hour for up to 72 h to monitor closure of the wound. For assessment of invasion, the scratch wounded cell layer

was overlaid with matrigel before returning the plates to the incubator and images were acquired in the same manner.

2.5. Gene expression analysis

Total RNA was isolated from cultured cells by the Izol-RNA Lysis Reagent (5 PRIME). Reverse transcription was done using qScript cDNA synthesis kit (Quanta Biosciences). Analysis of relative gene expression was performed by quantitative PCR (qPCR) using the ABI Prism 7900HT (Life Technologies) and PerfeCTa SYBR green fast mix (Quanta Biosciences). All gene expression levels are normalized to β-actin levels and are calculated by the $\Delta\Delta C_q$ method (Livak and Schmittgen, 2001).

2.6. Immunofluorescence confocal microscopy

HBEC2 and HBEC12 control and transformed cells were grown on cover slips until confluent. Cells were fixed for 20 min in 4% paraformaldehyde and permeabilized with 0.1% Triton X-100 in PBS for 5 min. Cover slips were blocked by incubation with 5% BSA in 0.1% PBS-Triton X-100 for 1 h at room temperature and incubated overnight with primary antibodies against E-Cadherin (ab40772, Abcam), N-Cadherin (ab98952), FOXA1 (ab23738), or FOXA2 (sc6554, Santa Cruz Biotechnology) in 3% BSA in PBS, at 4 °C in a humidified chamber. Secondary antibody Alexa fluor 488-linked goat anti-rabbit IgG (A11070, Life Technologies), Alexa fluor 647-linked goat anti-mouse IgG (A323) or Alexa fluor 488-linked donkey anti-goat IgG (A11055), (all from Molecular Probes, Life Technologies) were left on the cells for 1 h at room temperature in a humidified chamber. To visualize cell nuclei cells were counterstained with Hoechst (Sigma-Aldrich). Fluorescence signals were monitored using a pinhole confocal microscope (Zeiss Oberkochen, Germany) and images were acquired with an AxioCam camera (Zeiss).

2.7. Chromatin immunoprecipitation (ChIP)

Chromatin cross-linking and nuclei isolation was performed using truChIP chromatin shearing kit (Covaris) according to the manufacturer's protocol. Briefly, chromatin was cross-linked for two minutes by adding formaldehyde (1%) to the cell culture media. The cross-linking was subsequently quenched for 5 min. The cell pellet was collected by scraping, then washed and lysed. Chromatin was sheared using Covaris ultrasonicator M220. The IP was performed as described by Cortazár et al., with some modifications (Cortazar et al., 2011). Briefly, 25 μg chromatin was used in each ChIP with 2 μg of the appropriate antibody against H3 (Abcam, Ab1791), H2A.Z (Abcam, Ab4174), macroH2A (Abcam, Ab37264), or negative control rabbit immunoglobulins (DacoCytomation, X-0903). For each cell line, chromatin from four independent experiments was pooled. Negative control ChIPs were conducted in parallel with every experiment. ChIP DNA-eluates were assayed by qPCR with primers targeting 6 regions in *FOXA1* and 4 regions in *FOXA2* (Supplementary Table 1). For all qPCR assays eluates from the negative control (mock IgG) were included. C_q values for each sample was normalized to the C_q of an non-immunoprecipitated (IP input) sample and all results are presented as per cent of input.

2.8. Statistical analysis

Calculation of mean and standard deviation of qPCR results was done using the Pandas package in Python. Gene expression plots were created in Mathematica (Wolfram). All gene expressions are expressed as mean values of fold change ± standard deviations (SD). Hypothesis testing was done by two-way ANOVA with Dunnett's multiple comparison test. Significance is indicated as $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$ in the figures. Cell migration and invasion was analyzed by a linear mixed model (Stata, StataCorp LP). An interaction term between treatment and time allowed us to estimate the difference over time between the transformed cell lines and the control cell line. To take into account the dependency in the data, random effects were added for wells, while a random slope for time accounted for different individual slopes of the curve for each well. Finally, an autoregressive (AR) structure of order 2 was assumed for the residuals.

3. Results

3.1. Transformation of HBEC2 and HBEC12

To gain insight into how tobacco related carcinogens transform human bronchial epithelial cells we used two HBEC lines immortalized with the *CDK4* and *hTERT* genes (HBEC2-KT and HBEC12-KT) (Ramirez et al., 2004). To test if exposure to CSC was sufficient to achieve transformation, HBEC lines were exposed for 12 weeks and then seeded in soft agar. In parallel, both cell lines were exposed to the pre-carcinogenic compound B[*a*]P or the direct acting carcinogen MNU. The toxicity of MNU was different between the two cell lines, thus different doses were used. HBEC12 was the most readily transformed cell line, with colonies being observed after all treatments (Table 1). In contrast, HBEC2 was only transformed after exposure to CSC. Transformed cells were subjected to a second round of soft agar colony formation to ensure stable transformation and true clonality. In total, 28 independent transformed cell lines were established (four clones from each exposure; termed A–D; Table 1). Neither HBEC2 nor HBEC12 treated with vehicle for 12 weeks showed signs of colony formation in soft agar.

3.2. Transformed cells display characteristics of EMT

The mechanisms controlling the early stages of carcinogen induced cell transformation are largely unknown. One suggested mechanism is that cells undergo EMT to gain invasiveness and metastatic capacity. While the vehicle treated cells in the present study retained an epithelial morphology after 12 weeks of culturing, the transformed cells displayed a morphology resembling that of mesenchymal cells (Supplementary Fig. S2). To investigate whether the transformed cell lines had transitioned from an epithelial to a mesenchymal cell type, the expression levels of several genes known to be involved in EMT were analyzed.

Table 1
Transformation of HBEC cells after 12 weeks of exposure to tobacco smoke carcinogens.

	DMSO	B[*a*]P (0.3 μM)	B[*a*]P (1.0 μM)	CSC (1 μg/ml)	CSC (3 μg/ml)	MNU
HBEC2	Neg.	Neg.	Neg.	+	+	(1 mM) Neg.
Names of transformed cell lines				T2KT-CSC-L (A–D)	T2KT-CSC-H (A–D)	
HBEC12	Neg.	+	+	+	+	(0.5 mM) +
Names of transformed cell lines		T12KT-B[*a*]P-L (A–D)	T12KT-B[*a*]P-H (A–D)	T12KT-CSC-L (A–D)	T12KT-CSC-H (A–D)	T12KT-MNU (A–D)

"+" indicates transformation, while neg. indicates that no colonies were detected after the particular exposure.

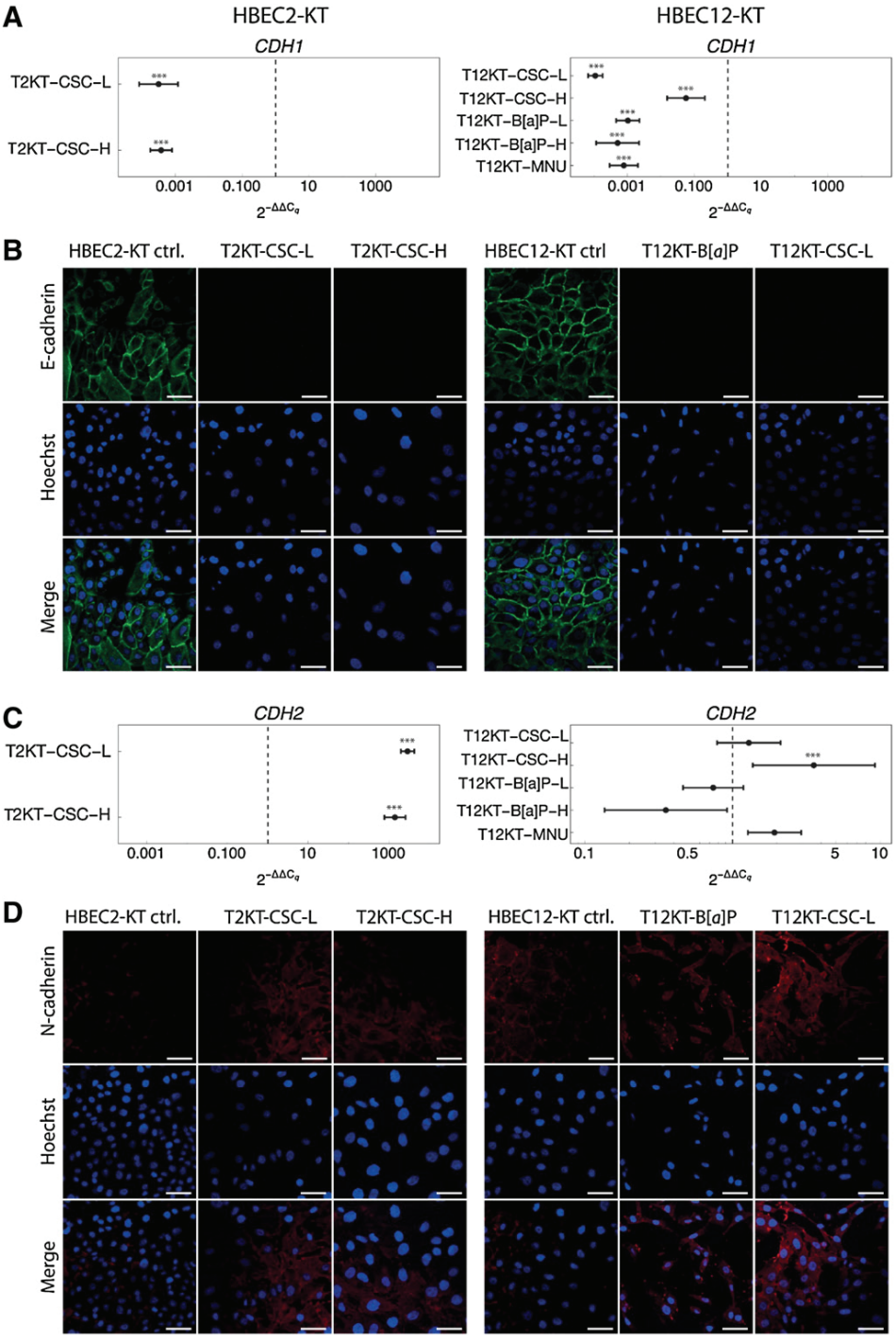

Fig. 1. Cadherin switch event in transformed HBECs. Cadherin expression was measured in untransformed and transformed HBEC2 and HBEC12 cell lines at the gene and protein level. (A) and (C) Expression of *CDH1* and *CDH2* mRNA in transformed HBEC2 and HBEC12 cells, respectively. Relative gene expression levels as compared to the particular untransformed control cells are presented (fold change; $2^{-\Delta\Delta Cq}$). (B) and (D) Representative images from confocal immunofluorescence microscopy of HBEC2 and HBEC12 control and transformed cells stained for E-cadherin (B) and N-cadherin (D), respectively. Nuclear/DNA staining was done by Hoechst. Scale bar: 50 μm. T2KT: transformed HBEC2, T12KT: Transformed HBEC12, CSC-L: cigarette smoke condensate 1.0 μg/μl, CSC-H: CSC 3.0 μg/μl, B[*a*]P-L: Benzo[*a*]pyrene 0.3 μM, B[*a*]P-H: B[*a*]P1.0 μM, MNU: Methylnitrosourea 0.5 mM. Gene expression data display mean ± SD, $n = 4$, where each replicate represents clones isolated from individual colonies in soft-agar. ***$p < 0.001$ (ANOVA Dunnet's posttest).

Statistically significant reduced levels of *CDH1* were observed in transformants of both HBEC2 and HBEC12, compared to untransformed control cells (Fig. 1A). The transformed cell lines originating from HBEC2 (CSC low and high dose) showed significantly reduced expression of *CDH1* ($>2 \times 10^3$ times). Similarly, in HBEC12 transformed clones exposed to either low or high dose of CSC, B[*a*]P, or single dose MNU, *CDH1* was significantly down-regulated between 10 and 10^4 times, compared to the untransformed controls. In contrast, transformed cells derived from HBEC2 showed a significant increase in *CDH2* expression ($>10^3$ times) (Fig. 1C). In transformed cells derived from HBEC12, only those arising from the high dose of CSC had significantly increased

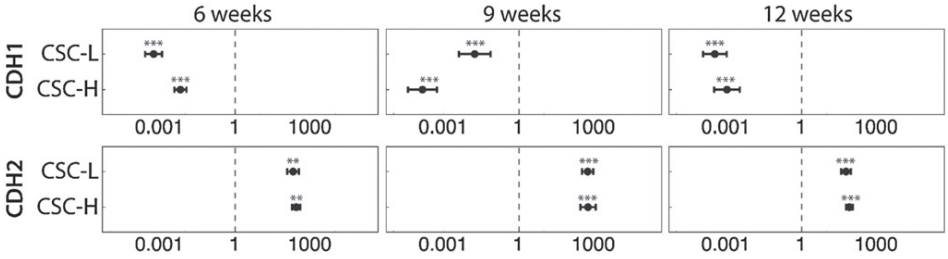

Fig. 2. Timing of the cadherin switch. Expression levels of *CDH1* and *CDH2* were measured prior to selection for transformed cells in soft agar in HBEC2 cells exposed to low and high dose CSC for 6, 9, and 12 weeks. Relative gene expression levels as compared to the particular control cells (DMSO) are presented (fold change; mean ± SD, $n = 4$). $^{**}p < 0.01$, and $^{***}p < 0.001$, (ANOVA Dunnet's posttest).

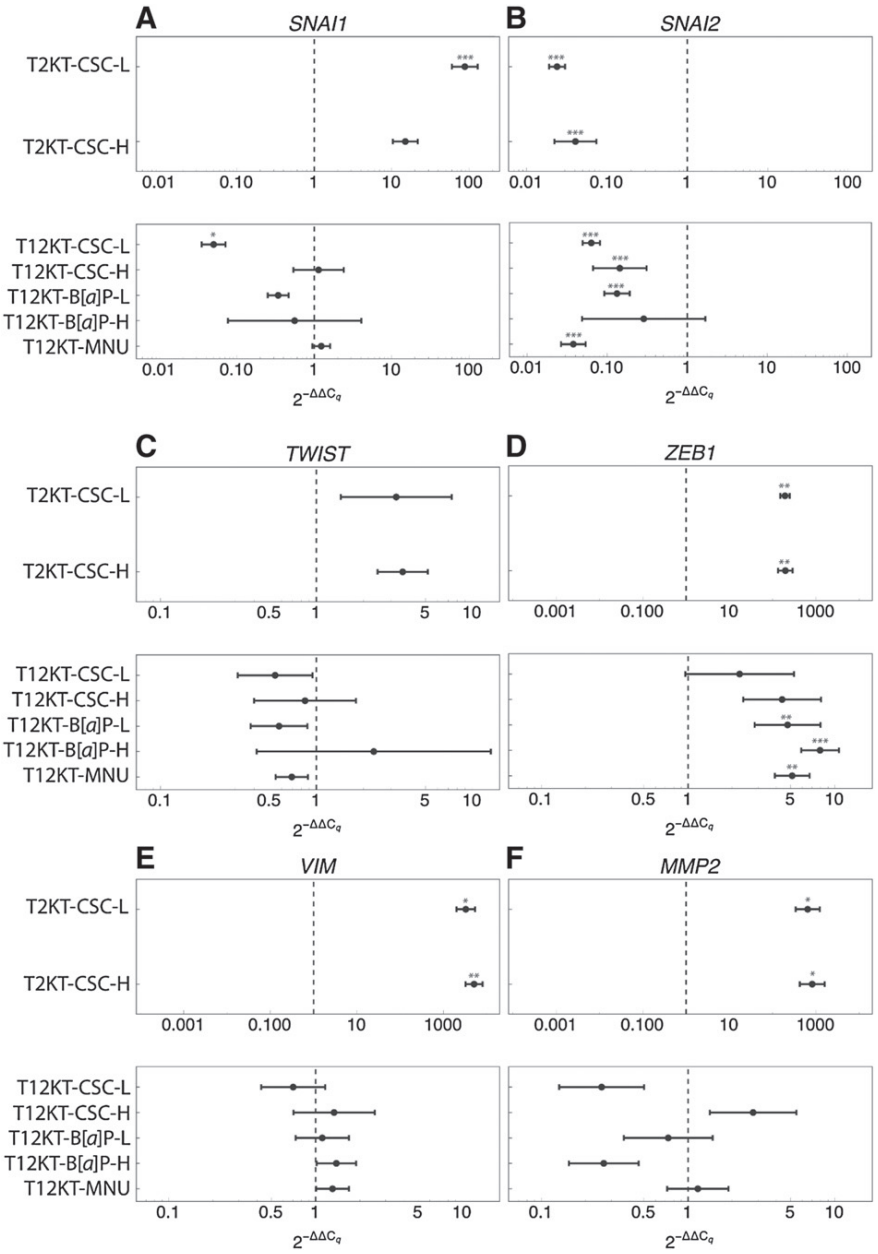

Fig. 3. Expression of EMT marker genes in transformed HBEC lines. The expression of the EMT marker genes (A) *SNAI1*, (B) *SNAI2*, (C) *TWIST*, (D) *ZEB1*, (E) *VIM*, and (F) *MMP2* was measured in untransformed and transformed HBEC2 and HBEC12 cell lines. Relative gene expression levels as compared to the particular control cells (DMSO) are presented (fold change; mean ± SD, $n = 4$). $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$, (ANOVA Dunnet's posttest).

CDH2 gene expression. All other HBEC12 transformants showed relatively unaltered expression of *CDH2*.

Cadherin deregulation in HBEC2 transformants was also evident at the protein level, when analyzed by fluorescence microscopy (Fig. 1B). In both untransformed HBEC2 and HBEC12, strong staining of E-cadherin was found in the membranes, especially of adherent cells indicating its being organized in junctional adhesion. In contrast, E-cadherin protein expression was abolished in transformed cell lines of both HBEC2 and HBEC12, as exemplified by B[*a*]P and CSC low dose, respectively. N-cadherin showed little staining in untransformed cells but was increased in transformed 2KT (T2KT-CSC-L and T2KT-CSC-H) compared to the controls (Fig. 1D). For T12KTs, a slight increase in the fluorescent signal in the T12-KT-CSC-L cells was apparent, despite a non-significant change in mRNA levels.

To ascertain the occurrence of the cadherin switch, expression of *CDH1* and *CDH2* was measured in CSC exposed HBEC2 at different time points prior to testing for soft agar colony formation. For short term exposure to CSC (1 or 3 weeks) no changes in the expression of *CDH1* and *CDH2* were apparent (data not shown). However, down-regulation of *CDH1* and up-regulation of *CDH2* was observed after 6, 9, and 12 weeks (Fig. 2). Thus, these data indicate that the timing of the cadherin switch is in line with or precedes transformation of the HBEC2 cell line.

Further evidence of EMT during *in vitro* cell transformation was obtained from gene expression analyses of typical EMT markers (Fig. 3). Expression of the EMT driver *SNAI1* was significantly up-regulated in T2KT-CSC-L (15-fold; Fig. 3A). In T2KT-CSCH, a non-significant increase of *SNAI1* mRNA was observed. In contrast, *SNAI2* was down-regulated in transformed HBEC2 as compared to

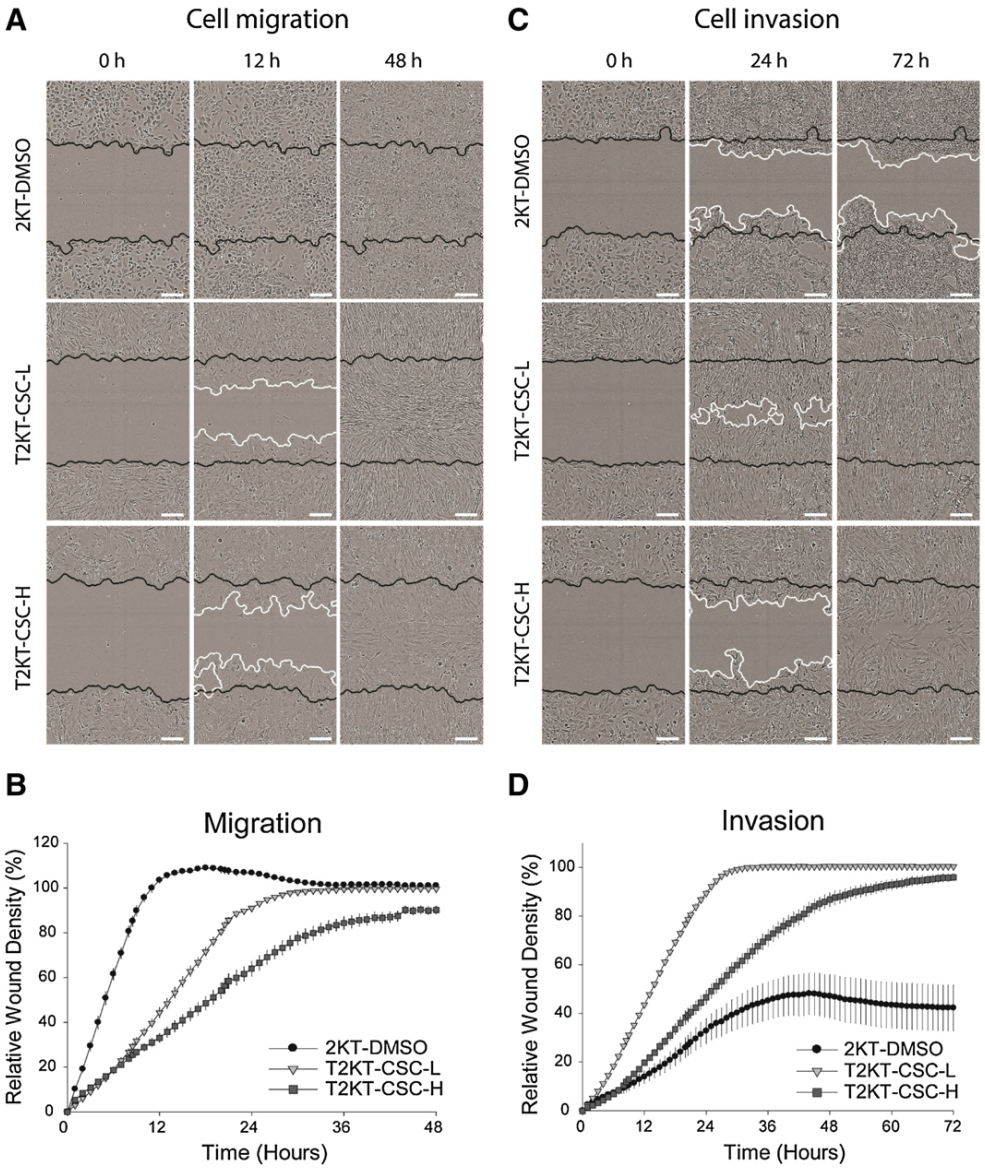

Fig. 4. *In vitro* migration and invasion capacity of HBEC2 cell lines. The capacity of untransformed and transformed HBEC2 cell lines to migrate or invade through a matrigel *in vitro* was measured by live cell imaging in a scratch wound healing assay (IncuCyte ZOOM). (A) Microscope images of migrating HBEC2, T2KT-CSC-L, and T2KT-CSC-H acquired at 0, 12, and 48 h. Black lines denote the original scratch wound and the white lines indicate the migration cell front. Scale bar: 150 μm. (B) Relative wound density. Migration of T2KT-CSC-L was significantly different from that of 2KT-DMSO after 1 h until 30 h and further forward, and T2KT-CSC-H was significantly different from 2KT-DMSO after 1 h (mixed models test, $p < 0.05$, $n = 8$). (C) Microscope images of invading HBEC2, T2KT-CSC-L, and T2KT-CSC-H acquired at 0, 24, and 72 h. Black lines denote the original scratch wound and the white lines the invasive cell front. (B) Relative wound density. Migration of T2KT-CSC-L was significantly different from that of 2KT-DMSO after 4 h, and T2KT-CSC-H was significantly different from 2KT-DMSO after 12 h (mixed models test, $p < 0.05$, $n = 8$).

the controls (approx. 25-fold; Fig. 3B). Furthermore, in the T2KT cell lines increased gene expression levels were observed for other EMT markers such as *ZEB1* (>100-fold), *VIM* (>10^3-fold), and *MMP2* (>10^3-fold). Expression of *TWIST* was also increased in T2KT cells, but not significantly.

Even though HBEC12 was the most readily transformed cell line, it showed a less clear pattern of EMT markers. No significant changes

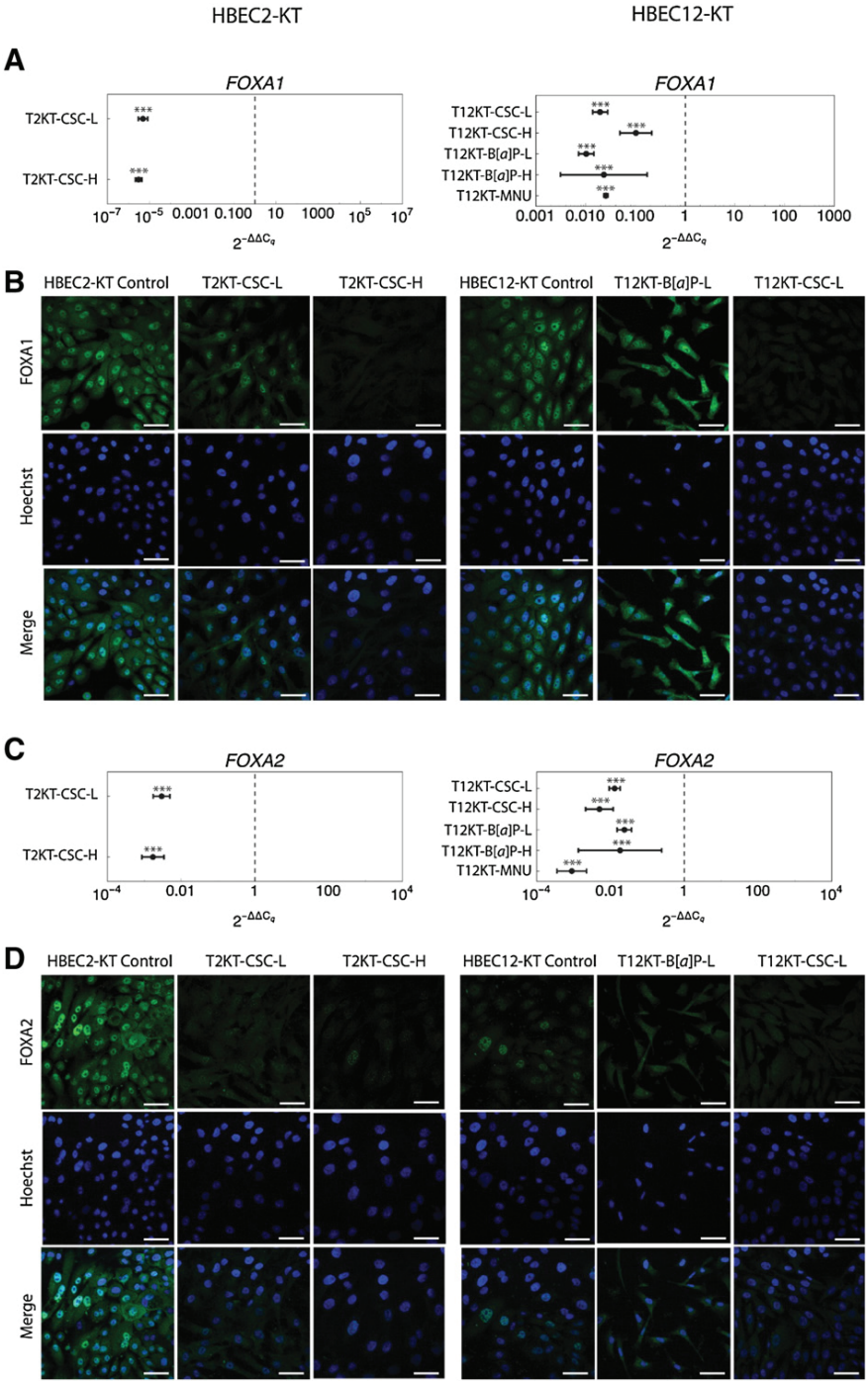

Fig. 5. *FOXA1* and *FOXA2* transcription factors are down-regulated in transformed cells. Expression of the pioneer transcription factors *FOXA1* and *FOXA2* was investigated in untransformed and transformed HBEC2 and HBEC12 cell lines at the gene and protein level. (A) and (C) Expression of *FOXA1* and *FOXA2* mRNA in transformed HBEC2 and HBEC12 cells, respectively. Relative gene expression levels as compared to the particular untransformed control cells are presented (fold change). (B) and (D) Representative images from confocal immunofluorescence microscopy of HBEC2 and HBEC12 control and transformed cells stained for FOXA1 (B) and FOXA2 (D), respectively. Nuclear/DNA staining was done by Hoechst. Scale bar: 50 μm. T2KT: transformed HBEC2, T12KT: Transformed HBEC12, CSC-L: cigarette smoke condensate 1.0 μg/μl, CSC-H: CSC 3.0 μg/μl, B[*a*]P-L: Benzo[*a*]pyrene 0.3 μM, B[*a*]P-H: B[*a*]P1.0 μM, MNU: Methylnitrosourea 0.5 mM. Gene expression data display mean ± SD, $n = 4$, where each replicate are clones isolated from individual colonies in soft-agar. ***$p < 0.001$ (ANOVA Dunnet's posttest).

were observed for *SNAI1*. However, *SNAI2* was significantly down-regulated in all but one of the T12KT transformants (Fig. 3). The expression of *ZEB1* was significantly increased in B[*a*]P and MNU transformed T12KTs, but not in the CSC transformed cells. No significant changes were observed for any of the other EMT markers in T12KTs.

3.3. Cell migration and invasion

One characteristic of EMT is that it increases the cells migration and invasion capabilities. A scratch-wound closure assay was performed to test the capability of the untransformed and transformed HBEC2 cells to migrate or to invade through a matrigel. All cell lines showed migration on the matrigel-coated surface (Fig. 4A). Interestingly, in case of HBEC2, untransformed cells showed the fastest rate of migration and complete closure was obtained after 12 h. T2KT-CSC-L and T2KT-CSC-H closed the wound after approximately 20 and 35 h, respectively. Statistically significant differences from the untransformed cells were observed between 1 and 30 h (T2KT-CSC-L) and from 1 h and beyond (T2KT-CSC-H) (Fig. 4B). In the invasion assay, the transformed HBEC2 cells readily invaded through the matrigel (Fig. 4C). The T2KT-CSC-L cell line closed the wound after approximately 30 h, whereas the T2KT-CSC-H showed almost complete wound closure after 72 h. In contrast, although some wound intrusion was apparent in the untransformed HBEC2 line during the first 40 h of the assay, this leveled off and further closure was not obtained. The invasiveness of T2KT-CSC-L was significantly different from that of untransformed HBEC2 after 4 h and beyond, and for T2KT-CSC-H it was significantly different from 12 h (mixed models, $p < 0.05$) (Fig. 4D). In addition to faster closure of the scratch wound, the transformed cells also invaded by an apparent different mechanism than the non-transformed cells including protrusions stretching into the matrigel (Supplementary Videos 1–3). Interestingly, a few hours after addition of the matrigel the control cells became less confluent on either side of the wound. When embedded in the matrigel the untransformed cells appeared to pack more tightly together creating open areas within the monolayer (Fig. 4C).

3.4. FOXA1/2 pioneer TF's are down-regulated during cell transformation

Deregulation of the transcription factors FOXA1 and FOXA2 has been implicated in tumorigenesis and EMT. Here, the expression levels of the *FOXA1* and *FOXA2* genes were down-regulated in all transformed cell lines compared to controls (Fig. 5). In T2KTs, *FOXA1* gene expression was reduced by 2×10^5 and 3×10^5-fold for low and high concentration of CSC exposure, respectively. In T12KTs, a smaller degree of down-regulation of *FOXA1* was observed ranging between 10-fold and 100-fold (Fig. 5A). The *FOXA2* expression level in T2KT was reduced approx. 300-fold and 600-fold for low and high dose CSC, respectively, and for T12KT the expression level was reduced between 40-fold and 10^3-fold (Fig. 5C).

By fluorescence microscopy analysis, untransformed control cells of both HBEC2 and HBEC12 showed nuclear staining of FOXA1 (Fig. 5B). FOXA1 expression was absent in T2KT-CSC-H and T12KT-CSC-L. Although *FOXA1* mRNA was almost absent in T2KT-CSC-L there was still some staining of FOXA1 in these cells. However, nuclear staining was significantly reduced compared to the untransformed HBEC2. In transformed T12KT-B[*a*]P-L cells no clear evidence of reduced nuclear staining was apparent. FOXA2 also showed nuclear staining in untransformed HBEC2 and HBEC12 cells, which was significantly reduced or almost abolished in the transformed cell lines (Fig. 5D). Thus, apart from FOXA1 in T12KT-B[*a*]P-L, immunohistochemical analysis confirmed the gene expression measurements.

Interestingly, down-regulation of *FOXA1* and *FOXA2* mRNA was observed in pre-transformed HBEC2 cells after 6, 9, and 12 weeks of carcinogen exposure (Fig. 6), indicating that down-regulation of these two transcription factors coincides or even precedes *in vitro* cell transformation.

3.5. ChIP-qPCR

To investigate possible mechanisms of altered *FOXA1* and *FOXA2* gene expression, analysis of binding of selected histone variants was performed. ChIP-qPCR assays for *FOXA1* were designed within or close to CpG-islands that are indicated to be bound by transcription factors in the ENCODE database and have chromatin state (ChromHMM) of promoter/enhancer (Fig. 7A). Binding of histone-H3 and the suppressive histone variant macroH2A to *FOXA1* was low at the transcription start site in untransformed HBEC2, indicative of an open chromatin state in the immediate vicinity of the transcription start site (TSS) (Fig. 7B and C). A minor increase in binding of both histone variants was observed at the TSS in both of the transformed cell lines. Histone-H3 and macroH2A showed increased binding in transformed cell lines at all the other sites analyzed, suggesting a more compact chromatin.

For the analysis of *FOXA2*, four genomic ChIP-qPCR assays were designed in possible regulatory regions (ENCODE) (Fig. 7E). Histone H3 and macroH2A show increased binding in transformed cell lines in all four assays with lower overall binding at the TSS (Fig. 7F and G).

Interestingly, binding of the histone variant H2A.Z, which has been associated with destabilized and actively transcribed chromatin, was increased in both *FOXA1* and *FOXA2* of transformed cells (Fig. 7D and H).

4. Discussion

HBEC2 and HBEC12 cells were exposed to CSC, B[*a*]P, or MNU for 12 weeks before selecting for transformed cells in a soft agar assay. Prior to this study these cell lines have been immortalized by the insertion of the *CDK4* and *TERT* genes (Ramirez et al., 2004). They were characterized as genomically stable, having few genetic changes and an intact tumor protein p53 checkpoint, making them well suited for lung carcinogenesis studies (Gazdar et al., 2010). While in the present study, neither of the unexposed control cell lines showed signs of transformation, tobacco carcinogen exposed cells formed colonies in soft agar. It should be noted, however, that whereas CSC was able to induce colony formation in both cell lines, B[*a*]P and MNU only produced colonies in HBEC12 after exposure for 12 weeks.

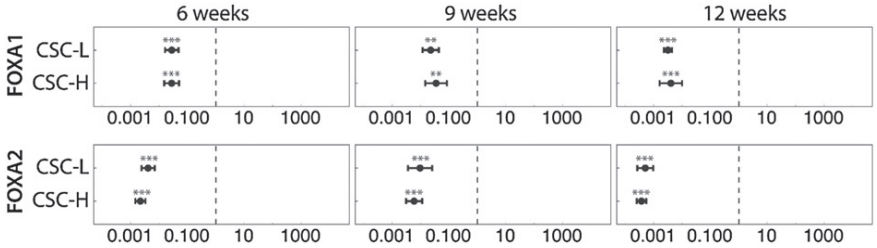

Fig. 6. Down-regulation of *FOXA1* and *FOXA2* in HBEC2 after long-term exposure to CSC. Expression levels of *FOXA1* and *FOXA2* were measured prior to selection of transformed cells in soft agar in HBEC2 cells exposed to low and high dose CSC for 6, 9, and 12 weeks. Relative gene expression levels as compared to the particular control cells (DMSO) are presented (fold change; mean ± SD, $n = 4$). $^{**}p < 0.01$, and $^{***}p < 0.001$, (ANOVA Dunnet's posttest).

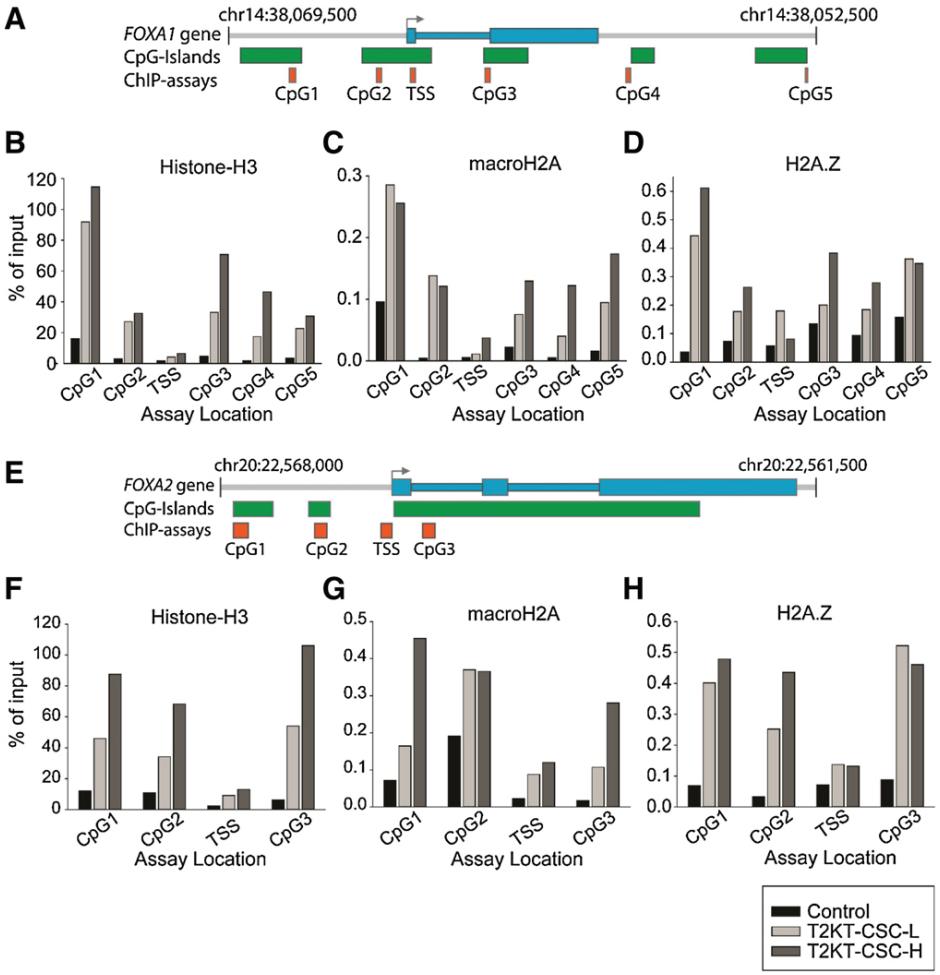

Fig. 7. Altered histone binding in genomic regions of *FOXA1* and *FOXA2*. (A) and (E) ChIP-qPCR assays for *FOXA1* and FOXA2 were designed in, or close to, five and four genomic CpG-islands, respectively. (B–D) *FOXA1* and (F–H) *FOXA2*. Binding of histone variants is presented as percentage of input on pooled chromatin from four independent experiments.

EMT is a transcriptional reprogramming of epithelial cells characterized by loss of cell-to-cell contact and gain of increased invasion capability. In this study, transformed cells showed characteristics of EMT, including a significant decrease of *CDH1* gene expression and concordantly reduced levels of E-cadherin protein. Loss of E-cadherin is associated with loss of epithelial phenotype and increased local invasiveness and metastasis (Onder et al., 2008; Vleminckx et al., 1991). The loss of E-cadherin and gain of N-cadherin (the "cadherin switch") is considered a hallmark of EMT. Although E-cadherin levels were reduced in both T2KTs and T12KTs, only the T2KT cell lines showed significantly increased levels of N-cadherin. The reason for this discrepancy is not known. The expression levels of *CDH1* and *CDH2* were significantly down- and up-regulated, respectively, prior to soft agar selection in cells exposed to CSC for 6, 9, and 12 weeks, compared to vehicle treated cells. However, the HBECs did not become transformed before 12 weeks of exposure. These results indicate that although deregulation of the cadherin switch may be important in enabling growth in soft agar, it is not sufficient.

Increased levels of *SNAI1* and *ZEB1* mRNA expression was observed in transformed HBEC2 cell lines, both of which are repressors of *CDH1* gene expression (Batlle et al., 2000; Eger et al., 2005). In HBEC12 transformed cells, however, no significant induction of *SNAI1* was observed. The level of *ZEB1* was increased in three out of five transformed cell lines (T12KT-B[*a*]P-H, T12KT-B[*a*]P-L, and T12KT-MNU). Similar to SNAI1 and ZEB1, SNAI2 is involved in repression of the *CDH1* gene (Conacci-Sorrell et al., 2003). However, the expression patterns of the repressors in the HBEC12 transformed cell lines were unclear and thus it is plausible that additional mechanisms may be involved in the regulation of *CDH1*. Indeed, it has been reported that reduced levels of *CDH1* can be attributed to increased DNA methylation levels in the promoter and translational suppression by binding of microRNA miR-9 (Lombaerts et al., 2006; Ma et al., 2010). Whether this is the case with the transformed cell lines in this study remains to be clarified.

Vimentin (VIM) and MMP2 are considered to play important roles in regulating the cells invasiveness by modulating the cytoskeleton and extracellular matrix, respectively (Giannelli et al., 1997; Ivaska et al., 2007). In the transformed HBEC2 cell lines an increase in expression of both these genes was observed, consistent with the invasive phenotype observed in the matrigel invasion assay. In HBEC12, the levels of *VIM* and *MMP2* were unchanged similarly to the transcription levels of the other EMT markers tested. These results suggest that although *VIM* and *MMP2* may be important in the cells ability to grow in soft agar, a change in the expression of these two genes is not the critical step for gaining this ability.

A decrease in the gene expression level of *SNAI2* was observed in all but one (T12KT-B[*a*]P-H) of the transformed cell lines. SNAI2 is, as SNAI1 and ZEB1, known to suppress expression of *CDH1*. In addition, *SNAI2* expression has been shown to be suppressed by direct binding of FOXA1 in its enhancer (Jin et al., 2013; Liu et al., 2012), thus it would be expected that decreased *FOXA1* expression levels would lead

to increased levels of *SNAI2*. However, this was not the case in either T2KT or T12KT cell lines suggesting that other mechanisms in addition to FOXA1 may be important in regulating *SNAI2* in human lung cells.

Consistent with the change in gene expression of the different EMT markers, T2KT cells were able to invade a matrigel faster than control cells. In addition, the morphology of invading cells was different in the transformed *vs* control cells (Supplementary Fig. 2). In contrast, increased migration rate was found in untransformed HBEC2 cells compared to the transformed derivatives. The common interpretation is that EMT and cell transformation involves increased invasiveness often in conjunction with increased migration. Interestingly, in agreement with our results it has been demonstrated that EMT can reduce cellular migration while increasing *in vitro* invasiveness, indicating that these processes may be uncoupled (Schaeffer et al., 2014).

In transformed cells originating from both HBEC2 and HBEC12 there was a significant reduction in the expression levels of *FOXA1* and *FOXA2*. The protein levels of both FOXA1 and FOXA2 in T2KT cells were also reduced as observed by fluorescent microscopy. For T12KTs, however, differences in the fluorescent signals were weak, but there was an indication of a difference in the protein levels. It is interesting to note that the reduction of FOXA1/2 in exposed cells occurs prior to clonal selection in soft agar, indicating that FOXA1/2 deregulation may be an early event.

FOXA1 has been shown to be important in transcriptional reprogramming of EMT in several cancers, including lung cancer, and >900 genes may be affected by FOXA1 deregulation (Jin et al., 2013; Wang et al., 2013). Hence, if the FOXA1 transcriptional level is deregulated by chemical exposure it may have great impact on the chromatin landscape and on the transcriptional program of the cell due to FOXA1s ability to bind compact chromatin and remodel chromatin structure. In the prostate, *FOXA1* has been found to positively regulate *CDH1,* resulting in reduced invasiveness *in vitro* (Jin et al., 2013). In a previous study on lung adenocarcinomas, *FOXA1* was overexpressed and amplified in a subset of the cases tested (Lin et al., 2002). In another study, alterations in *FOXA1* not only comprised gene amplification, but also deletion and sporadic mutation in lung adenocarcinoma (Li et al., 2013). Thus, it is apparent that regulation of *FOXA1* in lung cancer may be complex. However, down-regulation of *FOXA1* during *in vitro* transformation with concomitant appearance of EMT is in agreement with a role of *FOXA1* in maintaining the epithelial phenotype. Interestingly, Jin et al. (2013) found that while *FOXA1* is slightly up-regulated in localized cancer of the prostate it is down-regulated in metastatic prostate cancer. *FOXA2* may also suppress EMT and thereby metastasis in human lung cancer cell lines (Tang et al., 2011). In agreement with our study, FOXA2 was down-regulated in lung adenocarcinomas (Basseres et al., 2012).

The results of the present study indicate that the expression of the *FOXA* genes is down-regulated by long-term carcinogen exposure, and that this coincides with the occurrence of EMT. In our model, it is thus reasonable to speculate that deregulation of *FOXA1*/*FOXA2* and EMT are associated, and that these mechanisms are involved in the tobacco smoke carcinogen induced transformation of HBECs.

To get insight into the possible mechanisms involved in their down-regulation, binding of histone H3 and the histone variant macroH2A in the genomic regions surrounding *FOXA1* and *FOXA2* was quantitated by ChIP-qPCR. These analyses included CpG islands, chromatin state, and regions showing high TF binding in the *FOXA1* and *FOXA2* genes that were all identified from the ENCODE database (UCSC Genome Browser) (Fig. 7) (Kent et al., 2002; Rosenbloom et al., 2013). Lower histone binding was found in regions surrounding the transcription start site (TSS) in both *FOXA1* and *FOXA2*, consistent with nucleosome depletion around TSS (Lee et al., 2004). MacroH2A binds *FOXA2* in cell lines that do not express *FOXA2* while being absent in cell lines that do express the gene (Barrero et al., 2013). This is consistent with our finding of increased macroH2A binding in transformed cell lines. Interestingly, for all investigated chromatin locations, higher levels of histone binding were found in the transformed cell lines (T2KT-CSC-L, T2KT-CSC-H) compared to the control cell line. This result suggests a more compact chromatin state surrounding the *FOXA* genes in the transformed cell lines, consistent with the reduced level of transcription.

Using the same qPCR assays as above, the level of histone variant H2A.Z was found to be increased in the transformed HBEC2 cell lines when compared to the controls. H2A.Z preferentially binds in regions flanking TSS and is associated with nucleosome depletion and active gene expression (Barski et al., 2007; Li et al., 2012). In gene bodies, however, increased H2A.Z has been associated with repression of gene expression (Zemach et al., 2010). This is in accordance with our results for the GpG3 assays for both *FOXA1* and *FOXA2* (Fig. 7A and E). However, the finding of increased H2A.Z in the other *FOXA1* and *FOXA2* genomic regions tested was unexpected.

Exposure of HBECs to activated carcinogenic metabolites (such as BPDE) has previously been shown to transform HBECs *in vitro* (Damiani et al., 2008). In their paper, Damiani et al. (2008) showed that E-cadherin was down-regulated due to increased DNA methylation in the *CDH1* promoter of transformed cells. Consistent with their finding we find that all transformed cell lines have reduced levels of E-cadherin, confirming the importance of adherence proteins in the cell transformation process. Both HBEC2 and HBEC12 were responsive to B[*a*]P and CSC in regard to induction of typical PAH induced xenobiotic metabolism genes like CYP1A1 and CYP1B1 (data not shown), indicating their capacity to metabolically activate PAH pro-carcinogens (Uppstad et al., 2010). However, only in HBEC12 this resulted in transformed cells after exposure to B[*a*]P.

Similarly to the findings reported by Damiani et al. (2008)), our transformed cell lines were incapable of forming tumors in immunodeficient nude mice (data not shown). This suggests that even though the cells have gained some tumor cell characteristics including soft agar colony formation, EMT and *in vitro* invasiveness, they have not developed to become tumorigenic and the model may thus be representative of early stages of lung carcinogenesis.

In summary, our results show that long-term exposure of immortalized human bronchial epithelial cells to tobacco smoke carcinogens can lead to cell transformation *in vitro*. The transformed cells are able to grow anchorage independent and show increased invasiveness. In addition they have reduced expression of FOXA1 and FOXA2, co-occurring with EMT. These results indicate that EMT and down-regulation of FOXA1 and FOXA2 may be important in chemical carcinogen induced transformation of human bronchial epithelial cells, and should stimulate further studies on their role in lung cancer.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.tiv.2016.04.012.

Transparency Document

The Transparency document associated with this article can be found in the online version.

Acknowledgements

The study was supported by grants from the Norwegian Cancer Society (0971). The authors wish to thank Dr. Øivind Skare for assistance with statistical analysis.

References

Alexander, N.R., Tran, N.L., Rekapally, H., Summers, C.E., Glackin, C., Heimark, R.L., 2006. N-cadherin gene expression in prostate carcinoma is modulated by integrin-dependent nuclear translocation of Twist1. Cancer Res. 66, 3365–3369. http://dx.doi.org/10.1158/0008-5472.can-05-3401.

Barrero, M.J., Sese, B., Kuebler, B., Bilic, J., Boue, S., Martí, M., Izpisua Belmonte, J.C., 2013. Macrohistone variants preserve cell identity by preventing the gain of H3K4me2 during reprogramming to pluripotency. Cell Rep. 3, 1005–1011. http://dx.doi.org/10.1016/j.celrep.2013.02.029.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., Zhao, K., 2007. High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837. http://dx.doi.org/10.1016/j.cell.2007.05.009.

Basseres, D.S., D'Alò, F., Yeap, B.Y., Löwenberg, E.C., Gonzalez, D.A., Yasuda, H., Dayaram, T., Kocher, O.N., Godleski, J.J., Richards, W.G., Meyerson, M., Kobayashi, S., Tenen, D.G., Halmos, B., Costa, D.B., 2012. Frequent downregulation of the transcription factor Foxa2 in lung cancer through epigenetic silencing. Lung Cancer 77, 31–37.

Batlle, E., Sancho, E., Francí, C., Domínguez, D., Monfar, M., Baulida, J., García De Herreros, A., 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat. Cell Biol. 2, 84–89. http://dx.doi.org/10.1038/35000034.

Bremnes, R.M., Veve, R., Gabrielson, E., Hirsch, F.R., Baron, A., Bemis, L., Gemmill, R.M., Drabkin, H.A., Franklin, W.A., 2002. High-throughput tissue microarray analysis used to evaluate biology and prognostic significance of the E-cadherin pathway in non-small-cell lung cancer. J. Clin. Oncol. 20, 2417–2428. http://dx.doi.org/10.1200/JCO.2002.08.159.

Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.a., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., Fox, E.a., Silver, P.a., Brown, M., 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell 122, 33–43. http://dx.doi.org/10.1016/j.cell.2005.05.008.

Conacci-Sorrell, M., Simcha, I., Ben-Yedidia, T., Blechman, J., Savagner, P., Ben-Ze'Ev, A., 2003. Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of β-catenin signaling, slug, and MAPK. J. Cell Biol. 163, 847–857. http://dx.doi.org/10.1083/jcb.200308162.

Cortazar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., Steinacher, R., Jiricny, J., Bird, A., Schar, P., 2011. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature 470, 419–423. http://dx.doi.org/10.1038/nature09672.

Damiani, L.A., Yingling, C.M., Leng, S., Romo, P.E., Nakamura, J., Belinsky, S.A., 2008. Carcinogen-induced gene promoter hypermethylation is mediated by DNMT1 and causal for transformation of immortalized bronchial epithelial cells. Cancer Res. 68, 9005–9014. http://dx.doi.org/10.1158/0008-5472.CAN-08-1276.

Dela Cruz, C.S., Tanoue, L.T., Matthay, R.A., 2011. Lung cancer: epidemiology, etiology, and prevention. Clin. Chest Med. 32. http://dx.doi.org/10.1016/j.ccm.2011.09.001.

Delgado, O., Kaisani, A.A., Spinola, M., Xie, X.-J., Batten, K.G., Minna, J.D., Wright, W.E., Shay, J.W., 2011. Multipotent capacity of immortalized human bronchial epithelial cells. PLoS One 6, e22023.

Deutsch, L., Wrage, M., Koops, S., Glatzel, M., Uzunoglu, F.G., Kutup, A., Hinsch, A., Sauter, G., Izbicki, J.R., Pantel, K., Wikman, H., 2012. Opposite Roles of FOXA1 and NKX2–1 in Lung Cancer Progression. Genes Chromosom. Cancer 51, 618–629. http://dx.doi.org/10.1002/gcc.

Eger, A., Aigner, K., Sonderegger, S., Dampier, B., Oehler, S., Schreiber, M., Berx, G., Cano, A., Beug, H., Foisner, R., 2005. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene 24, 2375–2385. http://dx.doi.org/10.1038/sj.onc.1208429.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F., 2015. Cancer Incidence and Mortality Worldwide: Sources, Methods and Major Patterns in GLOBOCAN 2012. p. 386 http://dx.doi.org/10.1002/ijc.29210.

Gazdar, A.F., Gao, B., Minna, J.D., 2010. Lung cancer cell lines: useless artifacts or invaluable tools for medical science? Lung Cancer 68, 309–318.

Giannelli, G., Falk-Marzillier, J., Schiraldi, O., Stetler-Stevenson, W.G., Quaranta, V., 1997. Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5. Science 277, 225–228. http://dx.doi.org/10.1126/science.277.5323.225 (80-).

Haldorsen, T., Andersen, A., Boffetta, P., 2004. Smoking-adjusted incidence of lung cancer by occupation among Norwegian men. Cancer Causes Control 15, 139–147. http://dx.doi.org/10.1023/B:CACO.0000019485.74818.d6.

Hecht, S.S., 2006. Cigarette smoking: cancer risks, carcinogens, and mechanisms. Langenbeck's Arch. Surg. 391, 603–613.

Hecht, S.S., 2012. Research opportunities related to establishing standards for tobacco products under the Family Smoking Prevention and Tobacco Control Act. Nicotine Tob. Res. 14, 18–28. http://dx.doi.org/10.1093/ntr/ntq216.

Ivaska, J., Pallari, H.-M., Nevo, J., Eriksson, J.E., 2007. Novel functions of vimentin in cell adhesion, migration, and signaling. Exp. Cell Res. 313, 2050–2062. http://dx.doi.org/10.1016/j.yexcr.2007.03.040.

Jin, H.-J., Zhao, J.C., Ogden, I., Bergan, R.C., Yu, J., 2013. Androgen receptor-independent function of FoxA1 in prostate cancer metastasis. Cancer Res. 73, 3725–3736. http://dx.doi.org/10.1158/0008-5472.CAN-12-3468.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. Genome Res. 12, 996–1006. http://dx.doi.org/10.1101/gr.229102.

Lee, C.K., Shibata, Y., Rao, B., Strahl, B.D., Lieb, J.D., 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat. Genet. 36, 900–905. http://dx.doi.org/10.1038/ng1400.

Li, X., Chen, S., Sun, T., Xu, Y., Chen, Y., Liu, Y., 2013. The transcriptional regulation of SOX2 on FOXA1 gene and its application in diagnosis of human breast and lung cancers. Transcr. Regul. SOX2 FOXA1 Gene Its Appl. Diagnosis Hum. Breast Lung Cancers.

Li, Z., Gadue, P., Chen, K., Jiao, Y., Tuteja, G., Schug, J., Li, W., Kaestner, K.H., 2012. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell 151, 1608–1616. http://dx.doi.org/10.1016/j.cell.2012.11.018.

Lin, L., Miller, C.T., Contreras, J.I., Prescott, M.S., Dagenais, S.L., Wu, R., Yee, J., Orringer, M.B., Misek, D.E., Hanash, S.M., Glover, T.W., Beer, D.G., 2002. The hepatocyte nuclear factor 3 alpha gene, HNF3alpha (FOXA1), on chromosome band 14q13 is amplified and overexpressed in esophageal and lung adenocarcinomas. Cancer Res. 62, 5273–5279.

Liu, Y.-N., Abou-Kheir, W., Yin, J.J., Fang, L., Hynes, P., Casey, O., Hu, D., Wan, Y., Seng, V., Sheppard-Tillman, H., Martin, P., Kelly, K., 2012. Critical and reciprocal regulation of KLF4 and SLUG in transforming growth factor-initiated prostate cancer epithelial-mesenchymal transition. Mol. Cell. Biol. 32, 941–953. http://dx.doi.org/10.1128/MCB.06306-11.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25, 402–408.

Lombaerts, M., van Wezel, T., Philippo, K., Dierssen, J.W.F., Zimmerman, R.M.E., Oosting, J., van Eijk, R., Eilers, P.H., van de Water, B., Cornelisse, C.J., Cleton-Jansen, A.-M., 2006. E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br. J. Cancer.

Ma, L., Young, J., Prabhala, H., Pan, E., Mestdagh, P., Muth, D., Teruya-Feldstein, J., Reinhardt, F., Onder, T.T., Valastyan, S., Westermann, F., Speleman, F., Vandesompele, J., Weinberg, R.A., 2010. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nat. Cell Biol. 12, 247–256 (doi:http://www.nature.com/ncb/journal/v12/n3/suppinfo/ncb2024_S1.html).

Morel, A.P., Hinkal, G.W., Thomas, C., Fauvet, F., Courtois-Cox, S., Wierinckx, A., Devouassoux-Shisheboran, M., Treilleux, I., Tissier, A., Gras, B., Pourchet, J., Puisieux, I., Browne, G.J., Spicer, D.B., Lachuer, J., Ansieau, S., Puisieux, A., 2012. EMT inducers catalyze malignant transformation of mammary epithelial cells and drive tumorigenesis towards claudin-low tumors in transgenic mice. PLoS Genet. 8. http://dx.doi.org/10.1371/journal.pgen.1002723.

Nagathihalli, N.S., Massion, P.P., Gonzalez, A.L., Lu, P., Datta, P.K., 2012. Smoking induces epithelial-to-mesenchymal transition in non-small cell lung cancer through HDAC-mediated downregulation of E-cadherin. Mol. Cancer Ther. 47. http://dx.doi.org/10.1158/1535-7163.MCT-12-0107.

Onder, T.T., Gupta, P.B., Mani, S.A., YANG, J., Lander, E.S., Weinberg, R.A., 2008. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. Cancer Res. 68, 3645–3654.

Pozharskaya, V., Torres-González, E., Rojas, M., Gal, A., Amin, M., Dollard, S., Roman, J., Stecenko, A.a., Mora, A.L., 2009. Twist: a regulator of epithelial-mesenchymal transition in lung fibrosis. PLoS One 4. http://dx.doi.org/10.1371/journal.pone.0007559.

Ramirez, R.D., Sheridan, S., Girard, L., Sato, M., Kim, Y., Pollack, J., Peyton, M., Zou, Y., Kurie, J.M., Dimaio, J.M., Milchgrub, S., Smith, A.L., Souza, R.F., Gilbey, L., Zhang, X., Gandia, K., Vaughan, M.B., Wright, W.E., Gazdar, A.F., Shay, J.W., Minna, J.D., 2004. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. Cancer Res. 64, 9027–9034.

Robinson, J.L.L., Carroll, J.S., 2012. FoxA1 is a key mediator of hormonal response in breast and prostate cancer. Front. Endocrinol. (Lausanne) 3, 1–6. http://dx.doi.org/10.3389/fendo.2012.00068.

Rodgman, A., Perfetti, T.A., 2013. The Chemical Components of Tobacco and Tobacco Smoke. CRC Press.

Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G., Lee, B.T., Barber, G.P., Harte, R.A., Diekhans, M., Long, J.C., Wilder, S.P., Zweig, A.S., Karolchik, D., Kuhn, R.M., Haussler, D., Kent, W.J., 2013. ENCODE data in the UCSC genome browser: year 5 update. Nucleic Acids Res. 41, D56–D63. http://dx.doi.org/10.1093/nar/gks1172.

Schaeffer, D., Somarelli, J.A., Hanna, G., Palmer, G.M., Garcia-Blanco, M.A., 2014. Cellular migration and invasion uncoupled: increased migration is not an inexorable consequence of epithelial-to-mesenchymal transition. Mol. Cell. Biol. 34, 3486–3499. http://dx.doi.org/10.1128/mcb.00694-14.

Siegel, R.L., Miller, K.D., Jemal, A., 2015. Cancer statistics, 2015. CA Cancer J. Clin. 65, 5–29. http://dx.doi.org/10.3322/caac.21254.

Song, Y., Washington, M.K., Crawford, H.C., 2010. Loss of FOXA1/2 is essential for the epithelial-to-mesenchymal transition in pancreatic cancer. Cancer Res. 70, 2115–2125.

Tang, Y., Shu, G., Yuan, X., Jing, N., Song, J., 2011. FOXA2 functions as a suppressor of tumor metastasis by inhibition of epithelial-to-mesenchymal transition in human lung cancers. Cell Res. 21, 316–326.

Thiery, J.P., 2002. Epithelial-mesenchymal transitions in tumour progression. Nat. Rev. Cancer 2, 442–454. http://dx.doi.org/10.1038/nrc822.

Thiery, J.P., Acloque, H., Huang, R.Y.J., Nieto, M.A., 2009. Epithelial-mesenchymal transitions in development and disease. Cell 139, 871–890.

Uppstad, H., Øvrebø, S., Haugen, A., Mollerup, S., 2010. Importance of CYP1A1 and CYP1B1 in bioactivation of benzo[a]pyrene in human lung cell lines. Toxicol. Lett. 192, 221–228.

Ura, H., Bonfil, D.R., Reich, R., Reddel, R., Pfeifer, A., Harris, C.C., Klein-Szanto, A.J.P., 1989. Expression of type IV collagenase and procollagen genes and its correlation with the tumorigenic, invasive, and metastatic abilities of oncogene-transformed human bronchial epithelial cells. Cancer Res. 49, 4615–4621.

Vesuna, F., van Diest, P., Chen, J.H., Raman, V., 2008. Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer. Biochem. Biophys. Res. Commun. 367, 235–241. http://dx.doi.org/10.1016/j.bbrc.2007.11.151.

Vleminckx, K., Vakaet, L., Mareel, M., Fiers, W., Van Roy, F., 1991. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell 66, 107–119. http://dx.doi.org/10.1016/0092-8674(91)90143-M.

Wang, H., Meyer, C.A., Fei, T., Wang, G., Zhang, F., Liu, X.S., 2013. A systematic approach identifies FOXA1 as a key factor in the loss of epithelial traits during the epithelial-to-mesenchymal transition in lung cancer. BMC Genomics 14, 680. http://dx.doi.org/10.1186/1471-2164-14-680.

Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T., Hirohashi, S., 1995. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. Proc. Natl. Acad. Sci. U. S. A. 92, 7416–7419.

Zemach, A., McDaniel, I.E., Silva, P., Zilberman, D., 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328, 916–919. http://dx.doi.org/10.1126/science.1186366.

II

III