universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

“Identification of unknown epigenetically regulated genes in non-small cell lung cancer”

Durchgeführt an der Medizinischen Universität Wien
Univ. Klinik für Innere Med. I/Klinische Abteilung für Onkologie

Verfasserin

Valerie Nadine Babinsky

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2011

Studienkennzahl lt. Studienblatt:	A 441
Studienrichtung lt. Studienblatt:	Mikrobiologie und Genetik
Betreuer:	Ao.Prof. Dr. Wolfgang Mikulits

1 TABLE OF CONTENT

2 ACKNOWLEDGMENT / DANKSAGUNG

Diese Seite verschafft mir die Möglichkeit, jenen zu danken, die meinen Werdegang ermöglichen und trotzdem viel zu selten ein „DANKE“ hören.

Meine Eltern Chantal Babinsky-Chatain und Dr. Heinrich Babinsky, die wenn es nötig wäre, für mich die Welt auf den Kopf stellen würden; die mir von klein auf alles ermöglicht haben und mich mit Unterstützung und Liebe überhäuft haben. Ihr habt in großem Maße Anteil, an dem was ich bin und an meinen Erfolgen. DANKE

Meine Schwestern Mag. Fabienne Babinsky und Nathalie Babinsky, die mir zu jeder Zeit wahre Freunde sind und mir die nötige Motivation geben, alles schaffen zu können. DANKE

Mein Freund Peter Stekovits, der mir in schweren Momenten immer zur Seite steht und der mich immer wieder ermutigt, meinen Weg weiter zu gehen. DANKE

Besonderer Dank gebührt Frau Univ.-Prof. Dr. Sabine Zöchbauer-Müller, die mir durch diese Diplomarbeitsstelle, eine der wichtigsten Erfahrungen meines erwachsenen Lebens ermöglicht hat. DANKE

Mein täglicher Arbeitsalltag wurde erst durch meine Kollegen Dr. Gerwin Heller und Barbara Ziegler zu dem was es war: Eine wahre Freude! Dr. Gerwin Heller, der mich durch seine enorme Kompetenz und seine Arbeitsweise auf den Beginn meiner wissenschaftlichen Karriere bestens vorbereitet hat. Barbara Ziegler, die mir nicht nur technisch immer zu Seite gestanden ist, sondern auch auf persönlicher Ebene zu einer Freundin geworden ist, die ich in meinem Leben nicht missen möchte. DANKE

Besonders verbunden fühle ich mich Herrn Prof. Dr. Wolfgang Mikulits für die Begutachtung und die Benotung meiner Arbeit. DANKE

DANKE

3 INTRODUCTION

"Human cancers result from the accumulation of genetic alterations at specific chromosomal regions involving a multistep process" (2). For a long time mutations seemed to be the most important alterations in lung cancer cells. However, in recent years it has been shown that epigenetic changes, especially changes in DNA methylation patterns play an important role in the pathogenesis of malignant diseases. Both, hypomethylation of transcriptional inactive genomic regions and hypermethylation of cytosine- guanine (CpG) rich areas, called CpG islands, which are mostly located 5' of the transcription start site of actively transcribed genes, are frequently occurring changes in lung cancer cells (3-6). Alterations of DNA methylation patterns appear to be of tremendous importance in all steps of lung cancer pathogenesis occurring already in pre-carcinogenous lesions.

To analyze the alteration in DNA methylation patterns mainly primary NSCLC samples obtained from surgeries are used. However, DNA methylation cannot only be detected in primary NSCLC samples but also in blood, sputum, bronchial brushings, and bronchioloalveolar lavage (BAL) samples from NSCLC patients (7-14).

Costello et al. (15) reported that on average 600 CpG islands are targets for methylation in tumor cells. So far only ~ 150 genes were identified to be tumor-specifically methylated in NSCLC. Thus, increasing interest in large-scale and genome-wide analyze methods arose in recent years. Genome-wide methylation assays were developed mostly based on two methods. On the one hand on pharmacological re-activation of epigenetically regulated genes combined with gene expression microarray analysis and on the other hand on microarray analyses of DNA fragments after immunoprecipitation of methylated DNA fragments.

In this project two genome-wide methods are used to analyze DNA methylation in three adenocarcinoma cell lines (A549, NCI-H1993 and NCI-H2073). Results are verified with methylation specific PCR (MSP) to identify so far unknown targets for DNA methylation in NSCLC.

4 AIM OF STUDY

Silencing of certain tumor suppressor genes mediated by DNA methylation is a critical event in the pathogenesis of NSCLC. To date about ~150 genes are known targets for methylation in NSCLC. However, it has been reported that on average 600 CpG islands are methylated in a tumor (15). So far, DNA methylation in lung cancer was mainly investigated by single-gene methylation analyses. To get further insight into the significance of DNA methylation mediated gene silencing in NSCLC, we propose the following hypotheses:

- **Hypothesis 1:** The overwhelming majority of methylated genes in NSCLC are still unknown.
- **Hypothesis 2:** Methylation of unknown target genes in NSCLC is tumor-specific.
- **Hypothesis 3:** Methylated genes are involved in certain molecular pathways whose depletion may lead to the development of a malignant phenotype.

Ad hypothesis 1: Investigations of the methylation status of single genes revealed that DNA methylation plays an important role in the pathogenesis of NSCLC. Whole genome approaches allow the search for methylated genes in NSCLC on a large scale approach.

Thus, we performed a high-throughput search for genes whose expression is up-regulated by epigenetically active drugs in 3 NSCLC cell lines. This assay is based on the observation that epigenetic modifications are reversible. Therefore, we treated the NSCLC cell lines A549, NCI-H1993 and NCI-H2073 with the DNA methyltransferase inhibitor 5-aza-2´-deoxycytidine (Aza-dC) and/or the histone deacetylase inhibitor trichostatin A (TSA). Subsequently, Affymetix microarray analyses were performed using Affymetrix U133 plus 2.0 arrays.

For the vast majority of up-regulated genes the finding that they are inactivated by methylation will be new suggesting that these results will enhance our understanding of the molecular pathogenesis of NSCLC.

Ad hypothesis 2: We established an approach for whole genome analysis of DNA methylation: methylated DNA immunoprecipitation (MeDIP) combined with microarray analysis (MeDIP-chip). This approach is based on enrichment of methylated DNA fragments

using an anti-5-methylcytosine antibody followed by hybridization to CpG island microarrays. Using this approach we analyze the methylation patterns of 3 NSCLC cell lines and normal human bronchial epithelial cells (NHBEC) and compare the results to identify tumor-specifically methylated genes. Tumor-specific methylation of selected genes is confirmed by methylation-specific PCR in a set of primary tumor and corresponding non-malignant lung tissue samples of 10 NSCLC patients.

Ad hypothesis 3: Additionally, we want to identify the function of genes silenced by DNA methylation and their relevance in cancer related pathways. For this purpose *in silico* analysis are performed for Gene Ontology (GO) term enrichment.

5 BACKGROUND

5.1 Epigenetics: Histone protein modifications and DNA methylation

Literally "epi- (Greek: επί- over, above) genetic" incorporates all mechanisms that regulate gene expression but do not alter the DNA sequence itself leading to inheritable changes in the phenotype (16). Apart from methylation of the DNA strand itself, histone proteins can be modified in several different ways including methylation, acetylation, phosphorylation, ubiquitinylation and sumoylation (17, 18). By changing the conformation and the density of the DNA, epigenetic mechanisms are important regulators of gene transcription machinery. Epigenetic changes occur in the context of nucleosomes. Nucleosomes are the fundamental units of chromatin, consisting of a 146 bp DNA region, which is wrapped ~1.8 times around a protein octamer composed of 2 times H2A, H2B, H3 and H4 histone proteins (19). Histones of the H1 family (also called linker histones) are not connected to the histone octamer itself but interact directly with DNA between two of these octamers. This interaction leads to a higher level of organization which is called "solenoid" helical fibres (30nm fibres) (20-22). Schematically the DNA formation can be described as a "pearl necklace" (see Figure 1). The amino-terminal tails of histone proteins reach out of the "pearls" and are targets for post-replicative modifications.

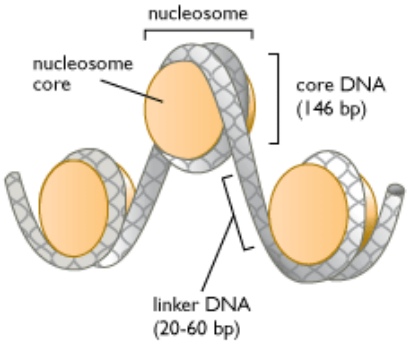

Figure 1
The fundamental unit of the chromatin: the nucleosome. 146bp long DNA regions are wrapped around a histone protein octamer. DNA regions lying between the nucleosomes are called linker DNA and are 20-60 bp long (1).

The combination of DNA methylation and histone protein modification results in the formation of either euchromatin or heterochromatin. The term euchromatin characterizes DNA regions showing a light compactation and consequently high gene transcription level. In contrast to that the term heterochromatin characterizes highly packed DNA regions which are transcriptionally inactive.

Recent studies suggest that DNA methylation is directly associated with certain modifications of histone proteins, as further discussed in section 5.1.2. The accruing epigenetic code of each cell is the outcome of a dynamic process, which is in general stable and can be inherited to following generations. Apart from its essential physiological function in each cell aberrant epigenetic regulation can lead to several pathologies such as cancer (23, 24).

5.1.1 Histone protein modifications

The most important histone modifications are acetylation, phosphorylation and methylation (25). Histone modifications affect DNA compactation in either cis-effect, meaning that the chromosomal structure is directly affected by the modification, or in trans-effect leading to an inhibition of other factors, for example DNA binding proteins that causes change of the chromatin formation (26, 27).

So far, histone acetylation is the best studied histone modification. During histone acetylation, acetyl groups ($-COCH_3$) are transferred from the donor acetyl-CoA to lysine residues in the histone tail, catalyzed by enzymes of the histone acetyltransferase (HATs) family (28). Generally, addition of an acetyl group to a histone tail diminishes the positive charge of the histone octamer resulting in a weaker interaction between the negatively charged DNA backbone and the histone octamer what leads to reduced chromatin compactation. The obtained formation favors higher transcriptional activity due to higher accessibility of the DNA for the transcription machinery (29). The antagonists of HATs are histone deacetylases (HDACs) which catalyze the elimination of acetyl groups from the histone tails leading to a repressive chromatin structure (heterochromatin).

Histone acetylation is a very dynamic process. Removing acetyl groups from the DNA, which can be undertaken within minutes leads to a conversion from actively transcribed DNA into inactive and densely packed DNA. On the contrary, addition of acetyl groups to the DNA takes longer suggesting that inhibition of gene expression can be enabled in a short period of time. Gene activation is a longer taking process (30).

Besides this molecular regulation mechanism phosporylation of histone tails is important for regulating gene expression. Phosphate groups ($-PO_4$) are mainly transferred to serine

residues within histone tail of H3 (31). Moreover, phosphorylation can not only be found within histone octamers but also on a conserved serine/threonine, proline, any amino acid, lysine/arginine (S/T-P-X- K/R) motif within the N-terminus of H1. In this way the direct interaction of H1 with the DNA is inhibited and the chromatin fiber is destabilized (32). The transfer of the phosphate group is catalyzed by enzymes of the kinase family.

Another important histone modification is mainly found on lysine residues: histone methylation. The transfer of the methyl group ($-CH_3$) is catalyzed by enzymes of the protein histone lysine methyltransferase. Histone methylation occurs within histone H3 (K4, K9, K27, K36, K79) and on one lysine residue within histone H4 (K20). Depending on the position of the methyl group this modification can operate either as a transcription activator or inhibitor by altering density of the chromatin compactation. While methylation at H3-K9 leads to transcriptional repression, methylation at H3-K4 is associated with transcriptional activation (33).

These modifications can affect each other in a synergistic or antagonistic way and are not independent transcription alteration factors. Thus, activation or inhibition of gene expression results from the sum of epigenetic modifications, the histone-code (34, 35).

5.1.2 DNA methylation

Apart from modifications of histone proteins which are involved in the regulation of gene expression (see section 5.1.1) DNA itself can be target for post-replicative modification by DNA methylation, thus, affecting the transcriptional activity of specific genes. DNA methylation regulates gene expression either directly or by influencing the modification of the histone proteins which influences gene expression (36).

Two different mechanisms for DNA methylation are known. D*e novo* methylation by adding a methyl group to a so far unmethylated cytosine and *maintenance* methylation which describes the copying process of pre-existing methylation patterns from the mother strand to the newly synthesized DNA strand after replication. *Maintenance* methylation facilitates inheritance of the DNA methylation patterns during cell division (37).

Proteins specifically recognizing different states of cytosine modification are necessary to fulfill these different tasks. Proteins responsible for *de novo* methylation bind unmethylated

cytosines and enzymes necessary for *maintenance* methylation have to bind specifically to hemimethylated DNA sections. Figure 2 shows these two mechanisms.

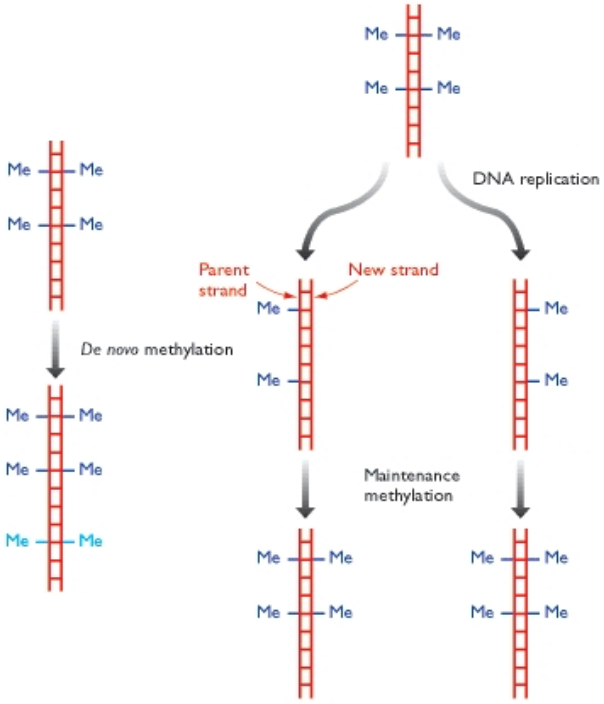

Figure 2
Illustration of the different methylation processes. *De novo* methylation is the introduction of methyl groups at previously unmethylated cytosines. *Maintenance* methylation ensures conservation of the methylation pattern during replication by copying the methylation pattern from the mother strand to the daughter stand (38).

5.1.2.1 CpG islands

In mammalian somatic cells cytosines in cytosine-guanine dinucletides (CpG) are the targets for DNA methylation (39). CpG dinucleotides are widely under re-presented in the whole genome. Only in some regions of the genome the estimated CpG ratio and even higher ratios can be found (40, 41). These regions are called CpG islands and are in general 0.5-4 kb in length (42, 43). Methylated CpGs are in most cases localized at transcriptionally inactive or even heterochromatin sections of the genome suggesting that methylation of a CpG island is correlated with compactation of the DNA (44). Computer analyses predicted that 28.890 CpG islands are present in a mammalian genome (45, 46). But 70%-80% of CpG dinucleotides in the mammalian genome are methylated (47). However, in 5' regions of approximately 60% of all genes CpG islands can be found (48). Thereby, CpG islands can overlap with the promoter region and can extend even into the exonic regions. With the

exception of X-chromosomally inactivated genes, imprinted genes and tissue specifically expressed genes CpG islands in 5' regions of genes are generally unmethylated (49).

5.1.2.2 CpG methylation

The transfer of a methyl group (-CH_3) from a universal methyl group donor called S-Adenosyl-L-Methionin (SAM or AdoMet) to a 5' carbon of cytosine bases is called DNA methylation (50). SAM is derived in a metabolic pathway called "one-carbon cycle" from the source material amino acid L-methionine. The amino acid like molecule is an important methyl group donor in many enzymatic pathways such as synthesis of hormones, neurotransmitters, nucleic acids, proteins, and phospholipids (51). The methyl group is bound to a charged sulfate atom and is then highly reactive towards polarisable nucleophiles such as N, O and S as well as towards activated C atoms (like carbanions).

The covalent attachment of the methyl group is catalyzed by enzymes of the DNA methyltransferase (DNMT) family (52). These enzymes have the ability to change the conformation of DNA to get access to the cytosine bases. SAM is inserted into the pocket of the enzyme and positioned by a methionin rich loop within the pocket increasing the affinity of the enzyme to the DNA 900-fold. By flipping the target cytosine 180° out of the helix the base looses the stacking interaction with its neighboring base and can also fit into the pocket of the active site of the enzyme (53). Additionally, the flipping mechanism minimizes the distance between the methyl group donor and methyl group acceptor (54). This change of conformation allows interactions between the catalytic side Cys81 of the enzyme and the cytosine. By interaction of the sulfur of Cys81 of the enzyme with C6 of the cytosine a covalent intermediate is formed and the methylation process becomes possible. 5-methyl-2'-deoxycytidine is formed by the attack of the formed carbanion to the methyl group of SAM. After the methyl transfer is performed a proton is abstracted from C5 of the cytosine base followed by a beta-elimination of enzyme-DNA educt and product dissociation (53, 55). Chemical structure formula of the educt and the product are shown in Figure 3.

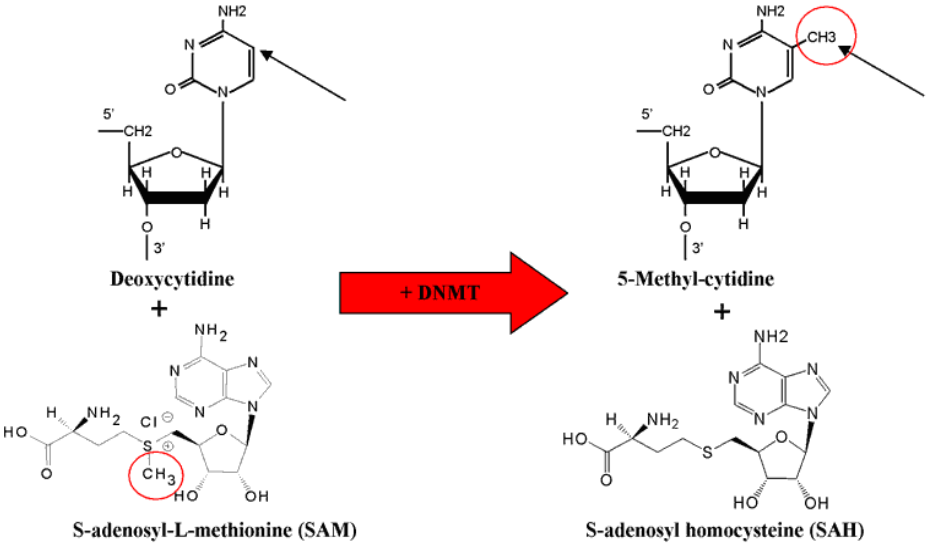

Figure 3
Transfer of a methyl group from a universal methyl group donor to a cytosine base is called DNA methylation; modified from (56)

5.1.2.3 DNA methyltransferases

Transfer of methyl groups from a universal donor to DNA is catalyzed by enzymes of the DNMT family (52). In mammalian cells three members of the DNMT family are known: DNMT1, DNMT2 and DNMT3. DNMT1 measures 193.5 kD and has the ability to interact with methylated as well as with unmethylated cytosines catalyzing both *de novo* methylation and *maintenance* methylation. However, DNMT1 shows in general 5 to 30-fold preference for hemimethylated DNA stands determining *maintenance* methylation during replication as its main function (57, 58).

Structural analyses of DNMT1 revealed that the C-terminus contains the catalytic domain, which shows a close relation to bacterial restriction methyltransferase than to other DNA methyltransferases in mammalians (59). Multiple domains important for different functions specific for eukaryotes are located near the N-terminus of DNMT1. For example import into the nucleus, targeting to the transcription loci or coordination and methylation during S-phase are regulated by this region of the protein. DNMT1 is essential for imprinted genes (49, 50). During the process of maintaining the methylation pattern in the newly synthesized DNA strand, DNMT1 can interact with a variety of co-proteins elevating the efficiency and the specificity of the enzyme.

DNMT1 is co-localized to the region of replication (60). By binding to “proliferative cell nuclear antigen” (PCNA) the ability to form a physical interaction to the DNA is heightened and the *maintenance* methylation capacity of DNMT1 is increased (61). The binding of the co-protein has positive impact on the methylation frequency but studies demonstrated that its presence is not absolutely necessary for the process of *maintenance* methylation.

For a long time DNMT1 has been the only DNA methyltransferase identified in mammals. In 1998 expressed sequence tag (EST) database searches revealed another DNA methyltransferase called DNMT2 (62). Studies showed that knockout mice for DNMT2 had neither deficiency in DNA methylation in embryonic stem cell nor was DNA methylation of newly integrated retroviral DNA impossible what raised the question of the presence of methylation catalyzing activity within this protein (63). The physiological function is so far not completely understood. Only three facts were determined: Firstly, DNMT2 shows structures known from other DNA methylating enzymes. Secondly, DNMT2 has a DNA binding domain and thirdly, DNMT2 only comprises a catalytic domain but no regulatory domain (64). These findings allow the hypothesis that DNMT2 is a member of the methyltransferanse family but research is far from understanding completely its physiologic function. Well conserved DNMT2 homologues were found in plants, vertebrates, D. Melanogaster and S. pompe proposing DNMT2 to be the most conserved member of the DNMT family throughout the phylogenetic tree (65).

Afterwards, members of the DNMT3 enzyme family were identified in the EST data base (DNMT3a, DNMT3b and DNMT3L) (66). DNMT3a and DNMT3b are essential for *de novo* methylation in mammals during development but can also lead to aberrant methylation patterns in cancer cells (67). In general DNMT3a and DNMT3b are highly expressed in undifferentiated embryonic stem cells but show minimal expression in differentiated somatic cells (68). DNMT3L is a protein related to DNMT3a and DNMT3b. No enzymatic activity can be found within DNMT3L, however by physically interaction with the enzymes DNMT3a and DNMT3b it modulates the catalytic activity of these proteins (69). Schematic structures of the DNMT family members are depictured in Figure 4.

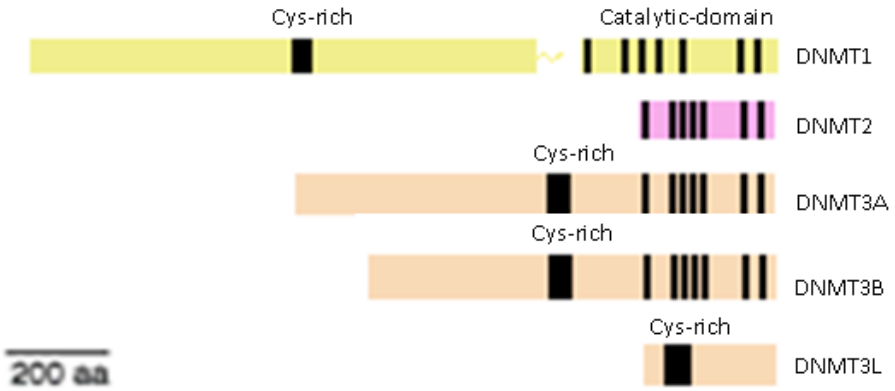

Figure 4
Known members of the DNMT super-family are summarized. DNMT1 is essential for *maintenance* methylation. Function of DNMT2 is so far unknown. DNMT3A andDNMT3B are essential for *de novo* methylation. DNMT3L is a co-factor protein for DNMT3A and DNMT3B.

5.2 Effects of DNA methylation on gene transcription

DNA methylation can alter gene expression either directly or indirectly. Direct inhibition of gene transcription can be found in genes which show a tissue specific gene expression pattern. The methyl group prevents the binding of transcription factors like MLTF, E2F or CREB due to steric interaction with the DNA binding proteins (70, 71). Iguchi-Ariga et al. (72) showed that a single methyl group on one CpG site within the CRE binding domain directly prevents the binding of the transcription factor CREB and inhibits the transcription. Apart from these cis-acting inhibition mechanisms, trans-acting ones are known. SP1 binding site for example attracts apart from SP1 other proteins, which make DNA methylation impossible (73). This “protection” can often be found in housekeeping genes (see Figure 5).

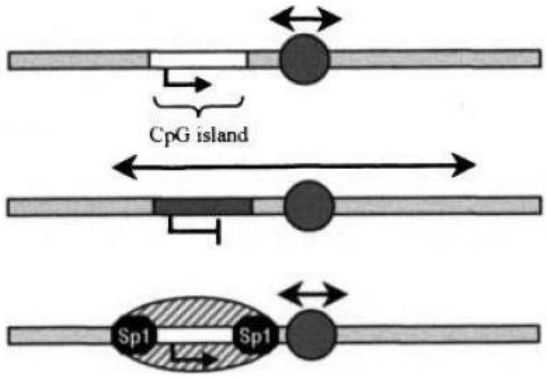

Figure 5
Promoter region with CpG island, without protection methylation spreads over the whole CpG island. SP1 binding site attracts "protection" proteins making methylation spreading over CpG island impossible; modified from (73)

For the indirect inhibition two protein super-families are needed: Kaiso like proteins and Methyl-CpG binding proteins (MBD) (74).

Kaiso like proteins interact physically with methylated cytosines via multiple zinc finger motives (75). A recent report showed that two proteins of this super-family, ZBTB4 and ZBTB38 bind to single methylated CpGs and inhibit transcription (76).

Methyl-CpG-binding protein 2 (MeCP2) was the first identified protein of the MBD family (77). It contains a N-terminal methyl-CpG binding domain (MBD) and a C-terminal transcriptional repression domain (TRD) (78, 79). Immunoprecipitation experiments revealed that TRD interacts with Sin3A (a transcriptional co-repressor) in mammals, which is part of a large inhibitory complex also comprising histone deacetylases HDAC1 and HDAC2 (80). Similarly HDAC1 and HDAC2 are attracted to DNA, chromatin formation becomes more condensed by eliminating acetyl groups and the expression of the affected gene is reduced (Figure 6) (80). Six members of the MBD family are known so far: MeCP1 (74), MeCP2 (81), MBD1 (82), MBD2 (83), MBD3 (83) and MBD4 (84). All these proteins show a high conservation throughout the vertebrates. Apart from the other family members MBD3 shows a mutation of its methyl-cytosine binding domain, preventing selective recognition of methylated DNA (85). However, MBD3 is of importance for gene expression regulation by recruiting co-repressor complexes to methylated DNA (86). The presence of MBD3 in mice cells was related with a higher frequency of adenocarcinomas suggesting that this protein plays an important role in carcinogenesis (87).

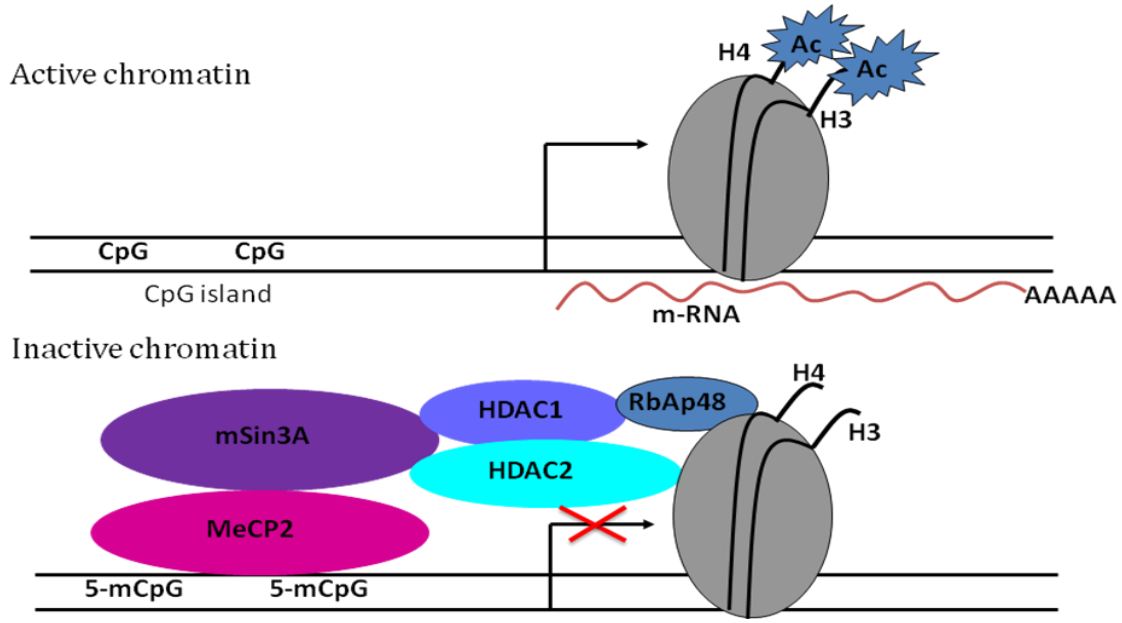

Figure 6
Methylated cytosines are recognized by MBD proteins. These interact with mSin3A, which is part of a great complex comprising HDAC. Contact between methylated DNA and HDACs is achieved; the chromatin formation tightens up. Modified from (36)

5.3 Epigenetic changes are reversible

Epigenetic changes are dynamic and reversible processes. Epigenetics takes part in facilitating the wide diversity of cell types (88). Besides, its physiologic function epigenetic changes can appear aberrantly during aging and development causing expression disequilibrium what may lead to pathologies (64, 65).

Research revealed that epigenetic changes can be controlled by intrinsic signals but also by external induction (89, 90). *In vitro* DNA methylation or histone protein modification can be pharmacologically reversed (91). Often used DNA de-methylating agents are 5-aza-2'-deoxycytidine (Aza-dC) and 5-azacytidine (5-AzaCR). Both of them are base analogous for deoxycytidine. Due to a modification on position 5 in the pyrimidine ring methyl groups can no longer be added, making maintenance of the DNA methylation pattern impossible (92). In these analogues the 5’ carbon is replaced by nitrogen (see Figure 7 and Figure 8). Before incorporation into DNA or RNA the base analogues are phosphorylated to form the nucleoside triphosphate (93).

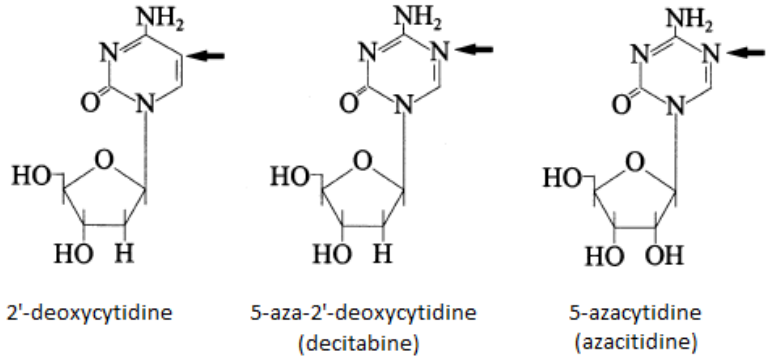

Figure 7
Structural formula of 2'-deoxycytidine and of frequently used base analogous 5-aza-2'-deoxycytidene and 5-azacytidine. Modification of the 5' position in the pyrimidine ring inhibits *maintenance* methylation.

Aza-dC is only incorporated into DNA but not into RNA and shows a higher cytotoxicity than 5-AzaCR. Incorporated in DNA the analogue binds covalently to different members of the DNMT family by forming unresolved thioether bond at the C6 (94). Due to this depletion of DNMT it is unavailable for *maintenance* methylation resulting in significant demethylation after repeated replication (95).

It has been shown that expression of several tumor suppressor genes (TSG) which are often inhibited by DNA methylation in cancer cells (e.g. *p16*, *RASSF1* and *CDH1*) can be restored by adding Aza-dC or 5-AzaCR *in vitro* (65-67). In clinical trials the anti-cancer function of Aza-dC was shown in acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), chronic myelocytic leukemia (CML) and hemoglobinopathies, however its efficacy in solid tumors is poor (96).

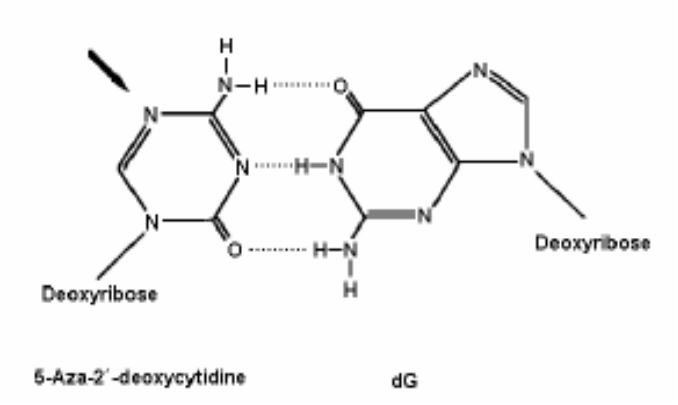

Figure 8
Base analogues are incorporated into DNA forming an unresolved thioether bond with DNMTs leading to unavailability of DNMTs. Maintenance of the methylation is impossible in the consequent duplication steps.

Besides DNA methylation also histone protein modifications can be induced or reversed by epigenetically active drugs. A commonly used one is trichostatin A (TSA) which inhibits the histone protein deacetylation (for the chemical structure formula see Figure 9) (97). TSA is an hydroxamic acid and the best studied histone deacetylase inhibitor (98). It forms a complex with a zinc atom at the base of the catalytic site of histone deacetylases and prevents interaction between the enzyme and the substrate (99). Thus, TSA inhibits removal of the acetyl groups resulting in active gene expression (100).

Pharmacologic manipulation with TSA leads to limited tumor cell growth *in vivo* with little or no toxicity (101). In addition histone deacetylase inhibiters block angiogenesis *in vitro* and *in vivo* due to normalization of the acetylation pattern of the TSG *p53* and *VHL*. Additional studies are necessary to identify a potential impact of TSA for cancer therapy.

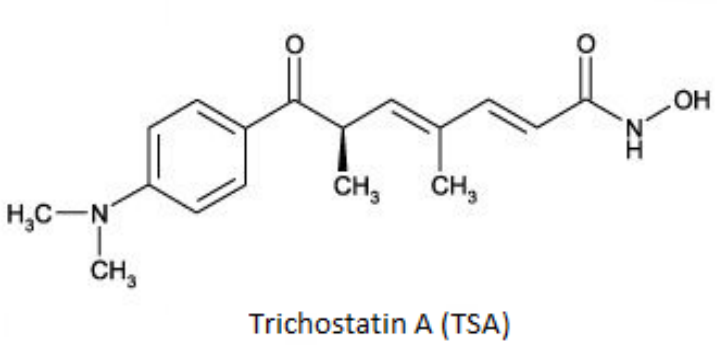

Figure 9
Chemical structure of the commonly used histone protein deacetylation agent TSA. By blocking the histone deacetylases TSA prevents the de-acetylation and thereby keeps the gene expression active.

DNA methylation and histone protein modifications strongly influence each other. Due to this close relation synergistic effects of DNMT inhibitors and HDAC inhibitors have been identified (102).

5.4 Lung cancer

Lung cancer is the number one cause of cancer deaths, killing one million people worldwide each year. Smoking is one of the leading causes of premature death and can be set in relation to 85% of all diagnosed lung cancer cases making tobacco smoke to the number one carcinogen for the lung (103, 104). Moreover length and heaviness of the smoking habit can be set in relation to probability of coming down with lung cancer. For a long life smoker the probability for lung cancer is 20-fold higher than the probability for a never smoker. The lung cancer risk after stopping smoking diminishes but only after 5 years of non-smoking and is never as low as the risk of a never smoker. However not only mainstream smoke but also second-hand smoke is known as pulmonary carcinogen (105).

Over 60 carcinogens can be found in tobacco smoke, which induce tumor genesis in vivo (106). Nicotine is not carcinogen itself but interacts with pathways which deregulate cancer related genes (107, 108). Nicotine interacts for example with ERK and PI3-K/mTOR pathways inducing fibronectin and promoting proliferation of NSCLC. Additionally, it is an addictive drug thus being the reason why people go on smoking (109). The best analyzed tobacco carcinogens are members of the polycyclic aromatic hydrocarbons (PAHs), N-nitrosamines and aromatic amines families. Recent reports determined tobacco-specific N-nitrosamines (TSNAs) and PAHs as the most harmful carcinogens in tobacco smoke (110). PAHs are formed due to the incomplete pyrolysis of tobacco leaves (111). TSNA derive from nicotine and related tobacco alkaloids (112). Three TSNAs, namely N-nitrosonornicotine (NNN), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), show high tumor inductive capacity in mice, rats and hamsters (113) (chemical structure formula depictured in Figure 10). They are all formed by nitrosation (N=O group is added) of secondary and tertiary amines and contain the functional group N-N=O.

In addition to the direct carcinogenicity in tobacco smoke smoking favors malignant transformation by acting as direct mucus irritant and inducing inflammation which leads to generation of free oxygen radicals (114).

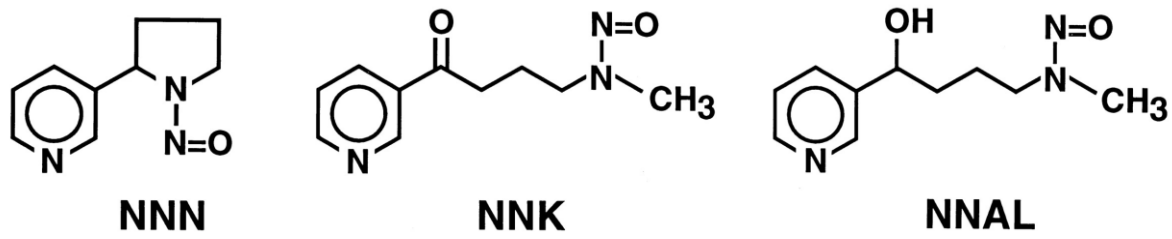

Figure 10
Structural formula of most important carcinogens derived from tobacco smoke; N-nitrosonornicotine (NNN), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL); (115)

For a long time lung cancer was a men's disease. However, in recent years a decrease in the incidence of lung cancer in men, but an increase of the incidence in women was observed. (116).

Lung cancer can be distinguished into two major types: small cell lung cancers (SCLC) and non-small cell lung cancers (NSCLC) (117). Thereby, 85% of all lung cancer patients suffer from NSCLCs (118). Furthermore NSCLCs are divided into three subtypes according their histological properties: adenocarcinomas, squamous cell carcinomas and large cell carcinomas (119). Lung cancer is very often a deadly disease with 5-year survival rates of only 14% (120).

Lung cancer is a multistep mechanism of changes in the genetic and the epigenetic code caused by exposure to carcinogens, like tobacco smoke or other harmful substances, over a long period of time (121). In this disease, like in many others an early prognosis augments the cure-chance drastically. Diagnostic techniques like spiral computed tomography, sputum cytology, histopathology or tumor-node-metastasis classification often failed to allow a sensitive diagnose. In recent years great efforts were taken to identify molecular changes relevant for the pathogenesis of NSCLCs like K-ras, p53 mutation status and microsatellite instability (122, 123). Furthermore DNA methylation was found to be of tremendous importance for development of NSCLCs (124).

5.4.1 DNA methylation as molecular biomarker in NSCLC

Epigenetics is one of the regulation mechanisms for gene expression which can be altered during carcinogenesis (125). In this work we focused on a part of epigenetics, the DNA methylation. Aberrations in DNA methylation patterns can occur early in carcinogenesis and can be found already in pre-carcinogenous lesions in the lung (126). Several genes including *p16*, *RASSF1A*, *APC*, *RARß-2*, *CDH1*, *CDH13*, *DAPK*, and *MGMT* were found to be frequently methylated in primary NSCLCs (127-135). All these genes were found to be associated to cancer related pathways such as apoptosis, cell proliferation or DNA repair.

Recently, genome-wide approaches for DNA methylation analysis were developed.

5.5 DNA methylation analyses

Different approaches have been developed to analyze DNA methylation: 1) Methylation sensitive restriction enzymes (136). 2) Methylation-specific polymerase chain reaction (MSP), 3) combined bisulfite restriction analyses (COBRA) and 4) bisulfite genomic sequencing. Due to the great number of possible targets for methylation in the genome the need for genome-wide techniques became apparent. Techniques, like the methylated DNA immunoprecipitation combined with microarray analyses (MeDIP-chip) or approaches with epigenetically active drugs combined with gene expression microarrays which allow the analysis of ~29.000 genes at a time, has become available over the last years (124).

5.5.1 Sodium bisulfite treatment

Most of the PCR based approaches for DNA methylation analysis are only applicable for sodium bisulfite treated DNA. During common PCR reactions methylation signals are lost, due to the fact that primers cannot distinguish between methylated and unmethylated cytosine. However, sodium bisulfite treatment makes distinguishing between methylated and unmethylated alleles possible (137). This approach was first identified in 1970 by Hayatsu et al. (138). Sodium bisulfite treatment can be divided into four steps: denaturation, addition of sodium bisulfite to the double bond between the position 5-6 of the cytosine, deamination and desulphonation by subsequent alkali treatment. During that treatment

unmethylated cytosines are deaminated to uraciles. In the next replication step uraciles are replicated as thymines. Due to the fact that the rate of deamination of unmethylated cytosines to uraciles is much higher than the deamination of methylated cytosines to thymines all remaining cytosines after the treatment are derived from methylated cytosine. Performed under proper condition (pH 4.8-5.8, 3.0-4.0 M sodium bisulfite concentration and 60°C temperature) the conversion rate of unmethylated cytosines to uraciles is 99% (137, 139-141).

A graphic summary of the process is provided in Figure 11. In ds DNA the base pairing locks the base cytosine in the inactive *anti* conformation. Single strand DNA provides the required *syn* conformation. Only in the *syn* conformation interaction between the position 6 of the pyrimidine ring and sodium bisulfate is possible (142, 143). Thus, denaturation of the template DNA is necessary.

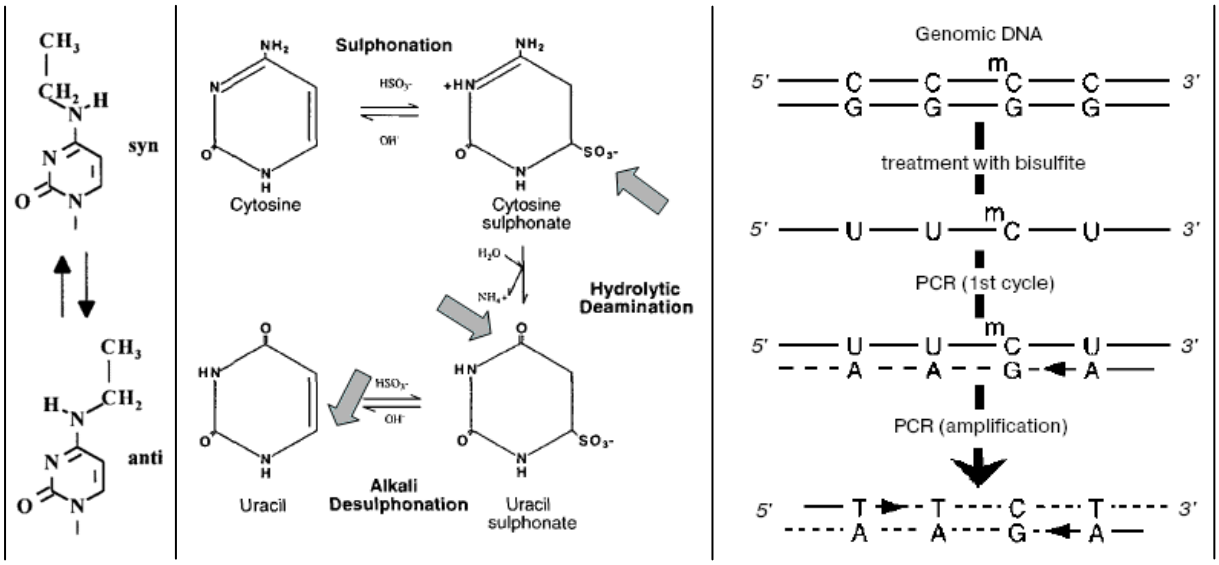

Figure 11
Incubation of DNA with sodium bisulfite requires DNA bases in *syn* conformation. Unmethylated cytosines are converted to uraciles in four steps. Uraciles are converted to thymines in the consequent PCR step. Only methylated cytosines remain cytosines after sodium bisulfite treatment. Modified from (144)

5.5.2 PCR based techniques for analyzing DNA methylation

Two different strategies are used for PCR based DNA methylation analysis. First, use of methylation-independent PCR primers (MIP). For this analysis only one primer set is designed which binds to both the methylated and the unmethylated allele. Due to the fact

that unmethylated sequences relatively poor of CG dinucleotides are amplified in a higher rate than methylated ones a PCR bias is only preventable by designing primers without any CG dinucleotides. An example for this approach is the COBRA technique. After bisulfite treatment of genomic DNA amplification with MIP primers is performed. Due to the bisulfite treatment new methylation-dependent restriction sites arise and methylation-depending retention of pre-existing sites can occur (145). If a cytosine is methylated it remains a cytosine during the sodium bisulfite treatment, an unmethylated cytosine is converted into a thymine and the restriction site is lost or a new one is formed (146). Subsequently, digestion with restriction enzymes specific for sites containing CpGs makes the distinction between methylated and unmethylated regions possible (147). The results are visualized with PAGE gel equipment.

Second, use of methylation-specific PCR primers (MSP) (see Figure 12). The MSP approach is one of the most used and most efficient techniques to analyze the methylation status of specific genomic regions. Genomic DNA is firstly treated with sodium bisulfite to allow a distinction between methylated and unmethylated alleles. Primers specific for either bisulfite converted methylated or bisulfite converted unmethylated region are design. The primer design step is the most critical one within this technique. It has to be guaranteed that primers only form duplexes with the region of interest but not unspecifically with other regions or other primers. The primers have to contain as many CpGs as possible, minimally 2 CpGs and the CpGs have to be as near as possible at the 3' end of the primer. The PCR program has to be as stringent as possible, which means high annealing temperatures and low cycle numbers to avoid false positive results (148). During PCR with primers specific for the bisulfite converted methylated allele only DNA with methylated primer-binding site can be amplified and visualized by a gel-electrophoresis equipment.

MSP is a fast and precise analysis which the analyses can be performed in a short time period. In addition only a small amount of DNA is needed. However, the MSP assay is only a qualitative but not a quantitative analysis.

To perform quantitative analyses fluorescence-based, quantitative real-time PCR approaches can be used. By intercalation of the fluorescent dye into the newly synthesized ds DNA the amount of produced DNA can be measured after every cycle of elongation (149).

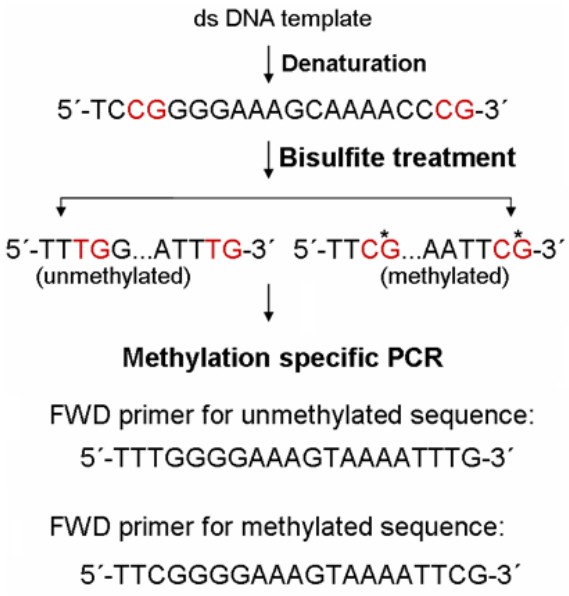

Figure 12
Schematic description of a MSP work protocol. Firstly, genomic DNA samples are treated with sodium bisulfite to convert unmethylated cytosines into thymines. Secondly, primers are designed for the methylated as well as for the unmethylated allele. By PCR amplification the methylation status of one specific gene is identified.

5.5.3 DNA methylation analysis on a whole genome level

Computer analyses revealed that approximately 29.000 CpGs are present in the human genome (46). On average 600 of these CpG islands are aberrantly methylated in a tumor (15). To analyze this huge amount of CpG islands without time consuming single loci determination of methylation patterns, whole genome approaches have been established during the last years. The first whole genome approach was restriction landmark genomic scanning (RLGS). With RLGS Dai et al. (150) identified 11 genes with different methylation patterns in lung tumor samples compared to corresponding non-malignant lung samples.

As mentioned in section 5.3 DNA methylation is reversible *in vivo* and *in vitro*. This fact can be used for whole genome approaches. By treatment of cells with DNMT inhibiting drugs like Aza-dC or 5-AzaC DNA methylation is reversible. Due to interaction between DNA methylation and histone deacetylation, a synergistic effect between DNMT inhibitors and HDAC inhibitors was observed *in vitro* (151, 152). Gene expression patterns of cells which are treated with demethylating and histone deacetylation inhibiting drugs and patterns of untreated cells are compared by expression microarray analyses. Expression of methylated genes is up-regulated by drug treatment compared to untreated cells (120, 153). Thus, genes

that are not expressed before treatment and expressed after treatment are genes of interest for further gene-specific methylation analyses. Using this approach Shames et al. (154) identified several unknown targets for aberrant methylation (e.g. *BNC1, LOX*) in NSCLC cell lines. The major disadvantage of this approach is that it can only be applied *in vitro*. It is known that methylation can accumulate in cell culture (120). Thus, results obtained from this approach have to be approved in primary tissue samples to be able to exclude artifacts caused by *in vitro* culturing. Additionally, treatment with DNMT inhibitors and HDAC inhibitors activates several stress responses in treated cells. Verification of the obtained results is an important step to exclude false positives from the results.

Recently, a new method to detect genome-wide DNA methylation has been described by Weber et al. (155). Based on combination of immunoprecipitation and microarray analyses the methylation status of all annotated genes can be identified. Genomic DNA is fragmented by sonication or restriction enzyme digestion obtaining fragments with a length of optimally between 200bp and 800bp. With an antibody specific for 5-methylcytosine methylated DNA fragments are enriched in an immunoprecipitated sample (IP) and separated from the unmethylated fragments. The IP sample is hybridized to a microarray chip together with a control sample containing the whole genomic DNA (Input). By labeling IP and Input with different fluorescent dyes the two samples will be compared. This new technique allows an unbiased detection of DNA methylation on a whole genome level. Using this approach Weber et al. (155) identified a great variety of genes which are targets for aberrant methylation in transformed cells like *FOF1*, *PAX6*, *TGFB2* or *ADAM12*.

5.6 DNA methylation and its impact on NSCLC

For a long time the origin of cancer has been put down to mutations and other alterations in the genomic sequence. However, during the last decade it became apparent that NSCLC arises from a multi step process implying aberration not only in the genetic but also in the epigenetic code (156). Particularly aberrations in DNA methylation patterns were found in NSCLCs. Promoter hypermethylation of TSGs occurs throughout the tumor genesis and can even be found in pre-carcinogenous lesions suggesting hypermethylation of certain genes to be one of the earliest hits in the pathogenesis of a tumor (157).

Methylation of certain genes can be detected in blood samples, in exfoliative material of the aero digestive tract epithelium and in sputum from lung cancer patients (120). It has been hypothesized that samples which are obtained through non invasive or minimal invasive methods could be of use for predicting possible disease recurrence after surgery or even an early detection of lung cancer (11, 158). Clinical studies on this subject are necessary. Using serum samples for analyses of the methylation status is controversial, since methylation frequency of genes detected in serum samples is lower than frequencies of methylation of the same genes analyzed in the corresponding primary lung tumor tissue. However, only genes that are methylated in the primary tumor show methylation in the serum sample (9). By contrast, it has been reported that sputum samples seem to be an adequate possibility (13).

During the last years a respectable number of genes could be identified which show tumor-specific methylation pattern. For example the methylation status of *RASSF1A*, *p16*, *CDH1*, *DAPK*, *APC, RARβ-2* and *MGMT* was intensively studied revealing that these genes undergo changes in their methylation pattern during the pathogenesis of NSCLC. TSGs which are methylated in tumor samples were often associated to pathways important for cell cycle regulation (e.g. *p16* and *RASSF1A*), regulation of cell adhesion (e.g. *CDH1*), regulation of growth (e.g. *APC and RARβ-2*) regulation of apoptosis (e.g. *DAPK*) or DNA repair (e.g. *MGMT*) (128-131).

Toyooka et al. (159) showed that methylation of *APC*, *CDH13* and *p16* differ between NSCLCs samples and neuroendocrine tumors showing a higher methylation rate in NSCLCs. *RASSF1A* on the contrary showed higher methylation frequency in neuroendocrine tumors. These data suggest that DNA methylation is tumor type-specific. Furthermore, it has been observed that aberration of the methylation pattern is even specific for the two main subtypes of NSCLC, adenocarcinomas and squamous cell carcinomas. *P16* for example is more frequently methylated in squamous cell carcinomas than in adenocarcinomas, whereas methylation of *APC* and *CDH13* occurs with a higher frequency in adenocarcinomas (160, 161).

The methylation frequency of *p16* is not only cancer type specific but can also be set in relation to the smoking habit of patients suffering from this cancer type. In general,

exposure to tobacco-related carcinogens increases the methylation frequency of *p16* which can even be found in non-carcinogenous lung tissue samples of ever smokers. These findings suggest that changes in the *p16* methylation pattern are one of the earliest events in the pathology of NSCLC (114, 126). Comparative analyses by Belinsky et al. (162) of *p16* and *DAPK* methylation revealed that changes in the methylation pattern occurring due to exposure to tobacco smoke persists for a long time even after stopping smoking.

Elevation of methylation frequency in cells after exposure to tobacco smoke can be due to enrichment of DNMT1. By working *in vitro* with immortalized bronchial epithelial cells Damiani et al. (163) showed that elevation of DNMT1 level after benzo(a)pyrene-diolepoxide 1 and/or methylnitrosurea exposure maybe of two reasons. Firstly, *DNMT1* expression is elevated after tobacco-smoke induction. Secondly, degradation of DNMT1 is inhibited due to over expression of the spindle check point protein MAD2.

Several studies suggested the methylation pattern is of prognostic relevance (164-166). *DAPK* methylation was found to be associated with a poor overall and disease-specific survival, due to a higher invasiveness of tumors after inactivation of *DAPK* (164). Interestingly, *CDH1* methylation correlated with a higher overall survival, probably due to a highly dynamic methylation and demethylation process during invasion of tumor cells into new tissues (167).

Disease recurrence occurs frequently in NSCLC patients. Brock et al. (168) analyzed the methylation pattern of genes including *p16*, *MGMT*, *DAPK*, *RASSF1A*, *CDH13*, *APC* and *ASC* and compared the results with clinico-pathological findings of the patients. Methylation of *p16, RASSF1A*, *CDH13* and *APC* in the primary tumors and corresponding tumor-negative lymph nodes was found to be associated with a high probability of recurrence. These findings suggest that methylation analyses might be helpful determining risk assessment for disease recurrence in NSCLC patients.

Moreover, several studies revealed that due to the reversibility of DNA methylation it is a potential target for anti-cancer therapy. Treatment with DNMT1 inhibiting agents such as Aza-dC or 5-AzaC in addition to HDAC inhibitors showed *in vitro* anti-carcinogenic attributes. Thus, by methylation inhibited TSGs can be reactivated and their expression level can be restored. 5-AzaC is used clinically for the treatment of patients suffering from certain types

leukemia and myelodysplastic syndromes (93, 169). However, currently 5-AzaC is not used for solid tumors. Juergens et al. (170) reported some benefits for NSCLC patients treated with 5-AzaC in a phase-II trial.

6 MATERIAL AND METHODS

6.1 Cell culture

Materials:

Growth Medium RPMI 1640+Glutamax	(Gibco; #145238)
Trypsin EDTA	(Gibco; #197812)
FCS	(Gibco; #179822)
Refobacin	(Merck; #112214)

It is necessary to work in a sterile surrounding to avoid contamination. Thus, the work had always to be performed in sterile bench with gloves. Before starting the work the sterile bench had to be cleaned, the UV-light had to be turned on for approximately 30 minutes and the cell culture medium had to be prepared. Therefore, 50ml FCS and 40mg of the antibiotic refobacin were added to 500ml growth medium. The growth medium can be stored at 4°C and had to be warmed up previous to each use (37°C). Cells were grown in cell culture flasks (BD Falcon, 750ml, and straight neck).

First of all, cells had to be washed. The old medium was decanted, then 5ml trypsin were added and incubated until cells start to detach from the flask ground. Then 20ml medium were added to stop the detachment reaction. Cells were washed of the surface and transferred into a 50ml falcon tube. Afterwards, cells were centrifuged for 5 minutes at 2000rpm, the medium was decanted and 20ml fresh growth medium were added to start a second washing step. Then cells were ready for DNA extraction-, cell passaging- or cell freezing-process.

6.1.1 Cell passaging

After the washing step cells were resuspended in new growth medium in the correct concentration to allow an optimal growth and were transferred to a new culture flask. For a proper growth, lung adenocarcinoma cell lines A549, NCI-H1993, NCI-2073 and NHBE cells

were incubated at 37°C, 5 % CO2 and 96% humidity in an incubator. Growth medium was changed every 3-4 days. Cells were grown to 80% confluence before the cells were split the next time.

6.1.2 Cell freezing

Materials:

Freezing Medium:	Growth medium + 5% DMSO	(Gibco; #145238)
	Cryotubes	(NUNC)

At a proper confluence of about 80%, cells were washed, resuspended in 3ml freezing medium and aliquots were made in 2ml cryotubes. Afterwards cells were stored in a cell-freezing box (filled with methanol to allow a slow freezing process of 1°C per hour) at -80°C over night and then they were stored at -195°C in liquid nitrogen.

6.1.3 Cell de-freezing

Cryotubes were heated at 37°C until the ice was gone. As soon as cells and the medium were defrosted the freezing medium was toxic for cells, so the following steps had to be performed as fast as possible. Immediately after the ice was completely melted, the cells had to be washed with 15ml growth medium. The medium was decanted and cells were resuspended in 5ml growth medium and transferred into a new culture flask for the growth process in the incubator. After 24 hours the start of the cell growth was controlled. If cells started to grow, the medium had to be changed and cells were bred until the needed confluence was reached.

6.1.4 Incubation of cells with 5-aza-2'-deoxycytidine

Materials:

5-aza-2'-deoxycytidine min. 95%	(Sigma Aldrich; #A3656)

At the beginning a stock solution was made out of 5mg Aza-dC and 250µl 1xPBS and then 20µl aliquots of this solution were stored at -80°C. Before starting the Aza-dC

incubation one aliquot was defrosted and 180µl growth medium was added. $5x10^6$ cells were grown in a $175cm^2$ cell culture flask. From the 200µl solution 25µl were again diluted with 25ml growth medium and after that they were added to cells. Due to instability of the cytosine-analogue in the growth medium 25µl of Aza-dC dilution had to be added every day for the duration of four days. If necessary the growth medium was changed and cells were split. After the four day incubation period cells were bred in a medium without Aza-dC for two or three days so that they regenerate and then RNA could be extracted (see 6.2.2)

6.2 Nucleic acid isolation

6.2.1 Isolation of genomic DNA

Materials:

PK-buffer (for ingredients see 6.6)	
Phenol S/P saturated	(Amresco; #K168-400)
Isoamyl-alcohol	(Sigma-Aldrich; #C549-10T)
Chloroform	(Sigma-Aldrich; #C-5312)
Sodium Acetate	(Sigma-Aldrich; #S9513)
EtOH (abs.)	(VWR; #20821.310)
cell scraper	(Falcon)

DNA was isolated from cell lines as well as from primary lung tumor tissue samples and corresponding non-malignant lung tissue samples using standard phenol/chloroform extraction.

Cells were bred in the culture medium until the flask showed 80% confluence. Then cells were scratched from the surface with a cell scraper (Falcon), transferred to a 1.5ml tube and incubated over night with 800µl PK-buffer at 55°C.

The tissue samples were shredded in liquid nitrogen with the help of a mortar. Small tissue fragments were transferred into a 1.5ml tube containing 800µl PK-buffer and incubated over night at 55°C. The consequent steps can be performed for cell material obtained from cell culture as well as from primary tumor samples. On the next day the

samples were centrifuged to collect the liquids from the cap of the tube. The cell/buffer solution was transferred to two 2ml Phase Lock Gel Light tube, 400µl phenol were added to each tube and after short incubation the samples were centrifuged for 5 minutes at 14000rpm. Through this step proteins were separated from DNA. Two phases were visible after the centrifugation. Due to the polar negatively charge of DNA, DNA was solved in the aquatic phase (the upper one). This one was decanted after the centrifugation step and transferred into a new 2ml Phase Lock Gel Light tube. To increase the purity of DNA an additional phenol extraction step was performed. Again after centrifugation the upper phase was transferred to a new 2ml Phase Lock Gel Light tube, 480µl chloroform as well as 20µl isoamyl-alcohol were added and the sample was again mixed and centrifuged for 5 minutes at 14000 rpm. After taking off the upper phase and transferring it to a 1.5ml tube, 1.5 volumes of EtOH (abs.) and 50µl 3M (pH 5.2) sodium acetate were added to adjust the salt concentration and mixed thoroughly. The precipitated DNA became clearly visible. Finally, sample was put on -20°C over night. On the next day the sample was centrifuged for 15 minutes on maximum speed (18000rpm) and at 4°C in a pre-cooled centrifuge. A pellet became visible. EtOH was discarded, then 600µl EtOH (70%) were added and again centrifuged for 10 minutes on maximum speed and at 4°C. Again EtOH (70%) was discarded, the sample was shortly centrifuged and the rest of EtOH (70%) was taken off with a pipette. The pellet was dried at 37°C for 5- 10 minutes. Depending on the size of the pellet 100-200µl ddH_2O were added. The concentration was measured with an “Eppendorf BioPhotometer”.

6.2.2 Isolation of total RNA

Materials:	Trizol reagent 15596-026	(Gibco; #149789)
	Chloroform	(Sigma-Aldrich; #C-5312)
	EtOH	(Merck; #1.00983.1000)

Cells were transferred to a 50ml tube and washed twice. The cell pellet was resuspended in 1ml trizol, transferred to a 1.5ml tube and incubated for 5 minutes at room temperature. 200µl chloroform were added for each ml trizol and the sample was mixed per hand very thoroughly. After 3 minutes of incubation the phases were separated from each other by

centrifuging the sample for 15 minutes at 11500rpm and 4°C. The upper phase was transferred into a new 1.5ml tube, 500µl EtOH (abs.) were added and incubated for 10 minutes at room temperature. After centrifuging (10 minutes at 11500rpm and 4°C) the supernatant was removed and EtOH (70%) was added to wash the pellet. After an additional 5 minutes centrifugation step at 7500rpm and 4°C the supernatant was decanted and the pellet was centrifuged briefly to be able to remove the whole amount of alcohol with a pipette. The pellet was dried at 37°C and then diluted in 50-200µl RNAse-free water depending on the size of the pellet. RNA was stored at -80 °C.

6.2.3 RNAse treatment of extracted DNA

To prevent interference of RNA with DNA which often causes false results DNA was RNAse treated. 1µl RNAse was added to DNA and incubated at 37°C for one hour. After incubation the probes had to be again phenol/chloroform extracted as described in 6.2.1. The only difference is that two hours of incubation in PK-Buffer was sufficient for digesting the proteins in the sample.

6.3 Methylation-specific PCR (MSP)

6.3.1 Sodium bisulfite treatment

In order to perform MSP analysis DNA samples had to be treated with sodium bisulfite. Unmethylated cytosines but not methylated ones were converted to uraciles. Accordingly they were converted to thymines in the following PCR step resulting in different DNA sequences for methylated and unmethylated regions. Sodium bisulfite treatment was performed by using the “EpiTect Bisulfite Kit” (Qiagen; #59104). In brief, 1µg genomic DNA of each sample was treated and eluted in 40µl of the elution buffer. Each step of this approach was performed according to the manufacturer’s instructions.

6.3.2 SssI CpG-methyltransferase treatment of genomic DNA

Materials:	10xNEB Buffer	(Biolabs; # B7002S)	2µl

diluted SAM	(Biolabs; # 9003S)	2µl
DNA		2 µg
*Sss*I methyltransferase	(Biolabs M; # M0226S)	1µl
add ddH_2O to a final volume of		20µl

In order to make sure that the PCR worked properly, a positive and a negative control were added to each analysis. The negative control contained only ddH_2O without any DNA and was required to report contaminations. The positive control was added to make sure that the PCR setup and the primers were working even if all analyzed samples showed a negative outcome. In order to produce the positive control genomic DNA was treated with *Sss*I methyltransferase, an enzyme that catalyzes *de novo* methylation of cytosines on CpG sites by using the methyl group donor S-adenosyl-L-methionin (SAM). This enzyme adds a methyl group to each CpG in the sample DNA. 2µg of the DNA were treated with 2µl 10x NEB buffer, 0.5µl 200x SAM and 12U *Sss*I methyltransferase in a total volume of 20µl. The samples were incubated at 37°C for one hour and consequently for 20 minutes at 65°C.

6.3.3 Primer design

The primer design is one of the most crucial steps during MSP. Thus, for each gene of interest 2-3 primer pairs were designed and tested to find the most efficient and most specific one. Genomic sequence of the genes of interest was obtained from Ensembl database (release 58) and the 5' promoter region of each gene was used for primer design. Using the program "Methyl Primer Express® Software v1.0" CpG islands located in the promoter regions were identified and primers for the methylated and the unmethylated allele were designed. Criteria for CpG island detection were as follows: minimum length of 200bp, 50% GC content and 0.60 ObsCpG/ExpCpG. In case of positive CpG island detection MSP primers were designed using following limitations: 100bp minimum and 250bp maximum length of PCR amplicon, primer length between 18bp and 22bp, Tm reaction between 56°C and 64°C and a minimum of 2 CpG sites per primer. Primer sequences designed for selected genes and the corresponding amplicon length are shown in Table 1.

After receiving the primers they were diluted to obtain a concentration of 1pmol/µl in ddH_2O and accordingly a "primer pair mix" of each primer pair was prepared. That means, 20µl of the diluted forward primer, 20µl of the diluted reverse primer, 160µl ddH_2O were mixed (to obtain a concentration of 0,1pmol/µl) and used for the following MSP analyses.

Using the designed primers and the prepared DNA, MSP was performed to analyze the methylation status of the gene of interest. The appropriate annealing temperature was determined by first, testing all primers with a PCR-program including an annealing condition at 60°C and 37 cycles. If the signal was strong MSP was repeated at a higher temperature and fewer cycles. If the signal was weak the temperature was decreased and the cycle-number increased.

Table 1
Summary of oligonucleotide primer sequences used for MSP analyses and amplicon length

Gene	Primer fwd (5´-3´)	Primer rev (5´-3´)	Amplicon length (bp)
ANKRD13D	AGAATAAGTCGCGTTCGTTC	TTTTCCGCTTAACGAAAAAA	154
AXUD1	GTAGTCGTCGGGTTATAGTTC	TCGTCAATAAAAAACTCCG	158
ALDH1A3	TGTTTATATATAAGGCGCGC	CAAATAATCTCGCCGTTCTA	129
BARHL2	CGATAGTTTTTTTCGTTGGTTC	ACCGAACGAACGAAAATCT	174
CD14	ATACGGTTAAGGTTTTTCGC	CTCGAACGTCAATTCCTTAA	122
CRY4D	TTTGTGCGGTTTTTGTTAAC	GCGATAAAAAAAACTCCGAA	171
CUL4A	CGCGGCGGCGGTTGGGTTCG	GCGCCCCCCGACTTAACGAACG	217
DLX4	AGGGGTGTTCGTGATAGATC	CTCTTCTTAAATCGCCGAAC	182
EOMES	TTGTTTTATATTTGCGTGTGC	AAAACACGTTTCGTTACGAA	215
EXOSC7	CGGGATTCGGTTTCGGGTTTCG	CCGCTCCTACAACCGCGACCG	120
FOXD3	GAGGCGTATCGTTGATTTATC	CAATACACGCGAAATACAAAA	176
FOXD4L1	TCGAGAGTATATCGAGGGC	CGCTAAACGTAAAACGCTTA	185
FOXD4L3	TCGAGAGTATATCGAGGGC	AATACCGCTAAACGTAAAACG	193
GAD1	AATGAAGAACGAGGGGTATC	AACCGACAACTACTCTACGCT	211
GDNF	TGGAATGGCGTATTAATTTC	ACCTCGCGAACTAATTTACA	241
GHSR	TGGTATTTTATCGGGTGAGC	CAAAACGCCGCTAAAAAA	170
GHSR	TGGTATTTTATCGGGTGAGC	CAAAACGCCGCTAAAAAA	170
GIPC1	CGAGAAGGATGTGTTTGTTTC	TAAACCTCCAATCGCGTC	105
GPR50	GTATCGTAGGGGAATTGTGAC	ACGAAAATATCACGCGACTA	197
HAND2	GGCGTTTAATGGTTTTTTTTC	TCGATTCCCCGAATCTATAA	218
HOXA9	GGCGTTTTCGAGTTTTTTAC	CGCAAATACGTCAAAAAAAA	223
HOXD11	GGGTTTTTACGGTAGGTTTTC	GCGCCGATACATAAAAAAA	225
HSPA2	GGTGGTCGTTAGTTGATTTC	CGACGCACGAATAAATAATAC	133

IRX4	TGGGTAGTTGCGAGTTTTTC	TACATAAATCGACGCTCCAA	190
ISL2	AAAGTCGGACGTTATTGGAC	AAAAACGCGATCGCTACTAC	215
KDELR3	GTCGTCGGGTCGTTTAGTTC	ACGACTACCACGTCTCCGA	228
MACF1	GTGTAGATAAAGGCGCGGTC	ACGAAAAAAACCGACCGA	171
NEFH	AGGTTATGATGAGTTTCGGC	CACGAATAAAAACCGCTAAAA	155
OSR1	AGTTCGGAAGTTCGGGTAGAC	CTACCTTCGAAACGAAACCG	191
OTX2	TCGATGTTTTTTTGCGAAC	ACGTTTCGACGACGTTTTAC	183
PBX3	TAGAGTTTAATGAGCGCGC	CGCGAACCGAACTAAAAA	163
POU4F2	AGATTGATAGTAGAGGCGGC	ACTAATCCCGACTACCGAAA	156
PRLHR	GAAGCGAGATCGGATTATC	TATTTACGATCCTCGCAATC	158
REC8L1	AGTTTGACGTATTACGTGGC	CGTTCGAATCTAAATCGAAA	229
SHOX2	TAGTTAATATGGCGTGGGC	CCGATATTATACCGCACAAA	160
SLFN13	GCGTTAGTGAGAATTTGGTC	CGACGACCGATACTCTATACTC	202
SLFN13	CGCGGTGTTTGGTTAATTTC	GAATCGAACGACACTCGATTC	247
SMO	GGGGAGTCGTGTATTTCGTTTCG	CGCCCCCGCCTACTCCTCG	170
STX16	TTTAGTTGTATAACGAACGCGC	GTACGAAACCCCCAAACG	102
TAL1	GGGGTGTGTGGTTTTTGC	CCGACTTAATCAACCCCG	115
THSD4	GGTTTCGTAGGAGTCGTAGC	CGACCCGTCTTAACCCTATA	222
TMTC1	TGAAAAGTCGGGAATTTTTC	CCCACTTACGATATCCGACT	155
ULBP1	TGGTTTTTTCGAATATCGC	CTTCAACCGAACATATTCGA	154
WFIKKN2	TGTGCGATTTTGGATAAGTC	CGACGACTAAAACGAATCAC	229
XRCC3	CGGAGTTTGAATTTCGTTTTATC	CCTTCTCCTCCAATAAACG	187
ZFYVE21	CGGTTCGTAGTTCGGTTCGGC	CCAATCCGTATAAAAACACCG	232
ZNF667	TACGTTTAGATGATCGGGAC	TATTACGCCTACGTAACCGA	221

6.3.4 MSP analysis

Materials:

Master mix (for one reaction):	ddH_2O		13,2µl
	dNTPs	(BIOLINE; #39025)	3,2µl
	PCR-Buffer	(QIAGEN;#203205)	2,5µl
	"Primer pair mix" specific		1,0µl
	for the gene of interest		
	Hot start TAQ polymerase	(QIAGEN; #203205)	0,2µl

Sodium bisulfite treated DNA, Agarose LE (Promega; #V-3125), Electrophorese buffer (sterile filtrated): 1x TAE (Tris/Acetic/EDTA Buffer; BIO-RAD; # 1610743), gel-red 10000x (Biotium; #41003/41003-1); Bromphenol blue (Sigma-Aldrich; #B-5525), 100bp DNA size marker (INVITROGEN; #15628-019), Loading buffer (0.05% bromphenol blue, 40% saccarose, 0.1M EDTA, 0.5% SDS and ddH_2O to a final volume of 10ml; the buffer has to be sterile-filtrated)

Ingredients for the MSP master mix (stored at -4°C) were defrosted and the PCR-tubes (200µl capacity) were marked. Then 20µl of master mix and 25ng DNA were transferred in a PCR-tube. After mixing the tubes were placed in the PCR cycler (GeneAmp PCR System 9700). To analyze MSP results, an agarose gel-electrophoresis was performed. An agarose gel was prepared out of 180ml 1xTAE and 3.6g agarose (to obtain a 2% agarose gel). After boiling in the microwave the agarose was cooled with cold water, 3µl gel-red were added and the gel-tray was filled with the agarose gel. By cooling the gel to room temperature the gel ingredients polymerized. 2µl loading buffer were added to all PCR samples and 12µl from that mixture were loaded onto the gel. In addition, 3µl of a 100bp size marker were loaded onto the gel. The separation process took approximately 30 minutes at 170 volt. Due to the ability of gel-red to fluorescence under UV-light the results became visible.

6.4 Expression microarray analysis

Isolated RNA was purified before the use of Affymetrix U133 plus 2.0 arrays.

6.4.1 RNA clean-up with the MinElute Kit (Qiagen)

Each step of this approach was performed according to the manufacturer's instructions. 100µg RNA were diluted in RNAse-free water to a total volume of 100µl, 350µl RLT Buffer were added and mixed by vortexing the sample. Afterwards 250µl EtOH (abs.) were added and mixed with a pipette. 700µl from this solution were transferred onto a "MinElute" column for a 22 minutes centrifugation step at 10000rpm. Afterwards the flow through was again transferred onto the membrane of the column for a second centrifugation step to increase the amount of bound RNA. Accordingly the column was transferred into a new 2ml collection tube, 500µl RPE buffer were added and the mixture was centrifuged again. The flow through could be discarded and 500µl EtOH (80%) were added to the column. After two minutes of centrifugation at 10000rpm the flow through was discarded and the column was placed in a new 2ml collection tube. To remove the remaining alcohol the column was centrifuged for five minutes at maximum speed. Again the column was transferred to a new 1.5ml collection tube and 14µl RNAse free water were added onto the middle of the columns membrane. To elute RNA from the column a centrifugation step of one minute at maximum speed was performed. The obtained concentration was checked and had to be above 1µg/µl to be able to perform the following steps.

6.4.2 Double strand cDNA synthesis

10µg of the purified RNA were diluted in RNAse- free water to a final volume of 9µl in PCR tubes and 2µl T7 oligo dT primer were added. The sample was incubated at 70°C for ten minutes followed by a 2 minutes incubation step at 4°C and then transferred on ice. To collect the evaporated liquids a short centrifugation step is performed. Now 4µl 5x first-strand reaction buffer, 2µl 0.1M DTT and 1µl 10mM dNTP mix were added and mixed thoroughly. After incubation the sample was put to 42°C for two minutes, 1.5µl SSII Super script reverse transcriptase were added and the sample was again incubated at 42°C for 60

minutes. The following steps had to be performed strictly on ice. 91µl DEPC-water, 30µl 5x second strand buffer, 3µl 10mM dNTP mix, 1µl E.coli DNA ligase, 4µl E.coli DNA polymerase I and 1µl E.coli RNAse H are added, mixed and the sample was incubated at 16°C for two hours. Now 2µl T4 DNA polymerase could be added and again incubated at 16°C for 5 minutes. The sample was put on ice and 10µl 0.5M EDTA (pH=8.0) were added.

6.4.3 Double strand cDNA clean-up

2600µl of cDNA binding buffer were added to the ds cDNA from step 6.4.2 and vortexed for 3 minutes. Using a color based pH indicator (yellow at pH 5) the adequate pH was adjusted by adding the right amount of 3M NaAc. The mixture was transferred in 500µl steps into a cDNA "Cleanup" spin column and consequently centrifuged at 10000rpm for one minute. The obtained flow through was discarded. Then the column was transferred to a fresh 2ml collection tube, 750µl cDNA washing buffer were added, the column was centrifuged for one minute at 10000rpm and the flow through was discarded. The whole amount of washing buffer was eliminated from the column by centrifuging at full speed for 5 minutes. Accordingly the column was transferred to a new collection tube. 14µl cDNA elution buffer were added onto the middle of the membrane and incubated for one minute. The cDNA was eluted from the column by a centrifugation step at full speed for one minute. Then the cDNA was stored at -80°C or was immediately used for the *in vitro* transcription process.

6.4.4 In vitro transcription

6µl of the purified ds cDNA were diluted in 16µl RNAse free water and 4µl 10x HY reaction buffer, 4µl 10x biotin labeled ribonucleotides, 4µl 10x DTT, 4µl RNAse inhibitor mix and 2µl T7 RNA polymerase were added to a final volume of 40µl. The ingredients were mixed thoroughly and were collected by spinning the tube. Then an incubation step of 16 hours at 37°C was performed. The solution had to be mixed every 30 to 40 minutes. The sample was stored at -80°C or the following clean-up step was performed immediately.

6.4.5 cRNA clean-up

The cRNA sample and 60µl DEPC water were vortexed together for three minutes, afterwards 350µl IVT cRNA binding buffer were added and vortexed again. After adding 250µl EtOH the solution was mixed. Then, the mixture was transferred to a cleanup spin column and centrifuged for 24 minutes at 10000rpm. The column was transferred to a new collection tube, 500µl IVT cRNA wash buffer were added and the column was again centrifuged for 44 minutes at 10000rpm. Consequently, the flow through is discarded. 500µl EtOH (80%) were added to the column and the same centrifugation step as described before was performed. To eliminate the whole amount of EtOH the column was centrifuged at full speed for five minutes and then the column was placed into a new collection tube. For the elution step 11µl RNAse free water were added onto the middle of the membrane and the column was centrifuged at full speed for one minute. Additional 10µl of RNAse free water were added onto the columns membrane. After the last centrifugation step at full speed for one minute the concentration of cRNA had to be determined and cRNA was stored at -80°C or immediately fragmented.

6.4.6 cRNA fragmentation

15µg of cRNA were fragmented in one approach. The cRNA was diluted in DEPC-water to a final volume of 16µl. 4µl fragmentation buffer were added and the sample was placed in the thermo-cycler at 94°C for 34 minutes.

During that time the microarray was prepared. (Important: the microarray had to be at room temperature 30 minutes prior to use, storage of the microarray at 4°C)

6.4.7 Microarray preparation and hybridization

PreHyb-buffer is prepared (volumes needed for 6 chips):

acetylated BSA	(50mg/ml)	15µl
HS-DNA	(10mg/ml)	75µl
2xMES-hyb.-buffer		750µl
DEPC-water		660µl

Then 230µl of the buffer were added to the microarray. After 15 minutes of incubation at 40°C on a rotating platform in an incubator the microarray was washed with 1x MES-hyb.-buffer, which had been previously made out of a 2x MES-hyb.buffer. 20µl of the fragmented cRNA were added to 230µl of hybridization buffer. The hybridization buffer contains:

HS-DNA	(10mg/ml)	15µl
acet. BSA		15µl
20 x eukaryotic controls	(Bios.)	15µl
Oligo B2		15µl
2 x MES-hyb.-buffer		750µl
DEPC-water		570µl

This approach was vortexed, collected by a short centrifugation step and incubated at 95°C for five minutes. Next it was centrifuged at full speed for five minutes. PreHyb-buffer was removed from the microarray and the cRNA-buffer mixture was added to the microarray. The hybridization was performed for 16 hours at 45°C on a rotating platform in an incubator.

6.4.8 Washing and scanning of microarrays

cRNA buffer mix was removed from the chip and transferred to a new 1.5ml tube. The microarray was washed immediately twice with 230µl 6xSSPE-T and the solution was left in the microarray. The microarray was put into the biopolis shared facilities (BSF) and it was

turned on for five minutes. The chip was washed twice with MES-wash buffer. The solution of the second washing step was left on the microarray and the chip was incubated for 30 minutes at 45°C (rotating). The streptavidin staining solution was prepared, containing:

Streptavidin, recombinant	(1mg/ml; in DEPC-water)	15µl
acet. BSA		75µl
2 x stain buffer		750µl
DEPC-water		660µl

After 30 minutes incubation the MES-wash buffer was discarded and 230µl of the streptavidin staining solution were added. After a 15 minutes incubation step at 40°C on the rotating platform the solution was discarded and the microarray was washed twice with 6x SSPE-T. Afterwards the microarray was incubated for five minutes in the BSF with 6x SSPE-T. During this time the biotinylated anti-streptavidin solution was prepared containing:

biotinylated anti-Streptavidin	6µl
acet. BSA	60µl
2 x stain buffer	750µl
DEPC-water	684µl

The microarray was washed once with MES-wash buffer and then 230µl biotinylated anti-streptavidin solution were added and incubated for 30 minutes at 40°C on a rotating platform. After that the buffer was discarded and the microarray was washed twice with SSPE-T. During five minutes incubation of the microarray in the BSF the streptavidin phycoerythrin (SAPE) solution was prepared, containing:

Streptavidin R-Phycoerythrin	(4°C, without light!)	15µl
acet. BSA	(50mg/ml)	75µl
2 x stain buffer		750µl

DEPC-water	660µl

The SSPE-T solution is discarded from the microarray and the microarray was washed once with MES-wash. Next 230µl SAPE-solution were added to the chip. After 15 minutes incubation at 40°C on a rotating platform the chip was washed twice with 6x SSPE-T and again incubated for five minutes in the BFS. Then the chip was filled completely with 6x SSPE-T (to make sure that no air bubbles remain in the microarray). Half an hour before the scanning step the scanner and the computer were turned on. Next the microarray was scanned and analyzed.

6.5 Methylated DNA immunoprecipitation (MeDIP)-chip analysis

MeDIP-chip analyses were performed in the three adenocarcinoma cell lines A549, NCI-H1993 and NCI-H2073 and in NHBE cells. Subsequent analyses were performed in 100 primary lung tumor samples and in 100 corresponding non-malignant lung tissue samples.

6.5.1 Fragmentation of genomic DNA

Genomic DNA was fragmented using the ultrasonic device “Disrupter” (Diagenode, Liége, Belgium) to obtain DNA fragments of 200-800bp in length. During the 10 minutes process DNA was cooled with ice-water to prevent DNA denaturation. To confirm fragmentation an agarose gel was loaded with 300ng of sonicated DNA in 2 µl loading buffer.

6.5.2 Methylated DNA immunoprecipitation (MeDIP)

Materials:	TE Buffer (for ingredients see 6.6)	
	5x IP buffer	
	Digestion buffer	
	Glycogen	(Roche Applied Science; #901393)
	anti-5-methylcytosine antibody	(Eurogentec; #080917)
	NaCl, 5M	(Ambion; #AM9759)
	Protein A agarose beads /	(Millipore; #16-157)
	Salmon sperm DNA	

3.2µg fragmented DNA was diluted in TE buffer to a final volume of 420µl. The DNA was heat-denaturated by incubation at 95°C for 10 minutes. Immediately the samples were put on ice for five minutes and evaporation was collected from the cap of the tube by a short centrifugation step in a pre-cooled centrifuge at 4°C. Next 20µl were removed from the sample for the input sample (referred to as input) and stored at -20°C. To the remaining 400µl 100µl 5x IP buffer and 3.5µl anti-5-methylcytosine antibody were added. The sample was incubated on a rotating platform for two hours at 4°C. During this incubation period antibodies bound to methylated CpGs.

80µl Protein A agarose beads (50% slurry) were pre-washed before incubation with the DNA-sample/antibody mixture. First the beads were centrifuged at 2000rpm for one minute and the supernatant was removed. Next 1ml 1x IP buffer was added to the beads mixed by resuspending and again centrifuged at 2000rpm for one minute. Again the supernatant was removed and the beads were resuspended in 40 µl 1x IP buffer. Pre-washed beads were added to the DNA/antibody sample followed by one hour incubation on a rotating platform at 4°C. During that incubation time the beads bound to the antibodies in the sample and a complex out of methylated DNA/antibody/beads was formed. Then the DNA/antibody/beads mixture was washed five times. Each washing step consisted of centrifugation at 2000rpm for one minute at 4°C, discarding the supernatant, adding 1ml of 1x IP buffer and incubation for five minutes on a rotating platform at 4°C. Unbound DNA was removed away and only methylated DNA bounded to the antibody/beads complex remained. Next DNA had to be separated from the DNA/antibody/beads complex. A digestion buffer was prepared. In the last washing step the sample was centrifuged and the supernatant was discarded. 250µl digestion buffer and 7µl proteinase K mix (10mg/ml) were added after the last washing step. For separation of the proteins from DNA one hour incubation at 55°C was performed. Then DNA was extracted from the mixture using standard phenol/chloroform extraction and EtOH precipitation. 250 µl phenol were added and mixed by vortexing. The DNA containing aqueous phase was separated from the organic phase by centrifuging for 10 minutes at 14000rpm and 20°C. The upper aqueous phase was transferred into a new 1.5ml tube and 250µl chloroform/isoamyl alcohol were added. The sample was again mixed by vortexing and centrifuged at 14000rpm for 10 minutes. The

supernatant was transferred into a new 1.5ml tube and 1µl glycogen (20mg/ml), 20µl 5M NaCl and 500µl EtOH were added. After mixing the sample by hand it was put at -80°C for 20 minutes. Next the sample was centrifuged in a pre-cooled centrifuge for 15 minutes at 4°C. A small pellet became visible. The supernatant was decanted and 500µl EtOH (70%) were added and the pellet was washed by vortexing the tube. Afterwards the sample was centrifuged for ten minutes at 14000rpm and 4°C. The supernatant was again decanted and a short centrifugation step ensured that the whole amount of alcohol was removed with a pipette. The pellet was dried on room temperature for 5 minutes and then dissolved in 20µl TE buffer (referred to as IP sample).

The DNA concentration of the IP and the input samples was determined with a Nanodrop 2000 (Thermo Scientific) yielding DNA amounts of 10-20ng/µl.

6.5.3 Whole genome amplification (WGA)

In general, ~5% of input DNA were recovered by MeDIP. However, for subsequent microarray analyses at least 4µg of immunoprecipitated DNA were required. Thus, DNA samples were amplified using WGA. A GenomePlex® Complete WGA Kit (WGA2) (Sigma-Aldrich) was used and a double determination was performed to reach even higher DNA amounts. Each step of this approach was performed according to the manufacturer's instructions.

6.5.3.1 Step 1 of WGA

A linker DNA-fragment was added to the sonicated DNA to allow amplification of different fragments with just one primer. We started with 20ng DNA per approach and ddH_2O was added to obtain a final volume of 11µl. Master mix was added and the DNA was heat-denatured in the PCR cycler at 95 °C for 3 minutes. After denaturation it had to be worked strictly on ice. The ligase-enzyme was added and then a PCR program was started as followed:

16 °C for 20 minutes

24°C for 20 minutes

37°C for 20 minutes

75°C for 5 minutes

Maintain at 4°C

6.5.3.2 Step 2 of WGA

Now amplification of the linker/DNA-fragments combination was possible. The master mix and the enzyme were added to the sample from step 1. A PCR amplification program was started in the GeneAmp PCR System 9700 as following:

Initial Denaturation	95°C for 3 minutes
Fourteen cycles each consisting of:	
Denaturation	94°C for 15 seconds
Annealing/Extention	65°C for 5 minutes

After finishing the program the sample was maintained at 4°C or store at -20°C until the following clean-up step was started. Each sample was amplified as duplicates.

6.5.4 Clean-up of amplified DNA

The clean-up step was performed on a specific post PCR-working place. DNA was purified using a GenElute™ PCR clean-Up Kit (Sigma-Aldrich). The two approaches of each sample were pooled in one 1.5ml tube before starting the clean-up. Each step was performed according the manufactory's manuals. The column had to be pre-washed. For that purpose 500µl column preparation buffer were put into the column and centrifuged at 12000rpm for one minute. The PCR sample was combined with 5-fold volume of binding buffer and transferred into the column. Since DNA bound to the column DNA was separated from the buffer solution by centrifugation at 12000rpm for one minute. Bounded DNA was washed with 500µl washing buffer. A centrifugation step of two minutes at 12000rpm made sure that the whole amount of liquid was removed from the column. By adding 40µl elution buffer onto the middle of the membrane and after centrifugation the column for 1 minute at 12000rpm DNA was eluted from the column. The amount of DNA was determined by using a Nanodrop 2000 (Thermo Scientific). DNA was stored in the elution buffer on -20°C.

6.5.5 Test for enrichment efficiency by quantitative real-time PCR

We used the quantitative real-time PCR to ensure the enrichment of the methylated DNA in the IP sample compared to the input sample. The amount of a gene that is known to be methylated in somatic cells like H19, an imprinted gene, was measured in the IP and the input sample and the results were compared. In addition, the amount of another gene, namely GAPDH, a housekeeping gene, which is unmethylated in somatic cells was analyzed and the results obtained from the two samples were compared. H19 was highly enriched in the IP sample compared to the input sample and GAPDH was in higher amounts present in the Input sample. Enrichment of methylated sequences was calculated and normalized to GAPDH. An example for calculating the enrichment fold change is shown in Table 2.

Table 2
Calculation of the fold change in a qRT-PCR approach for A549.

	GAPDH	***mean value***	***H19***	***mean value***	***Δ/ΔCt***	***DDCT (G)***	***Expr 2 exponential G***
A549 input	28,77	28,645	25,8	25,915	-2,73		**1**
A549 input	28,52		26,03				
A549 IP	38,29	38,19	23,47	23,70	-14,49	11,76	**3468,27**
A549 IP	38,08		23,92				

100ng of DNA were used for qRT-PCR analyses. H19 primers were designed in our laboratory. A TaqMan® Gene Expression Assays for GAPDH was designed by Applied Biosystems and were used as recommended from the manufacture. Reactions were performed in duplicates.

GAPDH:	TaqMan Master Mix	12,0µl
	dd H2O	11,4µl
	Primer	0,6µl
H19:	SYBR Green Master Mix	12,5µl
	dd H2O	10,5µl
	Primer	1,0µl

2µl of master mix were added to each reaction. 1µl (100ng/µl) of the DNA were added. Afterwards a stripe was put on top of the plate to close it thoroughly. A quantitative real-time PCR was performed using the ABI 7000 cycler.

6.5.6 MeDIP-chip analysis

After confirming the enrichment of methylated DNA, the samples were sent to NimbleGen service laboratories for a second quality control step, sample labeling and hybridization to NimbleGen´s HG18 CpG plus Promoter arrays comprising 385.019 probes. These probes cover all reported Human Refseq gene promoters (24.659) and all reported CpG islands (28.226) annotated on the UCSC genome browser. IP DNA was labeled with Cyanine 5 (Cy5) and input DNA was labeled with Cyanine 3 (Cy3). Both Cy3 and Cy5 belong to the cyanine dye super family and both are reactive fluorescent dyes soluble in water. Cy3 emits light with a wavelength of ~570nm, appearing green, Cy5 in contrast to that emits light with a wavelength of ~650/670nm which appear red.

IP and input samples were hybridized to the microarray chip. The promoter regions covered on the microarrays are 1kb and small CpG islands were extended at both ends for a total coverage of 700bp for more reliable detection. In addition, the arrays also contained DNA methylation positive control regions (H19/IGF2 cluster, KCNQ1 cluster, and IGF2R gene). Arrays were scanned on a GenePix 4200A Scanner using GenePix 6.0 Software and data were extracted from scanned images by using NimbleScan 2.3 (see Figure 13).

For each cell line analyzed, we obtained scaled log ratio data for individual samples (control versus sample), log ratio GFF files and genome annotation GFF files.

Figure 13
Schematic description of MeDIP-chip technology using NimbleGen microarray technique. First DNA is fragmented then methylated fragments are enriched in an IP sample. By comparing the IP sample to an input sample the methylated genes can be identified. Modified from NimbleGen®

6.5.7 Statistical analysis of microarray data

Data derived from MeDIP-chip were analyzed in cooperation with a team of biostatisticians from the Medical University of Vienna, Section of Medical Statistics, Vienna,

Austria under the supervision of Prof. Dr. Martin Posch. In brief, data were analyzed using the bioconductor package Ringo. Data were normalized with NimbleGen-normalization (tukey's biweight mean across each sample's log2 ratios were computed and subtracted from the individual log2 ratios). For smoothing the data, a window size of 750bp was slid along the chromosome; the intensity at the genomic position x0 was replaced by the median over all intensities of reporters inside the window that was centered at x0. Quality assessment was done by histograms of the normalized and smoothed data and by correlation plots (raw data, normalized data and normalized + smoothed data).

To test for potential differences in methylation between cases and controls two sample t-tests were calculated and p-values were corrected for multiple testing by the Benjamini-Hochberg correction. Detailed results about the pattern of methylation between tumors and non-malignant lung samples and the frequencies of methylation of certain genes can be obtained with this method.

Data derived from the gene expression microarray approach were analyzed as reported by Heller et al. (153) using Affymetrix Microarray Analysis Suite version 5 (MAS5) resulting in processed scanned chip images and a cell intensity file for each chip. Consequently, statistical analysis was performed using Bioconductor's affy package. Using CyberT algorithm the significance of changes in the expression level of each gene was calculated. These tests were performed using the R-based Flexarray 1.3 software. To set a gene as methylated changes in the expression level had to meet following criteria: gene expression levels obtained from drug treated cells had to be >2-fold higher than those obtained from untreated cells. Difference in gene expression between control and drug treated cells had to be statistically significant ($P<0.01$).

Using the Parent-Child-Intersection algorithm of Ontologizer the obtained results were analyzed with Gene Ontology (GO) analysis (171). P-values of these analyses were calculated using one-sided Fisher exact test corrected for multiple testing using the Bonferroni method as provided in Ontologizer.

6.6 Buffer ingredients

1xPBS:	NaCl	0.137M	(Sigma-Aldrich; #. S7653)
	KCl	2.7mM	(Sigma-Aldrich; #P5405)
	Na_2HPO_4	4.3mM	(Sigma-Aldrich; # S9390)
	KH_2PO_4	1.4mM	(Sigma-Aldrich; # P5655)
Electrophorese buffer:	1 x TAE		(BIO-RAD; #1610743)
Loading buffer:	Bromphenol blue	0,05%	(Sigma-Aldrich; #B-5525)
	Saccharose	40%	
	EDTA (pH 8,0)	0,1M	(Sigma-A.#43178-8)
	SDS	0,5%	(Sigma-Aldrich;#L-4522)
PK-buffer:		***10ml***	
	1,0M Tris (pH 8,0)	0.5ml	(Biomol; #50005)
	0,5M EDTA (pH 8,0)	20µl	(Sigma-Aldrich; #43178-8)
	SDS 20%	50µl	(Sigma-Aldrich; #L-4522)
	Proteinase K10mg/ml	0.2ml	(Roche Applied Science;#745723)
	H_2O	9.23ml	
TE Buffer	Tris HCL pH 8.0	10mM	(Biomol; #50005)
	EDTA	1mM	(Sigma-Aldrich; #43178-8)
5x IP buffer		***10ml***	
	100mM Na-Phosphat pH7	5ml	(Fluka; #82637)
	5M NaCl	1.4ml	(Ambion; #AM9759)
	10% Triton X-100	250µl	(Sigma-Aldrich; #23472-9)
	ddH_2O	3.35ml	
Digestion buffer		***10ml***	
	1M Tris HCl, pH 8.0	500µl	(Sigma-Aldrich; #T3038-1L)
	0.5M EDTA	200µl	(Sigma-Aldrich; #E7889)
	10%SDS	500µl	(Sigma-Aldrich; #L-4522)
	ddH_2O	8.8ml	

7 RESULTS

The aim of our study was to better understand the role of DNA methylation in the pathology of NSCLC. To reach this goal we used different strategies: Firstly, we performed gene expression microarray analysis of three lung adenocarcinoma cell lines (A549, NCI-H1993, NCI-H2073) before and after treatment with the epigenetically active drugs Aza-dC and/or TSA. Genes that showed no or low expression before and elevated expression after drug treatment could be identified as potentially epigenetically regulated. Secondly, we established a methylated DNA immunoprecipitation combined microarray (MeDIP-chip) approach and analyzed the genome-wide DNA methylation pattern of 3 adenocarcinoma cell lines (A549, NCI-H1993, NCI-H2073) and NHBE cells. Based on the microarray results we selected several genes for further single gene methylation analysis of primary tumor and corresponding non-malignant lung tissue samples of 10 NSCLC patients using methylation-specific PCR (MSP).

7.1 Genome-wide DNA methylation analysis of NSCLC cell lines

7.1.1 Gene expression analysis of the NSCLC cell lines A549, NCI-H1993 and NCI-H2073 before and after treatment with Aza-dC and/or TSA

Three NSCLC cell lines (A549, NCI-H1993 and NCI-H2073) were treated with the epigenetically active drugs Aza-dC, TSA or a combination of Aza-dC and TSA. Afterwards, expression microarray analyses were performed using Affymetrix HG-U133_plus_2.0 GeneChips as reported recently (153). To identify potentially methylated genes microarray data from untreated and drug treated cells were compared. Genes found to be up-regulated after drug treatment had to meet the following criteria: the fold change in expression after drug treatment had to be >2-fold and differences in expression between untreated and drug treated cells had to be statistically significant ($p<0.01$). The universally methylated transketolase-like1 (*TKTL1*) gene was used as a positive control for loss of DNA methylation

and induction of gene expression as reported recently (154). *TKTL1* was found to be up-regulated by Aza-dC in all cell lines analyzed. Each experiment was performed in duplicate with independently grown cells to verify the comparability of microarray data as shown in Figure 14.

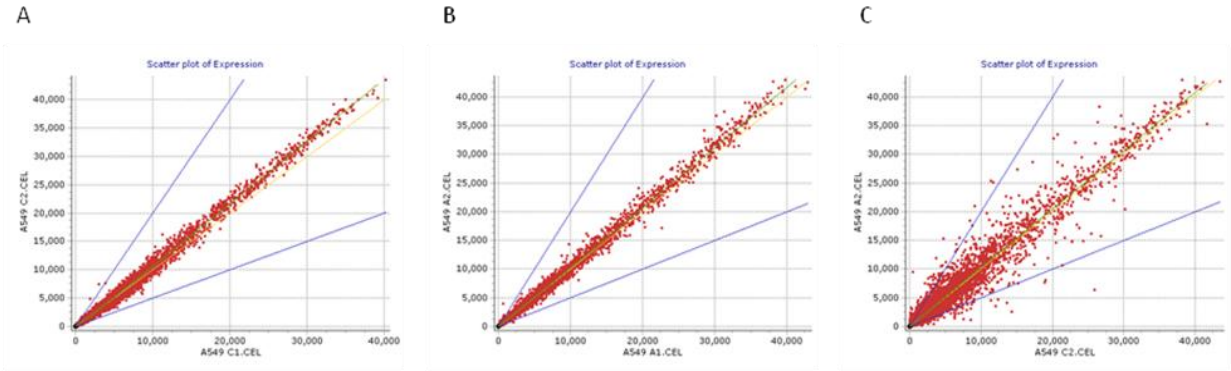

Figure 14
Scatter plots showing the comparability of expression microarray analyses. Comparison of microarray data of untreated cells (C1 vs. C2) are shown in (A) and of Aza-dC treated cells (A1 vs. A2) are shown in (B). Comparison of untreated and Aza-dC treated cells is shown in (C). Light blue lines indicate the achieved fold change in gene expression.

7.1.1.1 Aza-dC induced genes

To determine the effect of DNMT inhibition on gene expression of NSCLC cell lines, we analyzed gene expression profiles of 3 NSCLC cell lines before and after treatment with Aza-dC. As shown in Figure 15 up-regulation of gene expression as well as down-regulation of gene expression was observed. Since, we were interested in the identification of epigenetically inactivated genes we focused only on up-regulated genes.

1059 probe sets (1.9% of 54.675 transcripts analyzed) were found to be up-regulated in the cell line A549, 286 probe sets (0.5%) in NCI-H1993 cells and 1296 probe sets (2.3%) in NCI-H2073 cells, respectively. Fold changes in expression varied from 2-fold to 201.6-fold, from 2-fold to 28.6-fold and from 2-fold to 83.15-fold in the cell lines A549, NCI-H1993 and NCI-H2073, respectively. As multiple probe sets for the same gene are present on Affymetrix microarrays, Aza-dC induced probe sets represent 768 unique genes in the case of A549, 219 unique genes in the case of NCI-H1993 and 986 unique genes in the case of NCI-H2073. By comparing lists of up-regulated genes in the three cell lines we identified two genes commonly up-regulated in all three cell lines analyzed, 21 genes were up-regulated in both A549 and NCI-H1993, 6 genes were up-regulated in the cell lines NCI-H1993 and NCI-H2073

and 24 genes were commonly up-regulated in the cell lines A549 and NCI-H2073 (Figure 16). Overall, 1153 unique genes were found to be up-regulated in at least one of the three lung adenocarcinoma cell lines after exposure to Aza-dC.

7.1.1.2 TSA induced genes

Next, gene expression patterns of the three NSCLC cell lines were analyzed before and after treatment with TSA. Volcano plots of these results are shown in Figure 15. By gene expression microarray analyses of the three cell lines after TSA treatment we identified 1428 up-regulated probe sets (2.6% of 54.675 transcripts analyzed) in the cell line A549, 2070 up-regulated probe sets (3.8%) in the cell line NCI-H1993 and 453 up-regulated probe sets (0.8%) in the cell line NCI-H2073, respectively. These probe sets represent 847, 1337 and 287 unique genes in the cell lines A549, NCI-H1993 and NCI-H2073, respectively. Fold changes in expression varied from 2-fold to 63.37-fold, from 2-fold to 142.84-fold and from 2-fold to 47.33-fold in the cell lines A549, NCI-H1993 and NCI-H2073, respectively. By comparing gene expression data after TSA treatment of the three cell lines we identified 19 genes commonly up-regulated in all three cell lines analyzed, 92 genes were up-regulated in both A549 and NCI-H1993, 72 genes were commonly up-regulated in the cell lines NCI-H1993 and NCI-H2073 and 21 genes were up-regulated in the cell lines A549 and NCI-H2073 (Figure 16). Overall, 2254 unique genes were up-regulated in at least one of the three lung adenocarcinoma cell lines after exposure to TSA. By comparing microarray data from cells treated with Aza-dC and TSA we found that about 27% of identified genes after Aza-dC induction were additionally induced by TSA treatment (Figure 17).

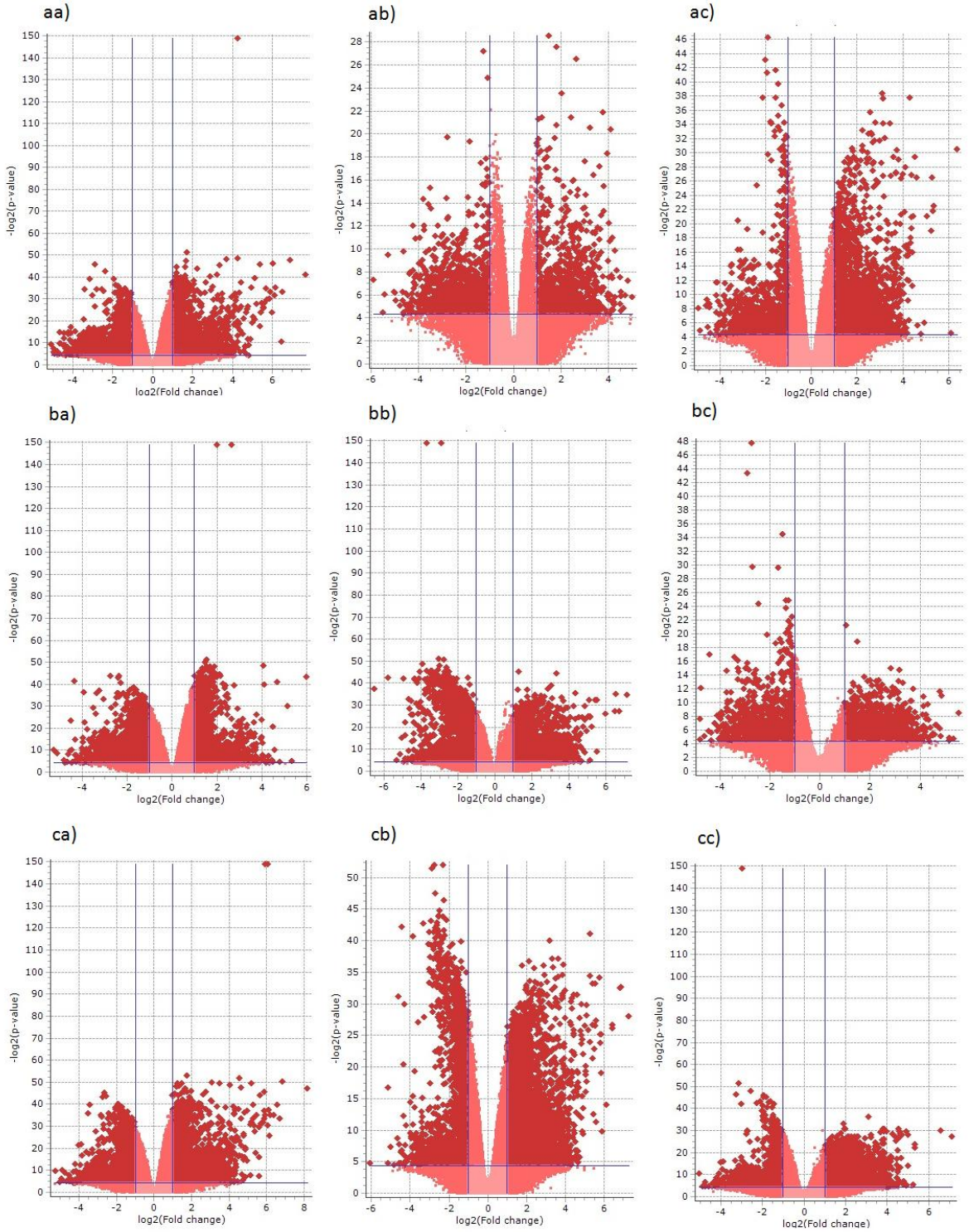

Figure 15
Volcano plots showing gene expression changes after treatment of NSCLC cell lines with Aza-dC, TSA or Aza-dC/TSA. Fold change and the statistical significance are plotted on the x- and y- axis, respectively. Changes in the expression level of A549 are shown in a) after aa) Aza-dC, ab) TSA, and ac) Aza-dC and TSA treatment. Changes in the expression level in NCI-H1993 in b) after ba) Aza-dC, bb) TSA, and bc) Aza-dC and TSA treatment and changes in the expression level in NCI-H2073 are shown in c) after ca) Aza-dC, cb) TSA, and cc) Aza-dC and TSA treatment. Genes which were statistically significant (p>0.01) up-regulated >2-fold are shown in dark red.

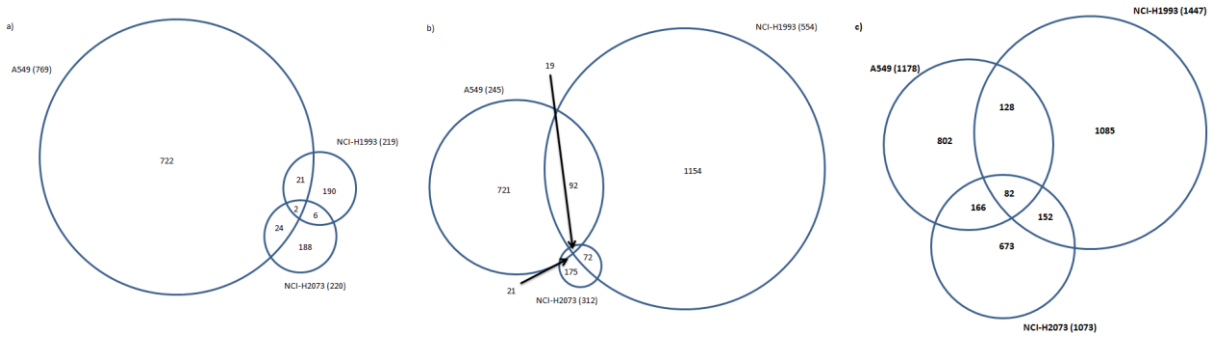

Figure 16
Venn diagrams summarizing the number of up-regulated genes in the cell lines A549, H1993 and H2073 after a) Aza-dC, b) TSA, and c) Aza-dC/TSA treatment. The region of overlap between all circles shows the number of genes up-regulated in all three cell lines analyzed. Regions of overlap between the circles of two cell lines indicate genes up-regulated in two of the three cell lines. Regions that do not overlap between circles indicate genes up-regulated in only one cell line.

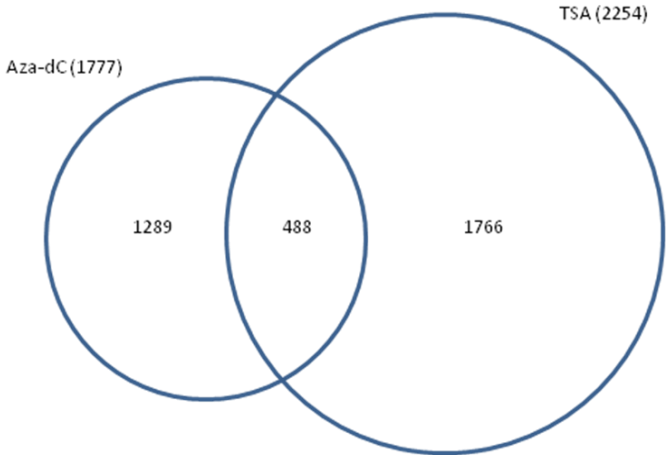

Figure 17
Venn diagram showing genes which were up-regulated after Aza-dC treatment and genes that were up-regulated after TSA treatment. The overlap between the circles shows genes that were up-regulated by both Aza-dC treatment and TSA treatment.

7.1.1.3 TSA and Aza-dC treated cells

Consequently, the lung adenocarcinoma cell lines A549, NCI-H1993 and NCI-H2073 were treated with a combination of the two epigenetically active drugs Aza-dC and TSA. Volcano plots of the results are shown in Figure 15. Treatment of cells with the combination of Aza-dC and TSA resulted in up-regulation of 1769 probe sets (3.2%) in the cell line A549, of 2366 probe sets (4.3%) in NCI-H1993 and of 1462 probe sets (2.6%) in NCI-H2073, respectively. Fold changes in expression varied from 2-fold to 289.96-fold, from 2-fold to 153.61-fold and from 2-fold to 139.11-fold in the cell lines A549, NCI-H1993 and NCI-H2073, respectively. In the cell lines A549, NCI-H1993 and NCI-H2073 these findings resulted in 1186, 1452 and 1073 unique genes up-regulated after Aza-dC and TSA, respectively. 82

genes were commonly up-regulated in all three cell lines analyzed. In addition, 128 genes were up-regulated in both A549 and NCI-H1993, 152 genes were commonly up-regulated in the cell lines NCI-H1993 and NCI-H2073 and in the cell lines A549 and NCI-H2073 166 genes were commonly up-regulated (Figure 16). 3088 unique genes were found to be induced in at least one of the three lung adenocarcinoma cell lines after exposure to the combination of Aza-dC and TSA.

An interesting aspect is that 91% (1055 unique genes) of all genes up-regulated after Aza-dC (1153 unique genes) were additionally up-regulated after treatment with the Aza-dC/TSA combination (see Figure 18).

It has been reported that Aza-dC and TSA act synergistically in inducing expression of genes which are regulated by DNA methylation (172). By comparing microarray results from Aza-dC and Aza-dC/TSA treated cells we observed a synergistic up-regulation of gene expression in 37% of up regulated genes in A549, in 23% of up regulated genes in NCI-H1993 and in 33% of up-regulated genes in the cell line NCI-H2073, respectively. Heat maps of these genes are shown in Figure 19. Examples for synergistically up-regulated genes, their chromosomal location, their physiological function and their relation to cancer are shown in Table 3. The genomic sequence of the 5' regions of these genes was obtained from ESEMBLE database (release 58). Using the software "Methyl Primer Express v 1.0" 568 unique sequences were analyzed concerning their CpG-content. In total, 77% of the synergistically up-regulated genes were associated to a CpG-island within the 5' region. These 438 genes were used for Gene Ontology (GO) analyses.

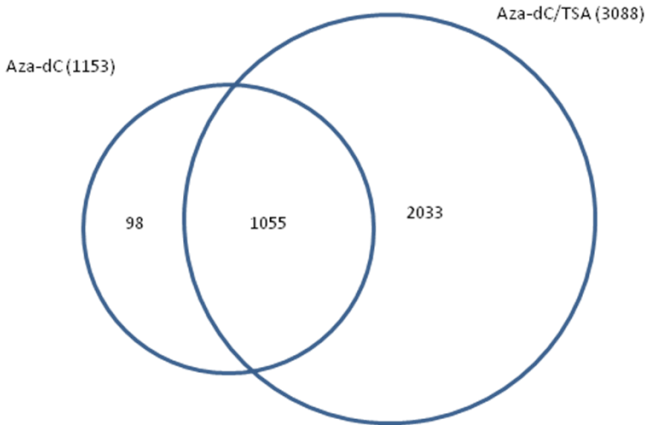

Figure 18
Venn diagram showing genes which were up-regulated after Aza-dC treatment and genes that were up-regulated after Aza-dC/TSA treatment. The overlap between the circles shows genes that were up-regulated by both Aza-dC treatment and the combinatorial treatment with Aza-dC/TSA.

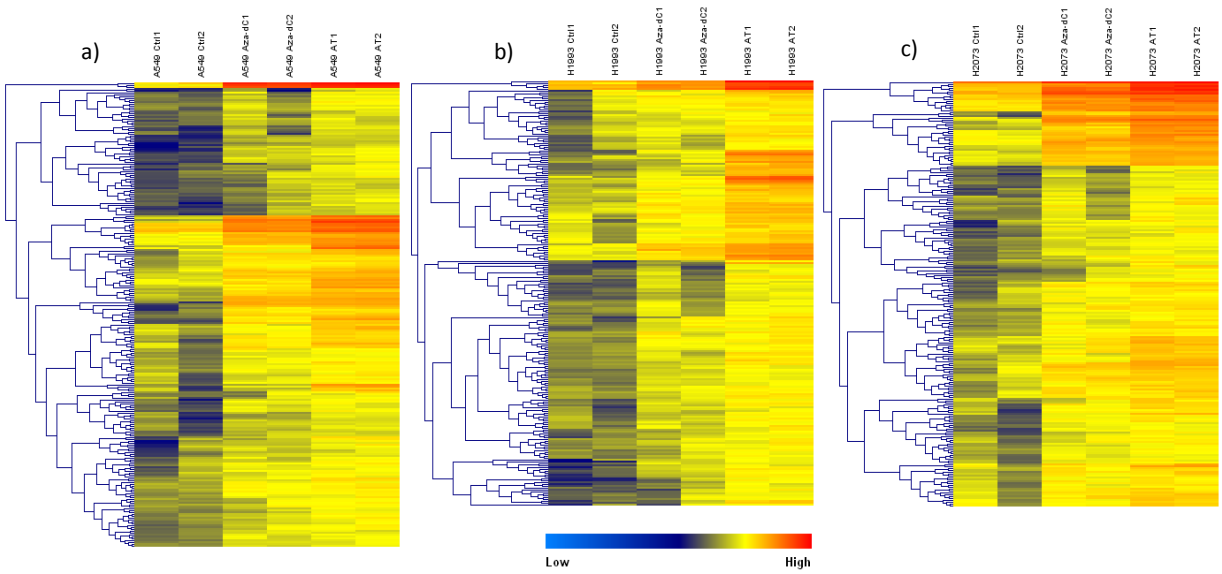

Figure 19
Heat maps of genes which are synergistically up-regulated by Aza-dC and TSA in the cell lines (a A549, (b NCI-H1993 and (c NCI-H2073.

Table 3
Examples of genes which were found to be synergistically up-regulated after drug treatment of A549, NCI-H1993 and NCI-H2073, their chromosomal location, their physiological function and knowledge about cancer association are described.

Gene	**Cytoband**	**Physiological function**	**FC after treatment with Aza-dC**	**FC after treatment with Aza-dC/TSA**	**Down-regulation reported in**	**CpG Island**
ALDH1A3	15q26.3	Cell detoxification	2,24	2,46	n/a	1
DLX4	17q21.33	Transcription factor activity	3,21	4,52	Lung cancer	1
TAL1	1p32	Transcription factor activity	5,55	7,11	n/a	1
TMTC1	12p11.22	Membrane integration	7,43	14,30	n/a	1
HEXIM1	17q21.31	Cyclin-dependent protein kinase inhibitor activity	6,72	8,12	Breast cancer	1
PTEN	10q23.3	EGFR downstream regulator	5,51	7,93	Colon, Prostate, Endometrial cancer	1
RECK	9p13.3	Endopeptidase inhibitor activity	4,44	6,74	Gastric, lung cancer	1
NEFH	22q12.2	Cell death	27,36	106,64	Esophageal cancer	1
NOX5	15q23	Cell proliferation regulation	2,25	3,40	n/a	1
KSR1	17q11.1-q11.2	Member of the MAPK signaling pathway, cell proliferation	2,03	8,12	Pancreas cancer	0
CPT1C	19q13.33	Lipid metabolic process	3,51	7,41	Skin cancer	1
HIST1H2AE	6p22.2	Histone protein, DNA compactation	10,69	16,18	n/a	1

HOXA9	7p15.2	Transcription factor, development regulation	2,39	6,35	Breast cancer	1
IRX2	5p15.33	Regulation of transcription	7,53	8,74	Breast cancer	1

FC, fold change; n/a, not available; information provided in this Table is depicted from references (131-136); (141); (143-144); (172)

7.1.1.4 Identification of cancer-associated genes

Next, we performed Gene Ontology (GO) analyses to identify genes which are involved in cancer-related molecular pathways and thus may have tumor suppressor function within the synergistically up-regulated genes associated to a CpG island in their promoter region. The population set was defined as all genes which are represented on Affymetrix HG-U133_plus_2.0 GeneChips. 438 genes were found to be synergistically up-regulated and were affiliated to a CpG island. These genes were analyzed using Ontologizer (171). Our analysis revealed up-regulation of already known targets for aberrant methylation such as *PTEN* (173). Even more interestingly we identified hitherto unknown target genes for methylation taking part in cancer-related pathways such as *RECK* (negative regulation of cell cycle), *NOX5 (*induction of apoptosis), *CD274* (negative regulation of cell proliferation) or *ATR* (negative regulation of DNA replication). A summary of results from GO analysis is shown in Figure 20.

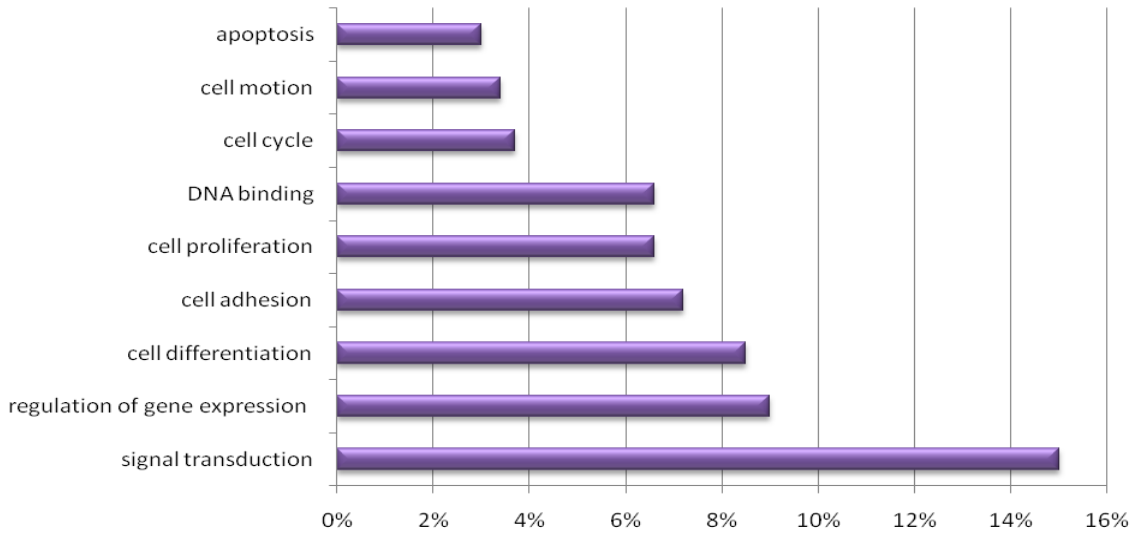

Figure 20
GO analyses showing the physiological function of synergistically up-regulated genes after Aza-dC and TSA treatment in three NSCLC cell lines A549, NCI-H1993 and NCI-H2073.

7.1.1.5 Single gene DNA methylation analysis of selected genes in tissue samples of NSCLC patients (n=10)

Selected genes found to be synergistically up-regulated by Aza-dC and TSA were further analyzed by methylation-specific PCR in primary tumors and corresponding non-malignant lung tissue samples from 10 NSCLC patients. The genes were randomly selected and included *TAL1*, *DLX4, TMTC1*, *NEFH, ALDH1A3, THSD4* and *MACF1*. Known targets for methylation in human tumors like *ALDH1A3,* as well as genes so far not further analyzed in cancer concerning DNA methylation like *NEFH* (microtubule cytoskeleton organization)*, TAL1* (cell proliferation)*, DLX4* (putative metastasis suppressor), *THSD4* (proteolysis) and even genes like *TMTC1* whose function is completely unknown, were part of the selection.

As a first step MSP assays were designed for the 5´ regions of the seven genes using Methyl Primer Express® Software v1.0. Primer sequences and amplicon length are shown in 6.3.3. As shown in Figure 21, methylation of the genes *THSD4, TMTC1*, *NEFH, DLX4, ALDH1A3*, and *TAL1* were found more frequently in primary tumor samples than in corresponding non-malignant lung tissue samples. In contrast, the gene *MACF1* was equally methylated in primary lung tumor samples and corresponding non-malignant lung tissue samples. Methylation was observed in primary NSCLC samples in different frequencies. The most frequently methylated genes were *THSD4* and *TMTC1* which were methylated in 90% and 80% of the analyzed tumor samples, respectively. Both genes were methylated in 0% of the analyzed non-malignant lung tissue samples. *NEFH, DLX4, ALDH1A3 and TAL1* were methylated in 70%, 70%, 50% and 30% of the NSCLC samples and in 0%, 10%, 0% and 10% of the matching non-malignant lung tissue samples, respectively. Analyzing the methylation status of *MACF1* revealed that methylation in the promoter region of this gene occurs not only in primary lung tumor samples analyzed but also in the corresponding non-malignant lung tissue samples.

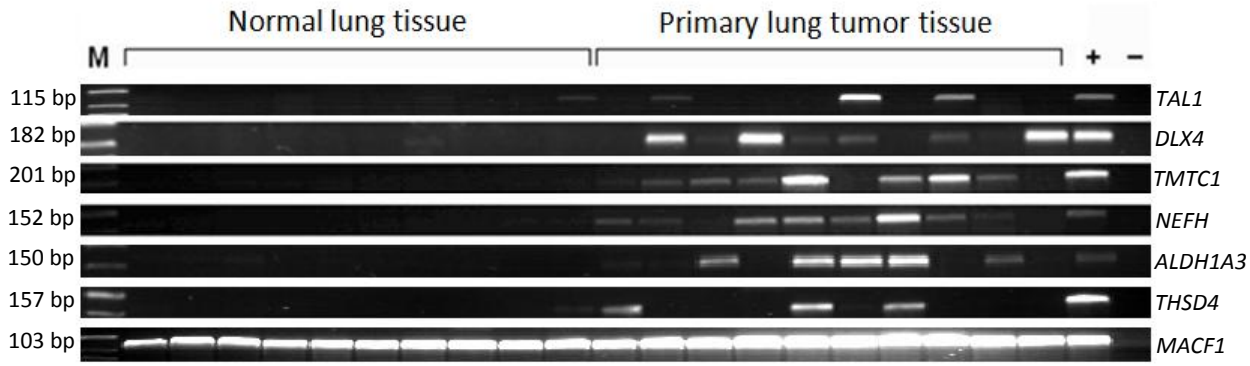

Figure 21
MSP analysis of methylation frequency of *TAL1*, *DLX4*, *TMTC1*, *NEFH*, *ALDH1A3*, *THSD4* and *MACF1* in primary lung tumor samples and corresponding non-malignant lung tissue samples of 10 NSCLC patients. Apart from *MACF1* all genes were found to be tumor-specifically methylated; + positive control; - negative control; M 100bp ladder.

7.1.2 Establishing MeDIP-chip

Immunoprecipitation of methylated DNA sequences in combination with subsequent microarray analysis (MeDIP-chip) is a recently developed technique for determination of genome-wide DNA methylation. Using an anti-5-methylcytosine antibody, methylated DNA fragments are enriched and subsequently hybridized to microarrays. A detailed description of the MeDIP approach is provided in section 6.5.2.

To optimize immunoprecipitation different setups were analyzed and evaluated. The approach was optimized for the following variables: duration of the sonication step, duration of the incubation step with the antibody as well as with the Protein A agarose beads and the number of washing steps. For testifying the enrichment of methylated DNA fragments and the exclusion of unmethylated fragments qRT-PCR for the imprinted gene *H19* (methylated control) and the housekeeping gene *GAPDH* (unmethylated control) was performed as described in 6.5.5. Since, the obtained amount of DNA after MeDIP is very low (10-20ng/µl) amplification of the IP samples was performed using a whole genome amplification approach (see section 6.5.3 for exact WGA process). According to suggestions from NimbleGen concerning fold enrichment of methylated DNA fragments we chose a fold change >500 as cut-off for efficient MeDIP assay.

Fragmentation of genomic DNA is the fundament for separating methylated DNA from unmethylated DNA. Thus, we aimed to obtain DNA fragments of 200-800 bp in length.

Minimal variation in the time of sonication revealed changes in the fragmentation state of the genomic DNA. Sonication for 5 minutes resulted in fragments of up to 2000bp in length. After sonicating genomic DNA for 10 minutes mainly fragments were present with a length between 150bp and 800bp in length. Sonicating genomic DNA for 15 minutes resulted in fragments which were 100-600bp. The difference between sonication for 5 minutes and sonication for 10 minutes was drastically. We chose 10 minutes sonication time with intervals of 30 seconds sonication followed by 30 seconds break as optimal setup because DNA fragments are long enough for the needed whole genome amplification step but not too long to ensure separating the methylated fragments from the unmethylated ones (see Figure 22).

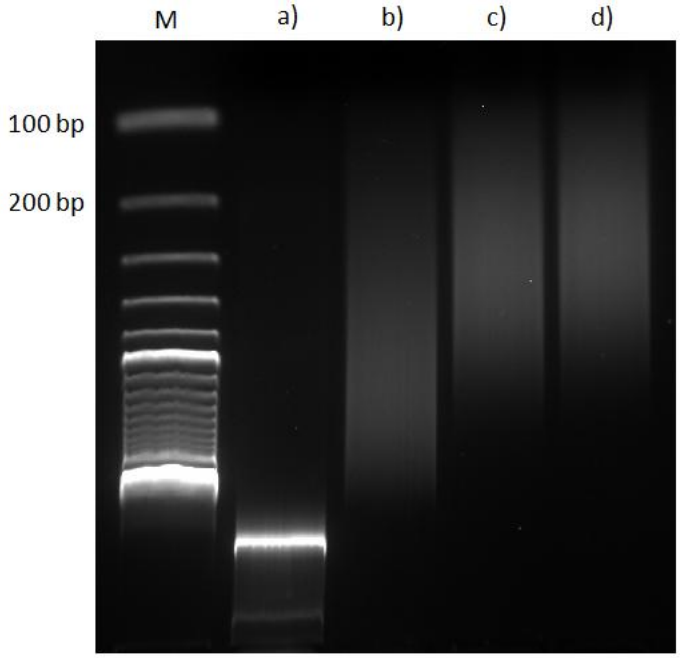

Figure 22
Agarose gel-electrophones image showing a) genomic DNA which was not sonicated, b) genomic DNA sonicated for 5 minutes, c) genomic DNA sonicated for 10 minutes and d) genomic DNA sonicated for 15 minutes. M, 100bp ladder.

For enrichment of methylated DNA fragments 3.2µg sonicated DNA were incubated with 3.5µg of an antibody specific for 5-methylcytosine for two hours at 4°C. Incubation time with the antibody as well as the consequent incubation with protein A agarose beads is sufficient for specific antibody binding to methylated fragments. However, if applied in the wrong way it can also lead to unspecific binding. At first, the incubation period with the antibody was optimized. One, two and three hours of incubation were tested. Incubation of genomic DNA fragments for one hour led to enrichment of low DNA amounts (<5ng/µl). However, for

subsequent experiments at least 10ng/µl DNA are required. Three hours of incubation yielded a high amount of DNA and unspecific antibody binding as determined by qRT-PCR. As shown in Figure 23 target sequences for primers specific for *H19*, which represents the methylated fragments, were present in equal amounts in the input sample and in the IP sample. Unmethylated fragments, which are depictured by amplification of *GAPDH*, were slightly enriched in the input sample compared to the IP sample. This amount of enrichment is not ensuring elimination of unmethylated fragments. The calculated fold change for enrichment of methylated DNA fragments in this approach was 34.78-fold.

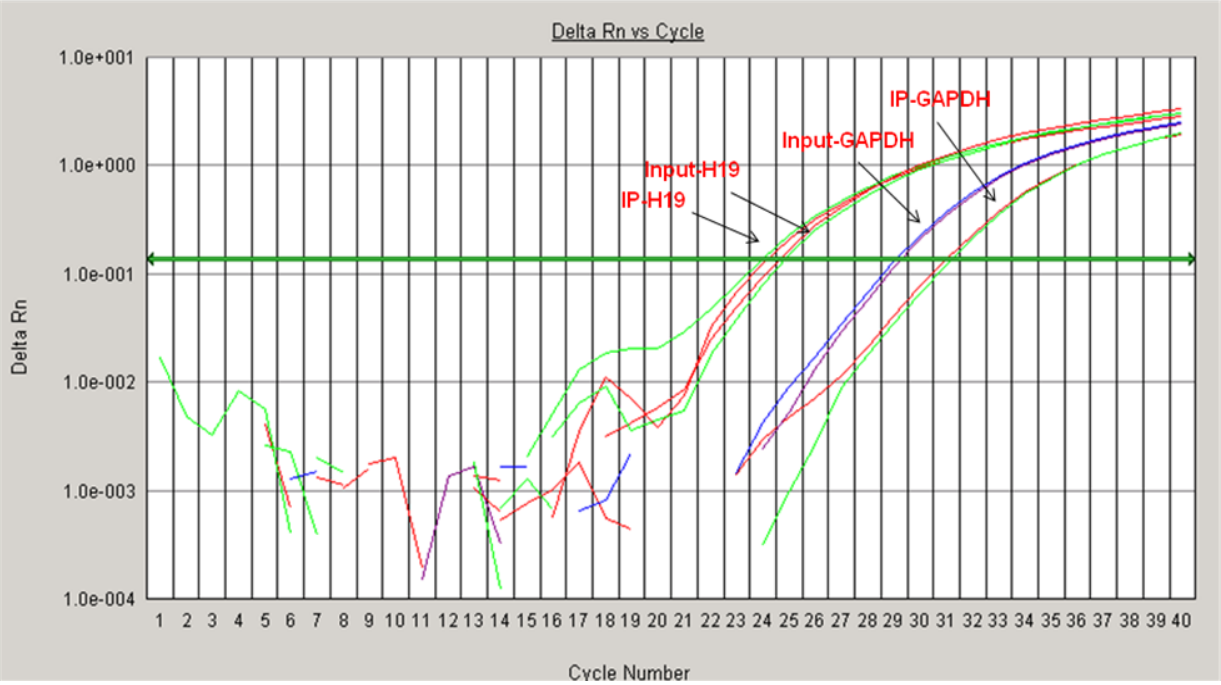

Figure 23
qRT-PCR for *H19* and *GAPDH* to test for enrichment of methylated DNA fragments after immunoprecipitation. Results of qRT-PCR analysis of the cell line A549 after incubation with the primary antibody for 3 hours are shown. Signals for *H19* and *GAPDH* in the IP sample were compared to signals of the input sample. Methylated and unmethylated fragments of the genomic DNA are present in both the IP and the input sample.

Incubation for two hours yielded sufficient DNA for further experiments and kept the specificity of the antibody on a high level.

Equal experiments were conducted for the incubation period of DNA-antibody complexes with protein A agarose beads. 3 different durations of incubation were tested: half an hour, one hour and two hours. After incubating only half an hour the DNA amount was too low for subsequent experiments (<5ng/µl). Setting the incubation time to two hours, we obtained unspecific interaction of protein A agarose beads with unbound DNA

fragments, leading to inefficient exclusion of unmethylated DNA fragments as shown in Figure 24. Methylated as well as unmethylated fragments are present in similar amounts in both the input and the IP sample. Amplification of *H19* is becoming detectable around cycle number 19 in the IP sample and only one cycle later in the input sample. This approach is not sufficient to enable the exclusion of unmethylated fragments from the IP sample and to enrich methylated ones in the IP sample. The fold change of enrichment was in this case only 38.45-fold.

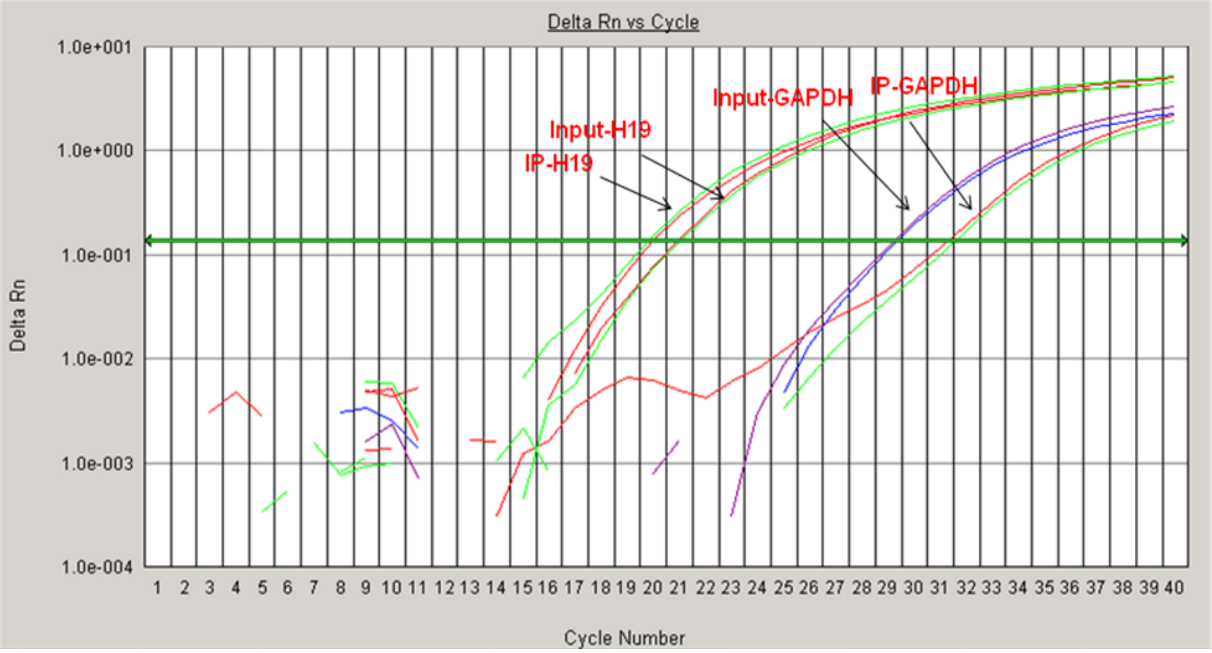

Figure 24
qRT-PCR for *H19* and *GAPDH* to test for enrichment of methylated DNA fragments after immunoprecipitation. Results of qRT-PCR analysis of the cell line A549 after incubation with protein A agarose beads for 2 hours are shown. Signals for *H19* and *GAPDH* in the IP sample were compared to signals of the input sample. Two hours of incubation allows unspecific interaction of the DNA fragment/antibody complex with the protein A agarose beads.

Incubation with protein A agarose beads was optimally set to one hour because of minimal unspecific binding events and sufficient DNA yield for subsequent experiments. During the consequent washing steps unbound DNA is eliminated from the IP sample. Insufficient washing of DNA/antibody/agarose complexes leads to remaining of unbound DNA within the sample. This situation was observed after washing only 3 times. Exclusion from the unbound DNA was achieved after 5 washing steps. Thus, 5 washing steps were set as optimal parameter.

In summary, optimal parameters for efficient enrichment of methylated DNA fragments by MeDIP are as follows: sonication of genomic DNA for 10 minutes, followed by incubation with 3.5 µg primary antibody for two hours and a consequently incubation for one hour with protein A agarose beads. Unmethylated fragments are eliminated with 5 washing steps. The graph of the qRT-PCR result for the cell line A549 is depictured in Figure 25. In this case enrichment of methylated fragments was archived in the IP sample in comparison to the input sample. Unmethylated fragments are present in the input sample but not in the IP sample. Amplification of *GAPDH* in the IP sample is only starting after the 40^{th} cycle, and can be seen as unspecific amplification. The fold change of enrichment for this sample amounted to 4069-fold.

Table 4
Summary of the optimization process of MeDIP approach. Duration and intensity of DNA sonication, duration of the incubation steps, number washing steps were optimized.

DNA sonication	5 minutes	Most DNA fragments are too long (up to 2000bp). Methylated and unmethylated regions can lie on one fragment
	10 minutes	Optimal length of DNA fragments (200bp - 800bp)
	15 minutes	DNA fragments are too small. Thus, consequent amplification is inefficient (100bp – 600bp)
Incubation time with primary antibody (AB)	1 hour	Time is too short for optimal DNA binding resulting in very low amounts of recovered DNA (<5ng/µl).
	2 hours	Optimal incubation period. Efficient enrichment of methylated DNA with a low amount of unspecific binding.
	3 hours	Efficient enrichment of methylated DNA with a high amount of unspecific binding. (see Figure 23)
Incubation time with Protein A beads	30 minutes	Subsequent qRT-PCR reveal that the amount of methylated DNA is significantly reduced (<5ng/µl).
	1 hour	Subsequent qRT-PCR show adequate amount of methylated DNA, unmethylated DNA is mostly

		excluded
	2 hours	Subsequent qRT-PCR displays that longer incubation increases the number of unspecific bonds (see Figure 24)
Number of washing steps	10	Unbound DNA is eliminated, very time consuming, minimal amounts of methylated DNA are lost in every further step
	5	Unbound DNA is eliminated; only minimal loss of methylated DNA
	3	Subsequent qRT-PCR reveals significant amount of unmethylated DNA, washing was insufficient

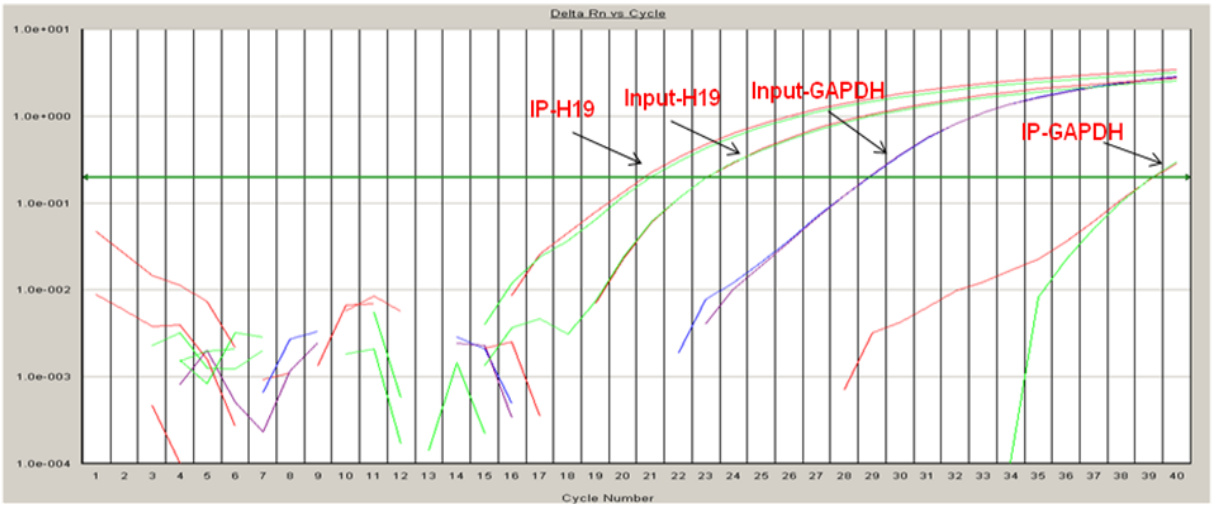

Figure 25
The graph depicturing the result of the qRT-PCR analysis of the cell line A549 after MeDIP under optimal conditions. The IP sample was compared to the Input sample. Primers specific for H19 and primers specific for GAPDH were used for this approach. Enrichment of methylated regions is confirmed shown by H19. Unmethylated regions shown by GADH are only produced unspecific.

7.1.2.1 MeDIP-chip analysis

After successful establishing the MeDIP assay the lung adenocarcinoma cell lines A549, NCI-H1993 and NCI-H2073 and Normal Human Bronchial Epithelial Cells (NHBEC) were consequently analyzed by MeDIP-chip. Labeling and microarray hybridization were performed by "NimbleGen" using "HG18 CpG plus Promoter" arrays comprising 385.019 probe sets. These probes cover all reported Human Refseq gene promoters (24.659) and all reported CpG islands (28.226) annotated on the UCSC genome browser. Arrays were scanned on a GenePix 4200A Scanner using GenePix 6.0 Software and data were extracted from scanned images by using NimbleScan 2.3 extraction software (NimbleGen Systems, Inc.). Raw data (.pair) files were used for detailed statistical analysis as described below. In addition, processed data files for each cell line analyzed were generated and provided by NimbleGen using default setting from NimbleScan 2.3 software.

To verify the MeDIP-chip results NimbleGen includes genes into the analyses with known methylation status. Since, methylation of *H19* is known to occur in all somatic cells this gene was used as our positive control. The house keeping gene *GAPDH*, which is unmethylated in all somatic cells was used as negative control. All cell lines revealed methylation of the *H19*

promoter region. The promoter region of GAPDH was unmethylated in all cell lines analyzed (see Figure 26).

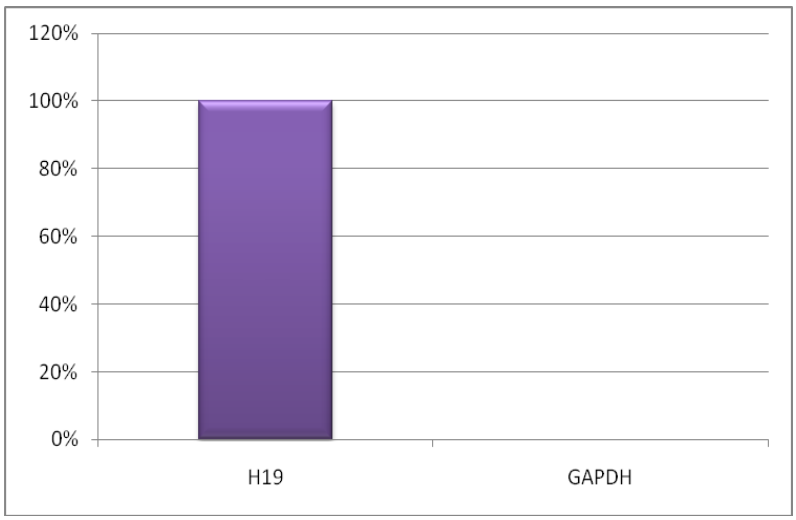

Figure 26
Quality control step: NimbleGen includes genes with known methylation status into the microarray analyses to confirm the obtained data. Example genes are *GAPDH* and *H19*. *GAPDH* which is a house keeping gene and in this position never methylated in somatic cells was unmethylated all cell lines analyzed. *H19* which is a development gene and in this function methylated and silenced in all somatic cells was methylated in all cell lines analyzed.

Statistic analysis of MeDIP-chip data was performed in cooperation with biostatisticians from the Medical University of Vienna, Section of Medical Statistics, Vienna, Austria under the supervision of Prof. Dr. Martin Posch. Differentially methylated regions between lung cancer cell lines and NHBECs were determined as described in 6.5.7. Due to statistic deviation the cell line NCI-H2073 had to be excluded from further analysis.

This statistical analysis revealed that 31.786 probes were differentially methylated between lung cancer cell lines and NHBECs with a p-value <0.05. After correction for multiple testing (Benjamini-Hochberg) 375 probes remained with a p-value <0.05. These probes were annotated and dereplicated yielding 47 unique tumor-specifically methylated protein encoding genes. Detailed information about these genes is shown in Table 5.

Table 5
47 tumor-specifically methylated genes identified by MeDIP-chip analysis in NSCLC cell lines

Accession[a]	Gene name	Gene symbol	Cytoband
AK098187	Chromosome 14 open reading frame 39	*C14orf39*	14q23.1
AAH15763	Epidermal growth factor receptor pathway substrate 8-like protein 1-like	*EPS8L1*	19q13.42
NP_001073927	Homeo box even-skipped homolog protein 2	*EVX2*	2q31.1
Q6ZQR6	n/a	*FLJ45983*	10p14
U13223	Forkhead box D4	*FOXD4*	9p11-q11
AF452723	Forkhead box D4 like 1	*FOXD4L1*	2q14.1
AL031777	Histone 1, H2ae	*HIST1H2AE**	6p22.2-p21.1
AF531289	Histone 1, H2bf	*HIST1H2BF*	6p21.3
NM_005519.1	Homeo box (H6 family) 2	*HMX2**	10q25.2-q26.3
AI581335	Homeo box A2	*HOXA2**	7p15-p14
AK056567	Homeo box A3	*HOXA3*	7p15-p14
AI277552	Homeo box A4	*HOXA4*	7p15-p14
AF010258	Homeo box A9	*HOXA9**	7p15-p14
AF016029	Homeo box B3	*HOXB3*	17q21.3
AF069333	Insulin-like growth factor 2 receptor	*IGF2R*	6q26
AF319967	Iroquois Homeobox protein 2	*IRX2**	5p15.33
BC020470	LIM Homeobox 1	*LHX1*	17q12

AJ277915	LIM Homeobox 9	*LHX9**	1q31-q32
XR_041919.2	Phosphodiesterase 4D interacting protein-like	*LOC199882*	1q21.2
AK057663	Hypothetical protein LOC285074	*LOC285074*	2p11.1
XM_370758.3	Hypothetical gene supported by BX248251	*LOC387978*	14q12
XM_931960	Hypothetical protein LOC643982	*LOC643982*	17q25.3
XM_928920.1	Homo sapiens similar to forkhead box D4 like 3, transcript variant 1	*LOC653427*	9q21.11
BX537794	Nuclear factor I/X (CCAAT-binding transcription factor)	*NFIX*	19p13.3
AB007948	Nicotinamide nucleotide adenylyltransferase 2	*NMNAT2**	1q25
NM_145260.2	Odd-skipped related 1	*OSR1*	2p24.1
AB037505	Orthodenticle homolog 2 (Drosophila)	*OTX2*	14q21-q22
AK074881	Paired box gene 6 (aniridia, keratitis)	*PAX6*	11p13
AF152305	Protocadherin alpha 1	*PCDHA1**	5q31
AF152306	Protocadherin alpha 10	*PCDHA10**	5q31
AF152307	Protocadherin alpha 11	*PCDHA11**	5q31
AF152308	Protocadherin alpha 12	*PCDHA12**	5q31
AF152309	Protocadherin alpha 13	*PCDHA13**	5q31
AF152310	Protocadherin alpha 2	*PCDHA2**	5q31
AF152311	Protocadherin alpha 3	*PCDHA3**	5q31
AF152312	Protocadherin alpha 4	*PCDHA4**	5q31
AF152313	Protocadherin alpha 5	*PCDHA5**	5q31

AF152314	Protocadherin alpha 6	*PCDHA6**	5q31
AF152315	Protocadherin alpha 7	*PCDHA7**	5q31
AF152316	Protocadherin alpha 8	*PCDHA8**	5q31
AC005609	Protocadherin alpha 9	*PCDHA9**	5q31
AB015671	Paired-like Homeobox 2b	*PHOX2B*	4p12
AF048720	Paired-like Homeo domain transcription factor 2	*PITX2*	4q25-q27
NM_004575	POU domain, class 4, transcription factor 2	*POU4F2*	4q31.2
AF022654	Short stature Homeobox 2	*SHOX2**	3q25-q26.1
AAH36458	Solute carrier family 32 (GABA vesicular transporter), member 1	*SLC32A1*	20q11.23
NM_152476	Zinc finger protein 560	*ZNF560*	19p13.2

[a] GenBank accession number

* Asterisks indicate genes which are identified by MeDIP-chip and whose expression was up-regulated after Aza-dC or Aza-dC/TSA treatment

7.1.2.2 Functional analysis of tumor-specifically methylated genes

We used Ontologizer (171) to functionally categorize the 47 tumor-specifically methylated genes. As shown in Figure 27 two major groups of tumor-specifically methylated genes were identified: transcription factors (represented as GO:0030528, transcription regulator activity) and cell adhesion molecules (represented as GO:0005509, calcium ion binding). Detailed results of GO analysis are shown in Table 6.

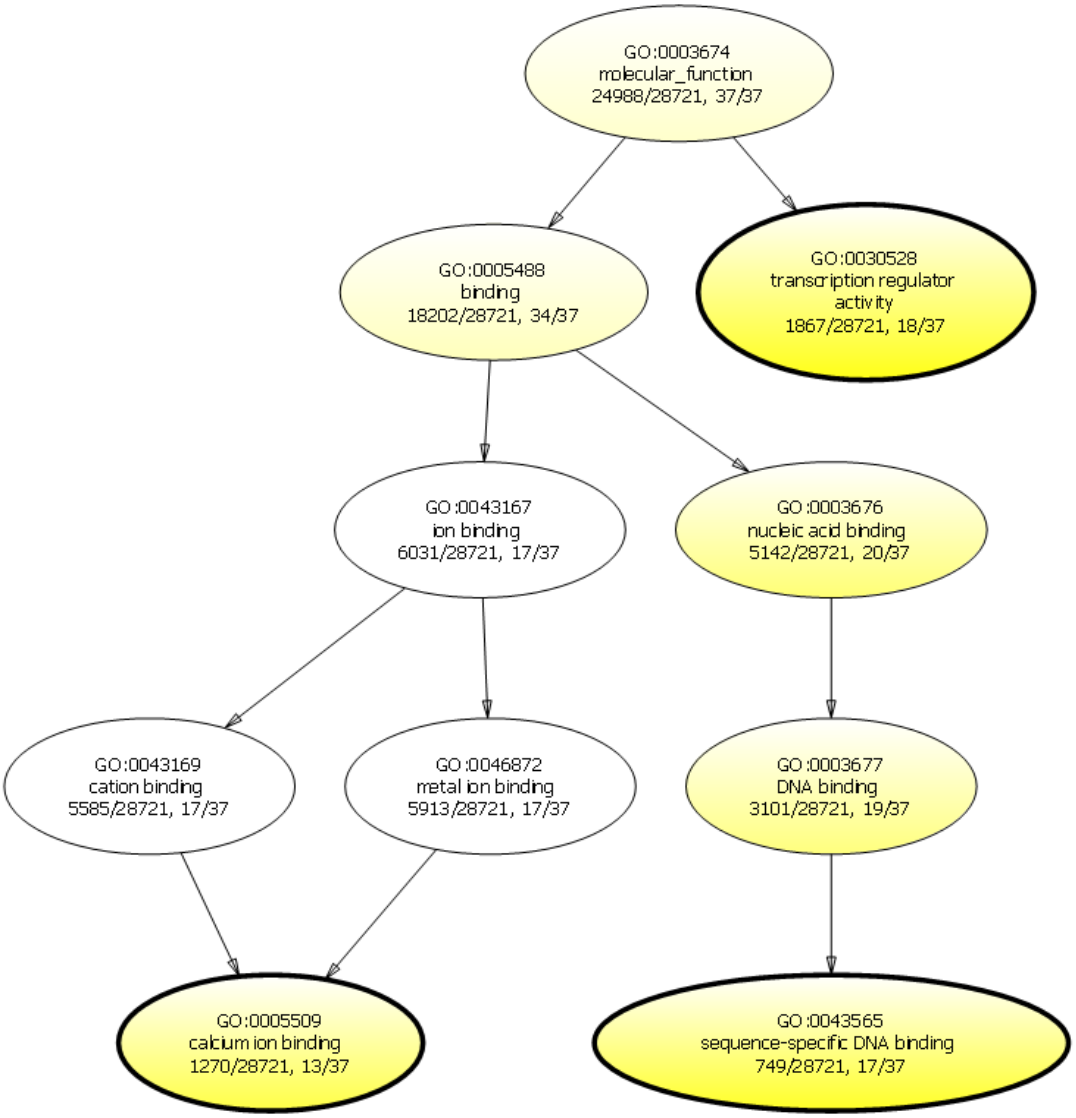

Figure 27
Enrichment of GO terms by means of molecular function. Enriched nodes are colored in different intensities of yellow depending on adjusted p-values (more significant enrichment corresponds to more intense yellow). First ratio corresponds to the proportion of GO terms in the human genome and the second ratio corresponds to the proportion of GO terms in the study group.

Table 6
Functional characterization of 47 tumor-specifically methylated genes according to GO categories

GO ID	Biological process	Study Count	Population Count	Adjusted p-value
GO:0032502	Developmental process	26	3529	2.56E-10
GO:0032501	Multicellular organismal process	26	3926	1.63E-09
GO:0022610	Biological adhesion	13	879	1.30E-07
GO:0009058	Biosynthetic process	20	5866	2.65E-05
GO:0006139	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	20	5056	2.65E-05
GO:0016337	Cell-cell adhesion	13	343	0.00018279
GO:0008150	Biological_process	37	21165	0.0004549
GO:0007156	Homophilic cell adhesion	13	167	0.00191185
GO:0007224	Smoothened signaling pathway	2	22	0.00225362
GO:0010467	Gene expression	19	5014	0.0167489
GO:0048856	Anatomical structure development	24	2232	0.0167489
GO:0046530	Photoreceptor cell differentiation	2	11	0.0299079
GO:0034960	Cellular biopolymer metabolic process	19	5594	0.0299079
GO:0007399	Nervous system development	17	837	0.0380023

GO ID	Molecular function	18	1867	2.67E-09
GO:0043565	Sequence-specific DNA binding	17	749	2.14E-07
GO:0005509	Calcium ion binding	13	1270	0.00018279
GO:0003676	Nucleic acid binding	20	5142	0.00453889
GO:0003677	DNA binding	19	3101	0.0129824
GO ID	**Molecular function**	14	3812	0.00161287
GO:0005575	Cellular_component	37	22009	0.00161287
GO:0005634	Nucleus	21	6172	0.0299079

The study count is the total number of genes with the corresponding GO annotation which were detected as tumor-specifically methylated in NSCLC cell lines
The total count is the total number of genes present on the NimbleGen´s HG18 CpG plus Promoter microarray which have the corresponding GO annotation.
The adjusted p-value is the calculated p-value for the enrichment analysis adjusted for multiple comparisons using the Bonferroni method.

We showed that tumor-specific methylation of transcription factors is present in lung cancer cell lines. 19 of the 47 genes identified to be methylated in lung cancer cell lines could be affiliated to the molecular function of transcription factors. One important super family of transcription factors is the homeo box family. DNA methylation within the promoter region of 4 genes out of 11 genes within the gene cluster HOXA was identified. *HOXA2*, *HOXA3*, *HOXA4* and *HOXA9* are methylated in their promoter region in the NSCLC cell lines A594 and NCI-H1993 but show only low methylation signal in the NHBE cells analyzed (see Figure 28). Additional genes like *IRX2*, *LHX1*, *LHX9* and *OTX2* were affiliated to the super family of transcription factors. Since in most cases one transcription factor regulates several other genes, these 19 genes enlarge the scope of these findings tremendously.

28 genes were shown to be important for different development processes like cell proliferation or cell death. Both pathways are known to be de-regulated to a certain extend in cancer. *HMX2*, *HOXA3*, *LHX1*, *LHX9* and *PHOX2B* were shown to be methylated in the NSCLC cell lines but not in the NHBE cells and all of them are important for the regulation of cell proliferation. Genes like *IGF2R*, *OSR1* which are taking part in the regulation of the cell death were also among the genes identified as tumor-specifically methylated.

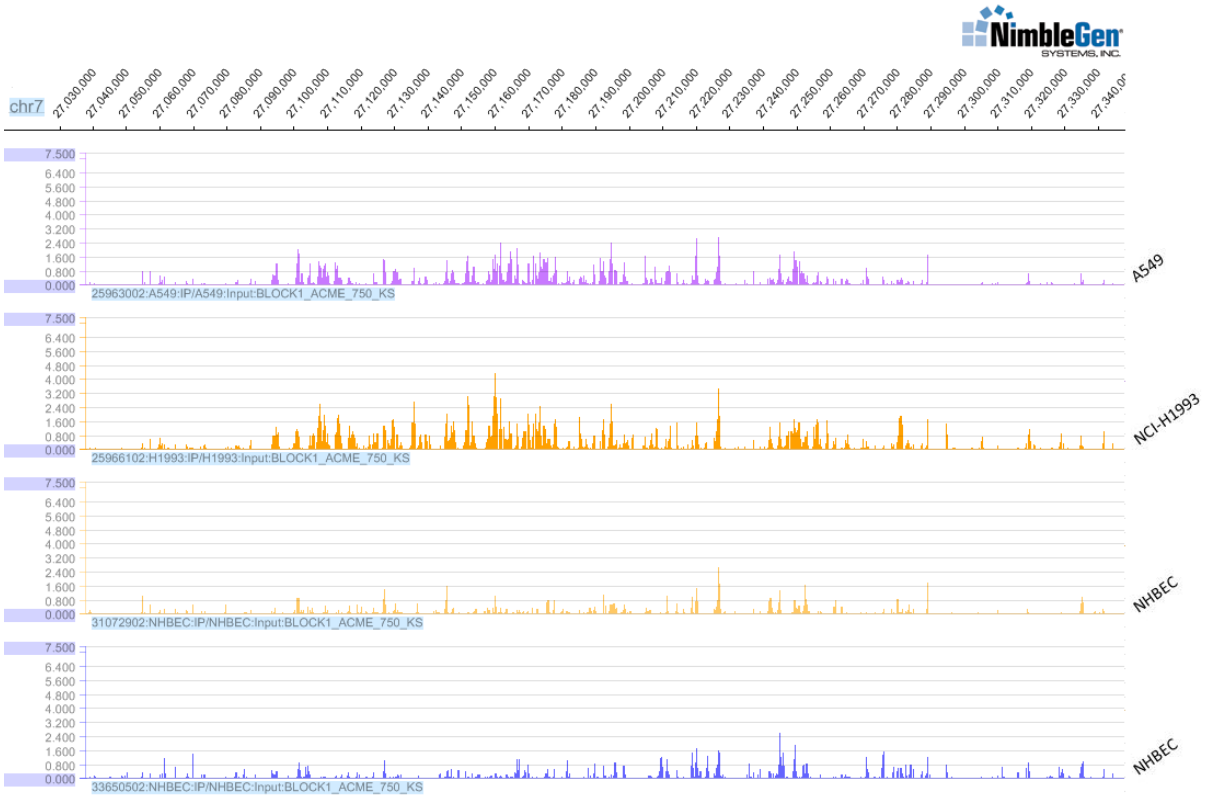

Figure 28
Result sheet obtained from NimbleGen; the region of the HOXA gene cluster on chromosome 7 is shown. Results for the NSCLC cell lines A549 and NCI-H1993 are compared to the results for NHBE cells. Strong methylation signal in the promoter region of the genes HOXA2, HOXA3, HOXA4 and HOXA9 is shown in the NSCLC cell lines.

Genes like *PCDHA1* and *PCDHA9* which are important for cell adhesion were identified as tumor-specifically methylated. Protocadherin genes are encoded in clusters. All 13 *PCDHA* genes are encoded in the cluster PCDH alpha. Our results reveal that all members of the PCDHA cluster (*PCDHA1*, *PCDHA2*, *PCDHA3*, *PCDHA4*, *PCDHA5*, *PCDHA6*, *PCDHA7*, *PCDHA8*, *PCDHA9*, *PCDHA10*, *PCDHA11*, *PCDHA12* and *PCDHA13*) are tumor-specifically methylated (see Figure 29).

Figure 29
Result sheet obtained from NimbleGen; the region of the PCDHA gene cluster on chromosome 5 is shown. Results for the NSCLC cell lines A549 and NCI-H1993 are compared to the results for NHBE cells. Strong methylation signal in the promoter region of every gene is shown in the NSCLC cell lines.

7.1.2.3 Comparison of gene expression microarray results and MeDIP-chip results of the cell lines A549, NCI-H1993 and NCI-H2073

Comparing results obtained from MeDIP-chip analysis and the genes identified as statistically significant up-regulated after Aza-dC and/or TSA induction in at least one of the three lung adenocarcinoma cell lines, 21 genes were commonly identified. This shows that 45% of the genes identified as methylated by the MeDIP-Chip method could be verified after induction with the epigenetically active drugs Aza-dC and/or TSA. Among these genes known target genes for methylation in caner like *HOXA9* and *IRX2* were found to be methylated by both techniques (174). *OTX2* which is a transcriptional regulator has been identified as important for cell cycle regulation and induction of senescence (175).

Tumor-specific methylation of 8 randomly selected genes was confirmed by MSP in primary lung tumor samples and corresponding non-malignant lung tissue samples of 10

NSCLC patients. As example *OTX2* was methylated in 80% of the analyzed tumors. In contrast, none of the non-malignant tumor samples showed methylation of the *OTX2* promoter region. Also methylation of the promoter region of *HOXA9* was identified to be tumor-specifically distributed. 90% of the tumor samples analyzed revealed methylation of the *HOXA9* promoter region in comparison to methylation of 10% of the promoter region in non-malignant lung tissue samples. MSP results obtained by analyzing 10 primary NSCLCs and matching non-malignant lung tissue samples for methylation of *OTX2* and *HOXA9* are shown in Figure 30. MSP results for additional genes are shown in Figure 21.

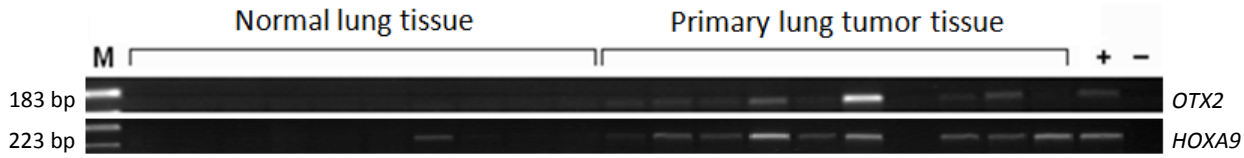

Figure 30
MSP analyses of methylation frequency of OTX2 and HOXA9 in primary lung tumor samples and corresponding non-malignant lung tissue samples of 10 NSCLC patients. All genes show a high degree of methylation in the tumor samples but only no or rare methylation signals in the corresponding normal tissue; + positive control; - negative control; M 100bp ladder

8 DISCUSSION

Changes in the epigenetic code, especially in DNA methylation patterns, are frequently occurring events in the pathogenesis of malignant diseases. To date about 100-150 genes are known to be aberrantly methylated in NSCLC. Most of this knowledge was gathered with time consuming DNA methylation analysis on single gene level. However, Costello et al. (4) reported previously that approximately 600 CpG islands are methylated in tumor cells. Thus, identification of novel target genes for methylation is of tremendous importance for the understanding of the molecular mechanisms leading to NSCLC. With the development of the microarray technique several years ago it became possible to analyze gene expression patterns in a genome-wide manner. Because DNA methylation, which is a reversible process, significantly contributes to the silencing of gene expression a combination of a pharmacological DNA demethylation assay combined with gene expression microarrays is a useful strategy to identify so far unknown methylated genes (12-14). Since this is a method which indirectly identifies methylated genes false positive results have to be expected. We used this approach on three NSCLC cell lines to better understand the impact of changes in the DNA methylation pattern on the pathogenesis of NSCLC. A major disadvantage of this technique is that it is only applicable to cells in cell culture but not to tissue samples. Cell lines are an optimal tool to get a first insight into a subject since cell lines are easy to cultivate, are highly available and depicture many features of *in vivo* cells. However, cells in cell culture can accumulate aberrations in gene expression, morphology, metabolism and DNA methylation (124, 176, 177). This suggests that research results obtained only from cell lines may not accurately reflect the situation in a cancer cell *in vivo*.

In recent years a novel genome-wide approach for methylation analysis called MeDIP-chip was developed. MeDIP-chip has the advantage to be applicable to tissue samples making a direct comparison of DNA methylation pattern of primary tumors and corresponding non-malignant tissue samples possible.

The aim of this project was to get more insight into the changes of DNA methylation patterns during pathogenesis of NSCLCs. To archive this aim we performed two different types of whole genome DNA methylation analyses. Additionally, selected genes were further

analyzed with the gene specific method MSP. In this way an optimal verification of the results obtained was ensured.

At first we analyzed the gene expression profile of three NSCLC cell lines before and after treatment with the epigenetically active drugs Aza-dC and/or TSA and compared them to each other. A similar approach was already applied to NSCLC cell lines by Shames et al. (154). They identified 132 tumor-specifically methylated genes by analyzing expression of ~47.000 transcripts after treatment of 7 NSCLC cell lines with the demethylating drug 5-aza-2´-deoxycytidine (Aza-dC) and by comparing these results with data of 3 normal human bronchial epithelial cells (NHBEC). Additionally, Zhong et al. (178) identified 214 genes whose expression was up-regulated after drug treatment by using a similar approach and treating 9 NSCLC cell lines with Aza-dC and TSA. Our results identified a large number of genes to be up-regulated after drug treatment.

Many of the up-regulated genes were found to be involved in cancer-related molecular pathways such as cell cycle, cell adhesion and apoptosis. Thus far, epigenetic silencing was unknown for most of these genes. However, we also identified genes whose epigenetic silencing in lung cancer has already been reported including *APC*, *p16*, *RASSF4*, *RASSF5*, *TIMP2* and *HIC1* (179, 180). Seven genes which were synergistically up-regulated after Aza-dC and TSA induction were selected for further MSP analyses of primary lung tumor samples and corresponding non-malignant lung tissue samples of 10 NSCLC patients. These analyses revealed tumor-specific methylation of six out of seven genes analyzed. The most frequently methylated genes in the primary NSCLCs were *THSD4* and *TMTC1*.

The function of *TMTC1* is nearly unknown. Due to the identification of tumor-specific methylation of *TMTC1* our data indicate that this gene is involved in the pathogenesis of NSCLCs. However, additional functional studies are necessary to characterize the association between *TMTC1* and cancer development.

Additionally genes like *ALDH1A* and *NEFH* were identified to be tumor-specifically methylated. *ALDH1A3* has previously been shown to be a target for methylation in different types of cancer including gastric and breast cancer. As reported previously changes in expression of this gene can be set into relation to progression of liver, gastric or breast carcinomas (181). Our results indicate that this observation may also be true for NSCLC. Interestingly, methylation of *NEFH* has recently been correlated with duration of tobacco

smoking and esophageal squamous cell carcinomas (ESCC) (182). Further studies are needed to determine if *NEFH* methylation already occurs in pre cancerous lesions of the lung. If *NEFH* can be found to be methylated in these tissues, this gene might be potentially helpful for early diagnosis of NSCLCs.

Recent studies revealed that *DLX4* has metastatic inhibiting function (183). *DLX4* was frequently methylated in primary lung tumor samples but was methylated only in 10% of the corresponding non-malignant lung tissue samples analyzed. Epigenetic inactivation of this gene might favor the metastatic potential of lung cancer cells.

Besides tumor-specific methylation of certain genes we also observed that drug treatment of cells up-regulate genes whose methylation is not tumor-specific. As seen in Figure 21 methylation of the gene *MACF1* was not only found in primary tumor samples but also in non-malignant lung tissue samples. Additional studies are necessary, to investigate *MACF1* methylation.

By comparing expression microarray results obtained from Aza-dC treated cells with results obtained from TSA treated cells, synergistically up-regulated genes were identified. A synergy in induction of gene expression of these two drugs was already observed (151). Several of the synergistically up-regulated genes are involved in cancer related pathways. One of them is *HEXIM1* which shows transcription factor activity, and negatively regulates the cell cycle as well as the expression of vascular endothelial growth factor in an estrogen dependent manner (184). Recently, *HEXIM1* de-regulation in ovarian cancer was identified but the mechanism of de-regulation for this gene is unknown so far (185).

Identifying *NOX5* to be synergistically up-regulated after Aza-dC and TSA is standing in controversy with findings from Hong et al. (186). They identified *NOX5* to induce pathogenesis of esophageal adenocarcinoma by mediating production of reactive oxygen species and cell proliferation (186).

ATM which has been shown to possess a key role in regulation of DNA replication in stress related situations was tumor-specifically methylated in the cell lines analyzed (187). There is still limited knowledge about the function of *ATM* in cancer and thus further studies are necessary. Besides genes whose involvement in cancer has already been investigated, we identified a large number of genes whose role in cancer has been unknown so far.

Further expression analysis as well as functional analysis for the cohort of genes analyzed would be of high interest.

Approximately 90% of the genes up-regulated after Aza-dC treatment were also found to be up-regulated in cells after TSA treatment. These findings not only confirm the already known fact of synergy between Aza-dC and TSA but also emphasize the knowledge about the strong linkage between DNA methylation and histone protein modification.

667 genes were synergistically up-regulated in at least one NSCLC cell line analyzed. 73% of these genes are involved in cancer related pathways including genes like *MAL* and *PCAF. PCAF* encodes a histone acetyltransferase which interacts physically and regulates the tumor suppressor PTEN. One important function of PTEN is to arrest cell cycle within the G1 phase and in this way inhibiting formation of malignancy (188). In addition, *PCAF* is known to be aberrantly methylated in esophageal squamous cell carcinoma.

RECK which is a membrane-anchored glycoprotein important for regulation of different metalloproteinasesis in normal cells was among the genes identified to be up-regulated after induction with epigenetic active drugs. It is known to be down-regulated in several cancer types and to be a target for methylation in NSCLC. Chang et al. (189) identified a strong correlation between inactivation of *RECK* and formation of lymph node metastasis. In most of the cases down regulation is achieved by aberrant DNA methylation in the promoter region. *RECK* is even considered as gene of interest for prognosis of peritoneal metastasis of gastric cancer (190). Further studies are necessary.

The second whole genome approach for analyzing DNA methylation patterns of the cell lines A549, NCI-H1993, NCI-H2073 and NHBEC is the MeDIP-chip method. Major advantages of this method are applicability to tissue samples. Additionally, the MeDIP-chip technique identifies methylation directly in contrast to approaches using epigenetically active drugs which re-express methylated genes. Since the identification of genes in the MeDIP-chip methods is based on immunoprecipitation step of methylated DNA fragments.

To test for efficient enrichment of methylated DNA fragments two quality control steps were performed. The first one was a qRT-PCR step using primers for genes known to be methylated and primers for genes known to be unmethylated. Housekeeping genes are known to be unmethylated in many types of somatic cells. *GAPDH* is a very famous representative of housekeeping genes. Thus, we chose this gene as reference gene in qRT-

PCR analyses to identify and quantify unmethylated fragments in the samples. *H19* was used as reference gene for methylated DNA fragments since, *H19* is completely silenced by DNA methylation after completion of different developmental processes (191, 192). The combination of these genes in a qRT-PCR allows an optimal comparison of the presence of methylated and unmethylated DNA fragments. As a second quality control step these two genes were included into the microarray as positive and negative control.

Analyzing NSCLC cell lines with the MeDIP-chip method revealed a huge amount of genes important in cancer related pathways. We identified tumor-specific methylation of single genes as well as tumor-specific methylation of whole gene clusters. In particular, members of the PCDHA and the HOXA cluster were found to be frequently methylated. All 13 representatives of *PCDHA,* namely *PCDHA1*, *PCDHA2*, *PCDHA3*, *PCDHA4*, *PCDHA5*, *PCDHA6*, *PCDHA7*, *PCDHA8*, *PCDHA9*, *PCDHA10*, *PCDHA11*, *PCDHA12*, and *PCDHA13* were methylated in the NSCLC cell lines but were unmethylated in the NHBEC cell line analyzed. PCDH genes can be subdivided into clustered and non-clustered protocadherins. All genes encoded in the PCDHA gene cluster on chromosome 5 take part in the formation process of protocadherin. Overall, knowledge about the involvement of clustered protocadherins in cancer is limited (193). Recently, *PCDHA* was identified as target for aberrant methylation in breast cancer (194, 195). Moreover, methylation of protocadherin clusters was investigated in Wilms tumors using the MeDIP-chip approach revealing that most members of the PCDHA, PCDHB and PCDHC clusters are tumor-specifically methylated (196). However, in our study we observed that only members of the PCDHA clusters are tumor-specifically methylated but not members of the PCDHB and PCDHG. Its relevance in NSCLC is unknown so far.

Members of the homeo box protein family were identified as a second group of clustered genes which are frequently methylated in the NSCLC cell lines analyzed. In general, HOX genes are members of a transcription factor family and are involved in embryonic development and in the control of differentiation of adult hematopoietic cells (197). 85% of genes encoded in the HOXA gene cluster were tumor-specifically methylated in the cell lines, analyzed suggesting that methylation of members of these gene cluster seems to be crucial for the pathogenesis of NSCLC. This observation is in concordance with previously reported data. Rauch et al. (198) reported that all 4 HOX gene clusters on chromosome 2, 7, 12 and 17 are preferential targets for methylation in lung cancer cell lines. Shiraishi et al. (199)

analyzed methylation of HOXA and HOXD cluster members in primary tumor samples and matching non-malignant lung tissue samples of 8 lung adenocarcinoma patients and found several of them to be tumor-specifically methylated. Identifying HOX genes to be tumor-specifically methylated in our study suggests that HOX genes may play an important role in the pathogenesis of NSCLC (Heller et al. unpublished).

These findings propose that not only expression of single genes is affected by DNA methylation but also whole gene clusters are regulated by DNA methylation in lung cancer. Further analyses are necessary to clarify if methylation of single genes attracts methylation to genes in the same cluster or if these genes were methylated independently from each other.

Additionally, genes that are involved in either de-regulation of cell proliferation or cell death were found to be tumor-specifically methylated in the cell lines analyzed. 27 of the tumor-specifically methylated genes were affiliated to pathways important for these two physiological processes. Differential expression of *PHOX2B* was identified in a large number of studies about neuroblastomas (200, 201). To our knowledge, this is the first report showing that *PHOX2B* is silenced by DNA methylation and is related to the pathogenesis of NSCLC.

21 genes commonly identified to be methylated by both methods used. These genes represent 45% of the genes identified as methylated by the MeDIP-Chip method. For example *HOXA9* is known to be a transcription factor involved in embryogenesis and differentiation of adult cells (198). *IRX2* encodes an iroquois-class homeo domain protein and is another transcription factor (202, 203).

4560 genes were identified as up-regulated in at least one cell line after drug treatment proposing epigenetic regulation as important factor for their regulation. MeDIP-chip analysis identified 47 genes to be methylated in a statistically relevant manner. Only 0.5% of genes identified by Aza-dC/TSA treatment were additionally identified by the MeDIP-chip method. This situation can be related to the fact that treatment with Aza-dC and/or TSA does not only alter DNA methylation patterns of cells but is also a stress factor for treated cells. This leads to the conclusion that identifying DNA methylation by indirect methods gives an excellent first insight into DNA methylation patterns but results have to be verified by direct DNA methylation analysis to identify false positive results.

Overall, our results stress the importance of DNA methylation for the pathogenesis of NSCLC. This project is one step on the long path of understanding the molecular mechanism of NSCLC, however many further studies are necessary.

9 REFERENCES

1. http://cyberbridge.mcb.harvard.edu/dna_2.html.
2. Nowell PC, Croce CM. Chromosomes, genes, and cancer. Am J Pathol 1986; 125: 7-15.
3. Flatau E, Bogenmann E, Jones PA. Variable 5-methylcytosine levels in human tumor cell lines and fresh pediatric tumor explants. Cancer Res 1983; 43: 4901-5.
4. Baylin SB, Hoppener JW, de Bustros A, Steenbergh PH, Lips CJ, Nelkin BD. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. Cancer Res 1986; 46: 2917-22.
5. Schulz WA, Steinhoff C, Florl AR. Methylation of endogenous human retroelements in health and disease. Curr Top Microbiol Immunol 2006; 310: 211-50.
6. Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet 1999; 21: 163-7.
7. Esteller M, Sanchez-Cespedes M, Rosell R, Sidransky D, Baylin SB, Herman JG. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res 1999; 59: 67-70.
8. Usadel H, Brabender J, Danenberg KD, et al. Quantitative adenomatous polyposis coli promoter methylation analysis in tumor tissue, serum, and plasma DNA of patients with lung cancer. Cancer Res 2002; 62: 371-5.
9. Fujiwara K, Fujimoto N, Tabata M, et al. Identification of epigenetic aberrant promoter methylation in serum DNA is useful for early detection of lung cancer. Clin Cancer Res 2005; 11: 1219-25.
10. Wang Y, Yu Z, Wang T, Zhang J, Hong L, Chen L. Identification of epigenetic aberrant promoter methylation of RASSF1A in serum DNA and its clinicopathological significance in lung cancer. Lung Cancer 2007; 56: 289-94.
11. Belinsky SA, Klinge DM, Dekker JD, et al. Gene promoter methylation in plasma and sputum increases with lung cancer risk. Clin Cancer Res 2005; 11: 6505-11.
12. Wang YC, Lu YP, Tseng RC, et al. Inactivation of hMLH1 and hMSH2 by promoter methylation in primary non-small cell lung tumors and matched sputum samples. J Clin Invest 2003; 111: 887-95.
13. Belinsky SA, Liechty KC, Gentry FD, et al. Promoter hypermethylation of multiple genes in sputum precedes lung cancer incidence in a high-risk cohort. Cancer Res 2006; 66: 3338-44.
14. Chan EC, Lam SY, Tsang KW, et al. Aberrant promoter methylation in Chinese patients with non-small cell lung cancer: patterns in primary tumors and potential diagnostic application in bronchoalevolar lavage. Clin Cancer Res 2002; 8: 3741-6.
15. Costello JF, Fruhwald MC, Smiraglia DJ, et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet 2000; 24: 132-8.
16. Weinhold B. Epigenetics: the science of change. Environ Health Perspect 2006; 114: A160-7.

17. Bhaumik SR, Smith E, Shilatifard A. Covalent modifications of histones during development and disease pathogenesis. Nat Struct Mol Biol 2007; 14: 1008-16.

18. Ouyang J, Gill G. SUMO engages multiple corepressors to regulate chromatin structure and transcription. Epigenetics 2009; 4: 440-4.

19. Kornberg RD, Lorch Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell 1999; 98: 285-94.

20. Grewal SI, Moazed D. Heterochromatin and epigenetic control of gene expression. Science 2003; 301: 798-802.

21. Fischle W, Wang Y, Allis CD. Histone and chromatin cross-talk. Curr Opin Cell Biol 2003; 15: 172-83.

22. Carruthers LM, Bednar J, Woodcock CL, Hansen JC. Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher-order chromatin folding. Biochemistry 1998; 37: 14776-87.

23. Kim JK, Samaranayake M, Pradhan S. Epigenetic mechanisms in mammals. Cell Mol Life Sci 2009; 66: 596-612.

24. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. Hum Mol Genet 2001; 10: 687-92.

25. Barth TK, Imhof A. Fast signals and slow marks: the dynamics of histone modifications. Trends Biochem Sci; 35: 618-26.

26. Boyes J, Bird A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell 1991; 64: 1123-34.

27. Eden S, Cedar H. Role of DNA methylation in the regulation of transcription. Curr Opin Genet Dev 1994; 4: 255-9.

28. Turner BM. Histone acetylation as an epigenetic determinant of long-term transcriptional competence. Cell Mol Life Sci 1998; 54: 21-31.

29. Brownell JE, Allis CD. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Curr Opin Genet Dev 1996; 6: 176-84.

30. Covault J, Chalkley R. The identification of distinct populations of acetylated histone. J Biol Chem 1980; 255: 9110-6.

31. Mahadevan LC, Willis AC, Barratt MJ. Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. Cell 1991; 65: 775-83.

32. Hill CS, Rimmer JM, Green BN, Finch JT, Thomas JO. Histone-DNA interactions and their modulation by phosphorylation of -Ser-Pro-X-Lys/Arg- motifs. EMBO J 1991; 10: 1939-48.

33. Du TT, Huang QH. [The roles of histone lysine methylation in epigenetic regulation]. Yi Chuan 2007; 29: 387-92.

34. Jenuwein T, Allis CD. Translating the histone code. Science 2001; 293: 1074-80.

35. Strahl BD, Allis CD. The language of covalent histone modifications. Nature 2000; 403: 41-5.

36. Bienvenu T, Chelly J. Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. Nat Rev Genet 2006; 7: 415-26.

37. Robertson KD, Uzvolgyi E, Liang G, et al. The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res 1999; 27: 2291-8.

38. Church D, Pruitt KD. Accessing the human genome. Curr Protoc Hum Genet 2002; Chapter 6: Unit 6 9.

39. Gruenbaum Y, Stein R, Cedar H, Razin A. Methylation of CpG sequences in eukaryotic DNA. FEBS Lett 1981; 124: 67-71.

40. Doerfler W. DNA methylation and gene activity. Annu Rev Biochem 1983; 52: 93-124.

41. Bestor TH, Coxon A. Cytosine methylation: the pros and cons of DNA methylation. Curr Biol 1993; 3: 384-6.

42. Bird AP. CpG-rich islands and the function of DNA methylation. Nature 1986; 321: 209-13.

43. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol 1987; 196: 261-82.

44. Lewis J, Bird A. DNA methylation and chromatin structure. FEBS Lett 1991; 285: 155-9.

45. Miller JH, Coulondre C, Farabaugh PJ. Correlation of nonsense sites in the lacI gene with specific codons in the nucleotide sequence. Nature 1978; 274: 770-5.

46. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860-921.

47. Cross SH, Bird AP. CpG islands and genes. Curr Opin Genet Dev 1995; 5: 309-14.

48. Wang Y, Leung FC. An evaluation of new criteria for CpG islands in the human genome as gene markers. Bioinformatics 2004; 20: 1170-7.

49. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 1993; 366: 362-5.

50. Bestor TH, Verdine GL. DNA methyltransferases. Curr Opin Cell Biol 1994; 6: 380-9.

51. Mischoulon D, Fava M. Role of S-adenosyl-L-methionine in the treatment of depression: a review of the evidence. Am J Clin Nutr 2002; 76: 1158S-61S.

52. Bestor TH. Cloning of a mammalian DNA methyltransferase. Gene 1988; 74: 9-12.

53. Holz B, Klimasauskas S, Serva S, Weinhold E. 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases. Nucleic Acids Res 1998; 26: 1076-83.

54. Lau EY, Bruice TC. Active site dynamics of the HhaI methyltransferase: insights from computer simulation. J Mol Biol 1999; 293: 9-18.

55. Cheng X, Roberts RJ. AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res 2001; 29: 3784-95.

56. www.med.ufl.edu/biochem/keithr/fig1pt1.html.

57. Yoder JA, Soman NS, Verdine GL, Bestor TH. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. J Mol Biol 1997; 270: 385-95.

58. Jeltsch A. On the enzymatic properties of Dnmt1: specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme. Epigenetics 2006; 1: 63-6.

59. Bestor T, Laudano A, Mattaliano R, Ingram V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol 1988; 203: 971-83.

60. Shimamura S, Ishikawa F. Interaction between DNMT1 and DNA replication reactions in the SV40 in vitro replication system. Cancer Sci 2008; 99: 1960-6.

61. Schermelleh L, Haemmer A, Spada F, et al. Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Res 2007; 35: 4301-12.

62. Yoder JA, Bestor TH. A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. Hum Mol Genet 1998; 7: 279-84.

63. Okano M, Xie S, Li E. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res 1998; 26: 2536-40.

64. Dong A, Yoder JA, Zhang X, Zhou L, Bestor TH, Cheng X. Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res 2001; 29: 439-48.

65. Schaefer M, Lyko F. Solving the Dnmt2 enigma. Chromosoma; 119: 35-40.

66. Bestor TH. The DNA methyltransferases of mammals. Hum Mol Genet 2000; 9: 2395-402.

67. Chen T, Li E. Establishment and maintenance of DNA methylation patterns in mammals. Curr Top Microbiol Immunol 2006; 301: 179-201.

68. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet 1998; 19: 219-20.

69. Suetake I, Shinozaki F, Miyagawa J, Takeshima H, Tajima S. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem 2004; 279: 27816-23.

70. Watt F, Molloy PL. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev 1988; 2: 1136-43.

71. Campanero MR, Armstrong MI, Flemington EK. CpG methylation as a mechanism for the regulation of E2F activity. Proc Natl Acad Sci U S A 2000; 97: 6481-6.

72. Iguchi-Ariga SM, Schaffner W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev 1989; 3: 612-9.

73. Jubb AM, Bell SM, Quirke P. Methylation and colorectal cancer. J Pathol 2001; 195: 111-34.

74. Meehan RR, Lewis JD, Bird AP. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Res 1992; 20: 5085-92.

75. Clouaire T, Stancheva I. Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cell Mol Life Sci 2008; 65: 1509-22.

76. Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA. A family of human zinc finger proteins that bind methylated DNA and repress transcription. Mol Cell Biol 2006; 26: 169-81.

77. Lewis JD, Meehan RR, Henzel WJ, et al. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 1992; 69: 905-14.

78. Nan X, Meehan RR, Bird A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Res 1993; 21: 4886-92.

79. Nan X, Ng HH, Johnson CA, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 1998; 393: 386-9.

80. Bird AP, Wolffe AP. Methylation-induced repression--belts, braces, and chromatin. Cell 1999; 99: 451-4.

81. Guy J, Hendrich B, Holmes M, Martin JE, Bird A. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet 2001; 27: 322-6.

82. Stancheva I, Collins AL, Van den Veyver IB, Zoghbi H, Meehan RR. A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. Mol Cell 2003; 12: 425-35.

83. Zhao X, Ueba T, Christie BR, et al. Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. Proc Natl Acad Sci U S A 2003; 100: 6777-82.

84. Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A. Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev 2001; 15: 710-23.

85. Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 1998; 18: 6538-47.

86. Ahringer J. NuRD and SIN3 histone deacetylase complexes in development. Trends Genet 2000; 16: 351-6.

87. Sansom OJ, Berger J, Bishop SM, Hendrich B, Bird A, Clarke AR. Deficiency of Mbd2 suppresses intestinal tumorigenesis. Nat Genet 2003; 34: 145-7.

88. Covic M, Karaca E, Lie DC. Epigenetic regulation of neurogenesis in the adult hippocampus. Heredity; 105: 122-34.

89. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis; 31: 27-36.

90. Cantor JP, Iliopoulos D, Rao AS, et al. Epigenetic modulation of endogenous tumor suppressor expression in lung cancer xenografts suppresses tumorigenicity. Int J Cancer 2007; 120: 24-31.

91. Li LH, Olin EJ, Buskirk HH, Reineke LM. Cytotoxicity and mode of action of 5-azacytidine on L1210 leukemia. Cancer Res 1970; 30: 2760-9.

92. Goffin J, Eisenhauer E. DNA methyltransferase inhibitors-state of the art. Ann Oncol 2002; 13: 1699-716.

93. Lubbert M. DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: clinical results and possible mechanisms of action. Curr Top Microbiol Immunol 2000; 249: 135-64.

94. Ferguson AT, Vertino PM, Spitzner JR, Baylin SB, Muller MT, Davidson NE. Role of estrogen receptor gene demethylation and DNA methyltransferase.DNA adduct formation in 5-aza-2'deoxycytidine-induced cytotoxicity in human breast cancer cells. J Biol Chem 1997; 272: 32260-6.

95. Juttermann R, Li E, Jaenisch R. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. Proc Natl Acad Sci U S A 1994; 91: 11797-801.

96. Daskalakis M, Blagitko-Dorfs N, Hackanson B. Decitabine. Recent Results Cancer Res; 184: 131-57.

97. Yoshida M, Horinouchi S. Trichostatin and leptomycin. Inhibition of histone deacetylation and signal-dependent nuclear export. Ann N Y Acad Sci 1999; 886: 23-36.

98. Jung M. Inhibitors of histone deacetylase as new anticancer agents. Curr Med Chem 2001; 8: 1505-11.

99. Marks PA, Richon VM, Breslow R, Rifkind RA. Histone deacetylase inhibitors as new cancer drugs. Curr Opin Oncol 2001; 13: 477-83.

100. Kouraklis G, Theocharis S. Histone deacetylase inhibitors: a novel target of anticancer therapy (review). Oncol Rep 2006; 15: 489-94.

101. Kim MS, Kwon HJ, Lee YM, et al. Histone deacetylases induce angiogenesis by negative regulation of tumor suppressor genes. Nat Med 2001; 7: 437-43.

102. Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nat Genet 1999; 21: 103-7.

103. Hatsukami DK, Stead LF, Gupta PC. Tobacco addiction. Lancet 2008; 371: 2027-38.

104. Li C, Schuetz JD, Naren AP. Tobacco carcinogen NNK transporter MRP2 regulates CFTR function in lung epithelia: Implications for lung cancer. Cancer Lett.

105. Besaratinia A, Pfeifer GP. Second-hand smoke and human lung cancer. Lancet Oncol 2008; 9: 657-66.

106. Hecht SS. Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention. Lancet Oncol 2002; 3: 461-9.

107. Zheng Y, Ritzenthaler JD, Roman J, Han S. Nicotine stimulates human lung cancer cell growth by inducing fibronectin expression. Am J Respir Cell Mol Biol 2007; 37: 681-90.

108. Sun X, Ritzenthaler JD, Zhong X, Zheng Y, Roman J, Han S. Nicotine stimulates PPARbeta/delta expression in human lung carcinoma cells through activation of PI3K/mTOR and suppression of AP-2alpha. Cancer Res 2009; 69: 6445-53.

109. Jain A. Treating nicotine addiction. BMJ 2003; 327: 1394-5.

110. Hecht SS. Tobacco smoke carcinogens and lung cancer. J Natl Cancer Inst 1999; 91: 1194-210.

111. Shields PG. Molecular epidemiology of smoking and lung cancer. Oncogene 2002; 21: 6870-6.

112. Hecht SS, Hoffmann D. Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. Carcinogenesis 1988; 9: 875-84.

113. Hoffmann D, Brunnemann KD, Prokopczyk B, Djordjevic MV. Tobacco-specific N-nitrosamines and Areca-derived N-nitrosamines: chemistry, biochemistry, carcinogenicity, and relevance to humans. J Toxicol Environ Health 1994; 41: 1-52.

114. Kim DH, Nelson HH, Wiencke JK, et al. p16(INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. Cancer Res 2001; 61: 3419-24.

115. Staretz ME, Murphy SE, Patten CJ, et al. Comparative metabolism of the tobacco-related carcinogens benzo[a]pyrene, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol, and N'- nitrosonornicotine in human hepatic microsomes. Drug Metab Dispos 1997; 25: 154-62.

116. Planchard D, Loriot Y, Goubar A, Commo F, Soria JC. Differential expression of biomarkers in men and women. Semin Oncol 2009; 36: 553-65.

117. Rubins JB, Ewing SL, Leroy S, Humphrey EW, Morrison V. Temporal trends in survival after surgical resection of localized non-small cell lung cancer. Lung Cancer 2000; 28: 21-7.

118. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 2008; 83: 584-94.

119. Anglim PP, Alonzo TA, Laird-Offringa IA. DNA methylation-based biomarkers for early detection of non-small cell lung cancer: an update. Mol Cancer 2008; 7: 81.

120. Heller G, Zielinski CC, Zochbauer-Muller S. Lung cancer: From single-gene methylation to methylome profiling. Cancer Metastasis Rev.

121. Lacroix L, Commo F, Soria JC. Gene expression profiling of non-small-cell lung cancer. Expert Rev Mol Diagn 2008; 8: 167-78.

122. Mao L, Lee DJ, Tockman MS, Erozan YS, Askin F, Sidransky D. Microsatellite alterations as clonal markers for the detection of human cancer. Proc Natl Acad Sci U S A 1994; 91: 9871-5.

123. Slebos RJ, Kibbelaar RE, Dalesio O, et al. K-ras oncogene activation as a prognostic marker in adenocarcinoma of the lung. N Engl J Med 1990; 323: 561-5.

124. Heller G, Zielinski CC, Zochbauer-Muller S. Lung cancer: from single-gene methylation to methylome profiling. Cancer Metastasis Rev; 29: 95-107.

125. Watanabe Y, Maekawa M. Methylation of DNA in cancer. Adv Clin Chem; 52: 145-67.

126. Kersting M, Friedl C, Kraus A, Behn M, Pankow W, Schuermann M. Differential frequencies of p16(INK4a) promoter hypermethylation, p53 mutation, and K-ras mutation in exfoliative material mark the development of lung cancer in symptomatic chronic smokers. J Clin Oncol 2000; 18: 3221-9.

127. Dammann R, Li C, Yoon JH, Chin PL, Bates S, Pfeifer GP. Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. Nat Genet 2000; 25: 315-9.

128. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. Cancer Res 2001; 61: 3225-9.

129. Zochbauer-Muller S, Fong KM, Virmani AK, Geradts J, Gazdar AF, Minna JD. Aberrant promoter methylation of multiple genes in non-small cell lung cancers. Cancer Res 2001; 61: 249-55.

130. Virmani AK, Rathi A, Sathyanarayana UG, et al. Aberrant methylation of the adenomatous polyposis coli (APC) gene promoter 1A in breast and lung carcinomas. Clin Cancer Res 2001; 7: 1998-2004.

131. Virmani AK, Rathi A, Zochbauer-Muller S, et al. Promoter methylation and silencing of the retinoic acid receptor-beta gene in lung carcinomas. J Natl Cancer Inst 2000; 92: 1303-7.

132. Toyooka KO, Toyooka S, Virmani AK, et al. Loss of expression and aberrant methylation of the CDH13 (H-cadherin) gene in breast and lung carcinomas. Cancer Res 2001; 61: 4556-60.

133. Toyooka S, Toyooka KO, Miyajima K, et al. Epigenetic down-regulation of death-associated protein kinase in lung cancers. Clin Cancer Res 2003; 9: 3034-41.

134. Belinsky SA, Grimes MJ, Casas E, et al. Predicting gene promoter methylation in non-small-cell lung cancer by evaluating sputum and serum. Br J Cancer 2007; 96: 1278-83.

135. Machida EO, Brock MV, Hooker CM, et al. Hypermethylation of ASC/TMS1 is a sputum marker for late-stage lung cancer. Cancer Res 2006; 66: 6210-8.

136. Singer-Sam J, LeBon JM, Tanguay RL, Riggs AD. A quantitative HpaII-PCR assay to measure methylation of DNA from a small number of cells. Nucleic Acids Res 1990; 18: 687.

137. Kristensen LS, Hansen LL. PCR-based methods for detecting single-locus DNA methylation biomarkers in cancer diagnostics, prognostics, and response to treatment. Clin Chem 2009; 55: 1471-83.

138. Hayatsu H, Wataya Y, Kai K, Iida S. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. Biochemistry 1970; 9: 2858-65.

139. Taylor KH, Kramer RS, Davis JW, et al. Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. Cancer Res 2007; 67: 8511-8.

140. Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. Nucleic Acids Res 1994; 22: 2990-7.

141. Hayatsu H, Negishi K, Wataya Y. Progress in the bisulfite modification of nucleic acids. Nucleic Acids Symp Ser (Oxf) 2009: 217-8.

142. Schweizer MP, Witkowski JT, Robins RK. Nuclear magnetic resonance determination of syn and anti conformations in pyrimidine nucleosides. J Am Chem Soc 1971; 93: 277-9.

143. Shapiro R, Cohen BI, Servis RE. Specific deamination of RNA by sodium bisulphite. Nature 1970; 227: 1047-8.

144. Azhikina TL, Sverdlov ED. Study of tissue-specific CpG methylation of DNA in extended genomic loci. Biochemistry (Mosc) 2005; 70: 596-603.

145. Sadri R, Hornsby PJ. Rapid analysis of DNA methylation using new restriction enzyme sites created by bisulfite modification. Nucleic Acids Res 1996; 24: 5058-9.

146. Oakeley EJ. DNA methylation analysis: a review of current methodologies. Pharmacol Ther 1999; 84: 389-400.

147. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 1997; 25: 2532-4.

148. Rychlik W. Selection of primers for polymerase chain reaction. Mol Biotechnol 1995; 3: 129-34.

149. Eads CA, Danenberg KD, Kawakami K, et al. MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 2000; 28: E32.

150. Dai Z, Lakshmanan RR, Zhu WG, et al. Global methylation profiling of lung cancer identifies novel methylated genes. Neoplasia 2001; 3: 314-23.

151. Meng CF, Zhu XJ, Peng G, Dai DQ. Re-expression of methylation-induced tumor suppressor gene silencing is associated with the state of histone modification in gastric cancer cell lines. World J Gastroenterol 2007; 13: 6166-71.

152. Cecconi D, Donadelli M, Dalla Pozza E, et al. Synergistic effect of trichostatin A and 5-aza-2'-deoxycytidine on growth inhibition of pancreatic endocrine tumour cell lines: a proteomic study. Proteomics 2009; 9: 1952-66.

153. Heller G, Schmidt WM, Ziegler B, et al. Genome-wide transcriptional response to 5-aza-2'-deoxycytidine and trichostatin a in multiple myeloma cells. Cancer Res 2008; 68: 44-54.

154. Shames DS, Girard L, Gao B, et al. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. PLoS Med 2006; 3: e486.

155. Weber M, Davies JJ, Wittig D, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 2005; 37: 853-62.

156. Guil S, Esteller M. DNA methylomes, histone codes and miRNAs: tying it all together. Int J Biochem Cell Biol 2009; 41: 87-95.

157. Lamy A, Sesboue R, Bourguignon J, et al. Aberrant methylation of the CDKN2a/p16INK4a gene promoter region in preinvasive bronchial lesions: a prospective study in high-risk patients without invasive cancer. Int J Cancer 2002; 100: 189-93.

158. Palmisano WA, Divine KK, Saccomanno G, et al. Predicting lung cancer by detecting aberrant promoter methylation in sputum. Cancer Res 2000; 60: 5954-8.

159. Toyooka S, Toyooka KO, Maruyama R, et al. DNA methylation profiles of lung tumors. Mol Cancer Ther 2001; 1: 61-7.

160. Buckingham L, Penfield Faber L, Kim A, et al. PTEN, RASSF1 and DAPK site-specific hypermethylation and outcome in surgically treated stage I and II nonsmall cell lung cancer patients. Int J Cancer; 126: 1630-9.

161. Jin M, Kawakami K, Fukui Y, et al. Different histological types of non-small cell lung cancer have distinct folate and DNA methylation levels. Cancer Sci 2009; 100: 2325-30.

162. Belinsky SA, Palmisano WA, Gilliland FD, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res 2002; 62: 2370-7.

163. 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: 9005-14.

164. Tang X, Khuri FR, Lee JJ, et al. Hypermethylation of the death-associated protein (DAP) kinase promoter and aggressiveness in stage I non-small-cell lung cancer. J Natl Cancer Inst 2000; 92: 1511-6.

165. Lu C, Soria JC, Tang X, et al. Prognostic factors in resected stage I non-small-cell lung cancer: a multivariate analysis of six molecular markers. J Clin Oncol 2004; 22: 4575-83.

166. Yanagawa N, Tamura G, Oizumi H, et al. Promoter hypermethylation of RASSF1A and RUNX3 genes as an independent prognostic prediction marker in surgically resected non-small cell lung cancers. Lung Cancer 2007; 58: 131-8.

167. De Schutter H, Geeraerts H, Verbeken E, Nuyts S. Promoter methylation of TIMP3 and CDH1 predicts better outcome in head and neck squamous cell carcinoma treated by radiotherapy only. Oncol Rep 2009; 21: 507-13.

168. Brock MV, Hooker CM, Ota-Machida E, et al. DNA methylation markers and early recurrence in stage I lung cancer. N Engl J Med 2008; 358: 1118-28.

169. Kaminskas E, Farrell A, Abraham S, et al. Approval summary: azacitidine for treatment of myelodysplastic syndrome subtypes. Clin Cancer Res 2005; 11: 3604-8.

170. Juergens R, Vendetti, F., Coleman, B., Sebree, R., Belinsky, S., Rudek M, et al. A phase II-trial of 5-azacitidine (5AC) and entinostat (SNDX-275) in relapsed advanced lung cancer (NSCLC): an interim analysis. Abstract# A66, 13th World Conference on Lung Cancer (WCLC) 2009.

171. Bauer S, Grossmann S, Vingron M, Robinson PN. Ontologizer 2.0--a multifunctional tool for GO term enrichment analysis and data exploration. Bioinformatics 2008; 24: 1650-1.

172. Yang X, Phillips DL, Ferguson AT, Nelson WG, Herman JG, Davidson NE. Synergistic activation of functional estrogen receptor (ER)-alpha by DNA methyltransferase and histone deacetylase inhibition in human ER-alpha-negative breast cancer cells. Cancer Res 2001; 61: 7025-9.

173. Virmani AK, Gazdar AF. Tumor suppressor genes in lung cancer. Methods Mol Biol 2003; 222: 97-115.

174. Rauch TA, Zhong X, Wu X, et al. High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci U S A 2008; 105: 252-7.

175. Bunt J, de Haas TG, Hasselt NE, et al. Regulation of Cell Cycle Genes and Induction of Senescence by Overexpression of OTX2 in Medulloblastoma Cell Lines. Mol Cancer Res; 8: 1344-57.

176. Bestor TH. Unanswered questions about the role of promoter methylation in carcinogenesis. Ann N Y Acad Sci 2003; 983: 22-7.

177. Sato N, Fukushima N, Maitra A, et al. Discovery of novel targets for aberrant methylation in pancreatic carcinoma using high-throughput microarrays. Cancer Res 2003; 63: 3735-42.

178. Zhong S, Fields CR, Su N, Pan YX, Robertson KD. Pharmacologic inhibition of epigenetic modifications, coupled with gene expression profiling, reveals novel targets of aberrant DNA methylation and histone deacetylation in lung cancer. Oncogene 2007; 26: 2621-34.

179. Yoshino M, Suzuki M, Tian L, et al. Promoter hypermethylation of the p16 and Wif-1 genes as an independent prognostic marker in stage IA non-small cell lung cancers. Int J Oncol 2009; 35: 1201-9.

180. Lin Q, Geng J, Ma K, et al. RASSF1A, APC, ESR1, ABCB1 and HOXC9, but not p16INK4A, DAPK1, PTEN and MT1G genes were frequently methylated in the stage I non-small cell lung cancer in China. J Cancer Res Clin Oncol 2009; 135: 1675-84.

181. Yamashita S, Tsujino Y, Moriguchi K, Tatematsu M, Ushijima T. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. Cancer Sci 2006; 97: 64-71.

182. Oka D, Yamashita S, Tomioka T, et al. The presence of aberrant DNA methylation in noncancerous esophageal mucosae in association with smoking history: a target for risk diagnosis and prevention of esophageal cancers. Cancer 2009; 115: 3412-26.

183. Tomida S, Yanagisawa K, Koshikawa K, et al. Identification of a metastasis signature and the DLX4 homeobox protein as a regulator of metastasis by combined transcriptome approach. Oncogene 2007; 26: 4600-8.

184. Ogba N, Doughman YQ, Chaplin LJ, et al. HEXIM1 modulates vascular endothelial growth factor expression and function in breast epithelial cells and mammary gland. Oncogene; 29: 3639-49.

185. Peedicayil A, Vierkant RA, Hartmann LC, et al. Risk of ovarian cancer and inherited variants in relapse-associated genes. PLoS One; 5: e8884.

186. Hong J, Resnick M, Behar J, et al. Acid-induced p16 hypermethylation contributes to development of esophageal adenocarcinoma via activation of NADPH oxidase NOX5-S. Am J Physiol Gastrointest Liver Physiol; 299: G697-706.

187. Flynn RL, Zou L. ATR: a master conductor of cellular responses to DNA replication stress. Trends Biochem Sci.

188. Okumura K, Mendoza M, Bachoo RM, DePinho RA, Cavenee WK, Furnari FB. PCAF modulates PTEN activity. J Biol Chem 2006; 281: 26562-8.

189. Chang HC, Cho CY, Hung WC. Downregulation of RECK by promoter methylation correlates with lymph node metastasis in non-small cell lung cancer. Cancer Sci 2007; 98: 169-73.

190. Du YY, Dai DQ, Yang Z. Role of RECK methylation in gastric cancer and its clinical significance. World J Gastroenterol; 16: 904-8.

191. Sasaki H, Ishihara K, Kato R. Mechanisms of Igf2/H19 imprinting: DNA methylation, chromatin and long-distance gene regulation. J Biochem 2000; 127: 711-5.

192. Wrzeska M, Rejduch B. Genomic imprinting in mammals. J Appl Genet 2004; 45: 427-33.

193. Waha A, Guntner S, Huang TH, et al. Epigenetic silencing of the protocadherin family member PCDH-gamma-A11 in astrocytomas. Neoplasia 2005; 7: 193-9.

194. Novak P, Jensen T, Oshiro MM, Watts GS, Kim CJ, Futscher BW. Agglomerative epigenetic aberrations are a common event in human breast cancer. Cancer Res 2008; 68: 8616-25.

195. Miyamoto K, Fukutomi T, Akashi-Tanaka S, et al. Identification of 20 genes aberrantly methylated in human breast cancers. Int J Cancer 2005; 116: 407-14.

196. Dallosso AR, Hancock AL, Szemes M, et al. Frequent long-range epigenetic silencing of protocadherin gene clusters on chromosome 5q31 in Wilms' tumor. PLoS Genet 2009; 5: e1000745.

197. Shah N, Sukumar S. The Hox genes and their roles in oncogenesis. Nat Rev Cancer; 10: 361-71.

198. Rauch T, Wang Z, Zhang X, et al. Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay. Proc Natl Acad Sci U S A 2007; 104: 5527-32.

199. Shiraishi M, Sekiguchi A, Oates AJ, Terry MJ, Miyamoto Y. HOX gene clusters are hotspots of de novo methylation in CpG islands of human lung adenocarcinomas. Oncogene 2002; 21: 3659-62.

200. Wilzen A, Nilsson S, Sjoberg RM, Kogner P, Martinsson T, Abel F. The Phox2 pathway is differentially expressed in neuroblastoma tumors, but no mutations were found in the candidate tumor suppressor gene PHOX2A. Int J Oncol 2009; 34: 697-705.

201. Serra A, Haberle B, Konig IR, et al. Rare occurrence of PHOX2b mutations in sporadic neuroblastomas. J Pediatr Hematol Oncol 2008; 30: 728-32.

202. Reggiani L, Raciti D, Airik R, Kispert A, Brandli AW. The prepattern transcription factor Irx3 directs nephron segment identity. Genes Dev 2007; 21: 2358-70.

203. Kadota M, Sato M, Duncan B, et al. Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. Cancer Res 2009; 69: 7357-65.

10 ABSTRACT

Lung cancer is the number one of death reasons among cancer patients causing ~ one million deaths each year worldwide. Adenocarcinomas represent the predominant histological subtype of non-small cell lung cancers (NSCLC). Beside genetic aberrations, epigenetic changes lead to de-regulation of gene expression during the pathogenesis of lung cancer. The so far best studied epigenetic change in NSCLCs is DNA methylation. Several genes, including *RASSF1A* or *DAPK* have already been identified to be methylated in NSCLC. To gain further insight into DNA methylation changes in NSCLCs we performed two different approaches for whole genome DNA methylation analysis in three lung adenocarcinoma cell lines. The first approach is based on the observation that epigenetic changes are reversible. Thus, cell lines were treated with the epigenetically active drugs 5-aza-2'-deoxycytidine (Aza-dC), a DNA methyltransferase inhibitor, and trichostatin A (TSA), a histone deacetylase inhibitor. Since synergistic effects of these two drugs are known cells were also treated with a combination of Aza-dC and TSA. Consequently, gene expression patterns of treated and untreated cells were compared by Affymetrix microarray analyses. Using this approach, we identified 4560 uniquely up-regulated genes after Aza-dC/TSA treatment. Since, Aza-dC and TSA are cytotoxic drugs up-regulation of certain genes may be due to secondary effects. To address this issue, we established an approach for direct identification of methylated DNA fragments: methylated DNA immunoprecipitation (MeDIP) combined with microarray analysis (MeDIP-chip). With this method methylation of ~28.000 CpG islands was analyzed in each of the 3 lung adenocarcinoma cell lines and in normal human bronchial epithelial cells (NHBEC). By comparing microarray data obtained from the adenocarcinoma cell lines and NHBE cells we identified 47 tumor-specifically methylated genes. Genes which were so far unknown to be de-regulated in lung cancer and which are affiliated to cancer associated pathways like cell differentiation, apoptosis or cell migration were identified (e.g. *POU4F2*, *PAX6* and *HOXA9)*. 21 out of the 47 genes identified were also found to be up-regulated after treatment with epigenetically active drugs.

Results of the two whole genome level approaches were verified using methylation-specific PCR (MSP). MSP analyses in NSCLCs and corresponding non-malignant lung tissue

samples obtained from 10 NSCLC patients verified the tumor-specific methylation of the genes identified.

In conclusion, using genome-wide approaches for DNA methylation analysis we identified a large number of so far unknown methylated genes in NSCLC.

11 ZUSAMMENFASSUNG

Um ein besseres Verständnis der Pathogenese des nicht-kleinzelligem Lungenkarzinoms zu erlangen, haben wir die Bedeutung der DNA Methylierung in NSCLC Zelllinien untersucht. Drei NSCLC Zelllinien wurden mit 5-aza-2'-deoxycytidine (Aza-dC) und/oder Trichostatin A (TSA) behandelt und das Geneexpressionsmuster von behandelten mit unbehandelten Zellen verglichen. Mit dieser Methode kann DNA Methylierung indirekt über die Analyse der differentiellen Gene Expression identifiziert werden. Wir haben 4560 Gene wurden identifiziert, deren Expression nach Behandlung mit den epigenetisch aktiven Agenzien in mindestens einer Zelllinie verstärkt war. 30% der verstärkt exprimierten Gene nach Behandlung mit Aza-dC wurden synergistisch hinauf reguliert nach Behandlung mit Aza-dC und TSA. Weiters wurden die drei NSCLC Zelllinien sowie „normal human bronchial epithelien cells“ (NHBEC) mit einer weiteren Genom-weiten DNA Methylierungsanalyse untersucht. Die Methode MeDIP-Chip wurde in den letzten Jahren entwickelt und kann DNA Methylierung direkt identifizieren. Basierend auf Immunopräzipitation werden in dieser Technik methylierte DNA Fragmente angereichert und mittels Microarray mit der Ausgangsprobe verglichen. Wir identifizierten 47 tumorspezifisch methylierte Gene. 45% dieser Gene waren auch in den Ergebnissen der ersten Methode präsent. Identifizierte Gene wurden durch eine genespezifische Methode, nämlich Methylierungs-spezifische PCR (MSP), bestätigt. Unter den identifizierten Genen befinden sich bereits bekannte Targetgene für Methylierung in NSCLCs sowie Gene, deren Methylierung bis dato unbekannt war. Eine Vielzahl der identifizierten Gene hat Einfluss auf Krebsassoziierte Signaltransduktionswege. Ein weiteres Resultat dieser Arbeit ist, dass neben der Methylierung einzelner Genen auch DNA Methylierung ganzer Gencluster wie HOXA und PCDHA identifiziert wurden.

Insgesamt ist es uns gelungen, eine Vielzahl bis dato unbekannter Targetgene für DNA Methylierung zu identifizieren und somit, die Wichtigkeit der DNA Methylierung in der Pathogenese von NSCLC hervor zu heben.

12 CURRICULUM VITAE

Valérie Nadine Babinsky

Born on October 16th, 1985 in Vienna, resident of Vienna

Dual Citizenship of Austria and Belgium

Family

Mother: Chantal Babinsky-Chatain; Translator

Father: Dr. Heinrich Babinsky; Communication and Event-Management

Siblings: Mag. Fabienne Babinsky, Nathalie Babinsky

Education

Sep 06-Mar 11	Specialization in Microbiology & Genetics **Major:** Cyto- and Development Genetics Genetics and Pathology
Jun 2006	First Diploma in Biology **University of Vienna, Grade: Very Good**
Sep 04 – Jun 06	Master Studies Biology **University of Vienna**
1996-2004	High School **Goethe Gymnasium Vienna 14**
1992-1996	Primary School **Felbigergasse Vienna 14**

Career

Dec 2009	Publication of study progress *Title: "DNA-Methylierung und Bronchialkarzinome"* **JATROS Hämatologie & Onkologie 7/2009**
Oct 2009	Presentation as part of the session "Best Abstracts" *Abstract: V582 Identification of unknown epigenetically regulated genes in non-small cell lung cancer* **Gemeinsame Jahrestagung '09 der Deutschen, Österreichischen und Schweizerischen Gesellschaften für Hämatologie und Onkologie**
Mar 09-Feb 10	Diploma-thesis on Epigenetics *Institute for Oncology* **General Hospital of Vienna**
Aug 2009	Internship on cellular quality control *Supervised by Dr. Erwin Ivessa* **Max F. Perutz Laboratories**
Jun 2008	Election for a tutor's position *Institute for organic chemistry* **University of Vienna**

Language skills

German, first mother tongue

French, second mother tongue

English, very good oral, writing and presentation skills

Additional skills

MS Office, Graphic design programs (Paint, Adobe Photoshop)

Drivers license class B since 2004

Valerie Babinsky

Stauffergasse 43, A-1140 Vienna, 0043/699/10181218
valerie@babinsky.info