Tumor methylation markers and clinical outcome of primary oral squamous cell carcinomas

Exploring the OSCC Methylome

Martijn Jacobus Antonius Maria Clausen

rijksuniversiteit
groningen

Tumor methylation markers and clinical outcome of primary oral squamous cell carcinomas

Exploring the OSCC Methylome

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 5 februari 2020 om 14:30 door

Martijn Jacobus Antonius Maria Clausen

geboren op 10 maart 1986

te Steenwijk

Promotores

Prof. dr. E. Schuuring
Prof. dr. J.L.N. Roodenburg

Beoordelingscommissie

Prof. dr. H. Hollema
Prof. dr. M. van Engeland
Prof. dr. A.J.W.P. Rosenberg

Paranimfen

Dr. G. Eising
Dr. T. Haer

TABLE OF CONTENTS

Chapter 1

General introduction

Head and neck cancer is the collective name for cancers that arise in the head-neck (HN) region. These are tumors of the upper aerodigestive tract, of which oral cavity and larynx cancer are the most common, but head and neck cancer also includes tumors in the floor of the mouth, tongue, but also in the vermillion border of the lips, from the salivary glands, thyroid gland, tonsils and nasal cavity and of the soft tissues and bones in this area. The common cancers in the HN region mainly develop from the mucosa, resulting in head and neck squamous cell carcinomas (HNSCC). Other head-neck cancer types are adenocarcinomas, sarcoma's, melanomas, squamous cell carcinomas of the skin and lymphomas that are more rare [1]. Complex skin cancers are often also treated in a multidisciplinary head and neck cancer center. The relative high frequency of HNSCC, compared to the other head and neck tumors, is because the epithelial layers of the oral cavity, pharynx and larynx are exposed to the same risk factors. The common risk factors are life style-related such as tobacco smoking especially when combined with alcohol use and dietary factors [2]. The continuous stress of the epithelial layers caused by exposure to tobacco smoke and alcohol causes wide-spread accumulation of genetic aberrations in the upper aerodigestive tract [3]. Other causes of HNSCC include viral infections of the Human papillomavirus in tonsillar carcinoma [4], the Epstein-Barr virus in nasal carcinoma or sunlight exposure in particular lip and skin cancers.

 An important part HNSCC are those located in the oral cavity [5]. Worldwide, an estimated 300,400 new cases of cancer of the oral cavity, including lip and predominantly oral squamous cell carcinomas (OSCC) were diagnosed in 2012 and 145,400 deaths associated with these cancers in 2012 [5]. Furthermore, the incidence of these cancers has also been steadily increasing over the last few years. E.g. in 2012 in USA the estimated new cases of cancer in the oral cavity or pharynx (OPSCC) was 40,250 and 7,850 deaths [5]. In the 2016 these numbers had increased to 48,330 new cases and 9,570 deaths. Because, both the overall new cases of all cancers and all cancer-related deaths increased with 3% from 2012 to 2016, the increase of 20% of new cases with OPSCC and 22% of deaths from 2012 to 2016 specifically, is highly significant [6], [7].

While the incidence of OSCC in the Netherlands is rather low, it has doubled over the last 25 years from 507 cases in 1990 to 906 in 2015 (www.cijfersoverkanker.nl, visited on 02-03-2019). Meanwhile, the 5-year-survival rate was 57% for OSCC diagnosed between 1991-1995 and only improved to a 62% 5-year survival for OSCC diagnosed between 2011-2015. Due to this stagnant 5-year survival combined with the increasing incidence, OSCC poses a big clinical challenge.

The biological behavior of OSCC is local destruction of tissue, anatomy and organs and regional metastases to the lymph nodes in the neck. Distant metastasis of OSCC is rather rare at the moment of diagnosis and has been report to occur in 6-12% of patients [8]–[11]. When distant metastases do occur, it is in a late phase and are mostly located in the lungs.

The treatment modalities for locoregional treatment are primary surgery and occasionally primary radiotherapy for the low stages of OSCC. High stages OSCC require combined treatment with both surgery as well as postoperative radiotherapy. High stage carcinomas that cannot be surgically removed, are solely treated with radiotherapy. Progress is made in combining radiotherapy with systemic treatment like chemotherapy or with biologicals in high risk cases. All mentioned treatments have serious side

effects, so a tailormade treatment plan, based on expected behavior of the tumor is mandatory for good survival and quality of life.

A dilemma in the treatment of OSCC is the management of the "clinical negative neck". This term refers to a common dilemma where there is no evidence for lymph node metastases based on clinical assessment and imaging, while due to tumor factors there is a substantial risk for the presence of microscopic metastases. The presence of lymph node metastasis was reported to reduce the five-year disease-free survival in patients with OSCC from 72 to 9% and had the highest hazard ratio of all reported clinical pathological predictors [12].

To assist in the detection of nodal spread in OSCC, specific clinicopathological traits that are associated with lymph node metastasis are included in risk assessment for LN metastasis. Certain tumor characteristics such as size, invasive behavior and the pattern of invasion are indicators of the tumor clinical behavior, and associated as well as predictive of lymph node metastasis [13]–[17]. Especially infiltration depth, perineural and lymphovascular invasion as well as histological differentiation have been found to be good clinicopathological predictors for cervical LN metastases in the neck [18]. These findings led to an important revision of the 8th edition of TNM staging guideline which includes now infiltration depth as part of the T-status in addition to conventional tumor diameter which was decisive in earlier version of the TNM staging [19], [20].

For patients with OSCC and a clinically negative neck (cN0) and a small tumor (T1-T2) [20], elective neck dissection is recommended [21]. In the past patients with cN0/T1-T2 were treated with a modified radical neck dissection, removing lymph node levels I to IV (Figure 1.1). Later these elective radical neck dissections were replaced by a "selective neck dissection" which removed a limited set of lymph nodes, level I to III (Figure 1.1). In about 70% of the cT1-T2/N0 OSCC cases, a neck dissection can be avoided with a "watchful waiting policy", which refers to an intensive follow up regime. The avoidance of a neck dissection is especially relevant since this surgical procedure, even a selective neck dissection, is associated with complications such as loss of shoulder function, edema and increase of costs [22]. Additionally, for the 30% of the cN0 OSCC patients who suffer a conversion to a regional recurrence, the watchful waiting ensures a proper follow-up to diagnose and treat late occurring LN metastases as soon as possible [23], [24]. A recently development in N-status determination of OSCC is the Sentinel Lymph Node Biopsy (SLNB). To perform a SLNB, a radioactive tracer is injected around the primary tumor which can then visualize the lymphatic drainage pattern as well the primary draining sentinel lymph node from the tumor location with a SPECT scan. During surgery the sentinel lymph node can be detected with a probe and is removed for histological assessment. Additionally, after resection of the sentinel lymph node, micro metastases in this particular lymph node can be identified using immunohistochemistry. SNLB is an important contribution to the assessment of the neck in OSCC. The negative predictive value (NPV) of SLNB in OSCC for nodal spread has been reported to be between 88 to 95% [25], [26]. The SLNB has a lower sensitivity and NPV in floor of mouth tumors for detecting occult metastasis due to "shine through phenomenon". This phenomenon is caused by the high proximity of the injection site of the tracer in the primary tumor and the lymph nodes in level I-A. As a result of the limited distance to the primary tumor, the high radioactivity of the site of tracer injection and the sentinel node cannot be properly distinguished [25], [27]. A drawback

of SLNB is the increased morbidity in case of a positive sentinel node [28], [29]. Patients with a positive sentinel lymph node are treated with a neck dissection were level I to V are removed. After a SLNB the neck dissection is more difficult due to residual wounds or scar tissue. Although SLNB is a good and less invasive procedure than an elective neck dissection, the current sensitivity of SLNB to detect occult metastases does not fully solve the dilemma in diagnosing nodal spread [30].
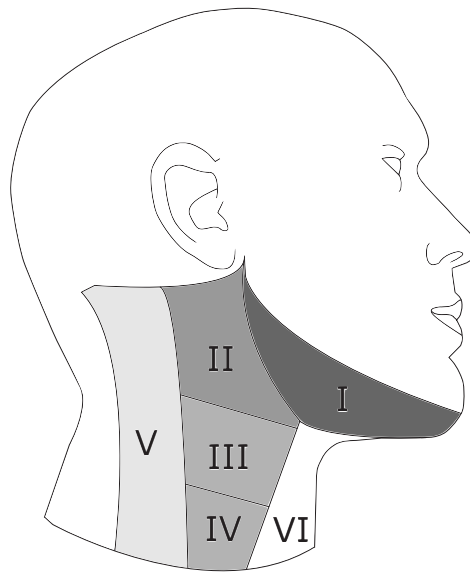
Figure 1.1. The six levels of lymph nodes in the lymphatic drainage patterns of OSCC tumors. Edited from [31] with permission.

Until now only tumor infiltration depth is helpful in predicting micro metastases in the neck of OSCC. The statistical validity of other clinical predictors is often insufficiently tested and subjected to a high degree of observer-bias [32]–[35]. Nevertheless, these clinicopathological features are often used as the decisive factor in the treatment strategy [36]. Therefore, there is a need for another approach to solve the dilemma around the negative clinical neck.

Molecular Tumor biomarkers predictive for N-status on Oral Squamous Cell Carcinomas

While great progress has been made in both the prediction of subclinical metastases in the neck of OSCC patients as well as treatment strategies, there is a group of tumors that behave differently and are not suited for staging and treatment with these methods. These are tumors with initially a low risk, but still develop later cervical metastases of an OSCC. Biomarkers may be helpful in selecting cases of OSCC that have a risk for subclinical metastases in the neck and need treatment for that. Molecular tumor biomarkers have been studied. These biomarkers are indicative of certain biological behaviors. These molecular drivers of particular phenotypes or growth patterns can severely impact the primary tumors

ability to metastasize to the lymph nodes. In the last decades, numerous molecular tumor markers have been identified and evaluated for association with LN status in OSCC (Table 1.1).

Table 1.1. Overview of single molecular biomarkers that have been shown to have aberrantly expressed proteins or mRNA levels in OSCC with LN metastases.

Biomarker	Alteration	Major pathway associated with biomarker	Reference
DF3/MUC1	Over expression	Cell cycle regulation, proliferation, apoptosis.	[37]
ALDH-1	Over expression	Cell cycle regulation, proliferation, apoptosis.	[38]
PTEN	Under expression	Cell cycle regulation, proliferation, apoptosis.	[39]
Bcl2	Over expression	Cell cycle regulation, proliferation, apoptosis.	[39]
Shp2	Over expression	Cell cycle regulation, proliferation, apoptosis.	[40]
MT3	Under expression	Cell cycle regulation, proliferation, apoptosis.	[41]
PDL-1	Under expression	Cell cycle regulation, proliferation, apoptosis.	[42]
ATG16L1	Over expression	Cell cycle regulation, proliferation, apoptosis.	[43]
ABCB5	Over expression	Cell motility, cell adhesion, microenvironment.	[44]
Twist	Over expression	Cell motility, cell adhesion, microenvironment.	[45]
E-cadherin	Under expression	Cell motility, cell adhesion, microenvironment.	[46]
Podolopin	Under expression	Cell motility, cell adhesion, microenvironment.	[46]
VEGF-C	Over expression	Cell motility, cell adhesion, microenvironment.	[47]
ITGA3	Over expression	Cell motility, cell adhesion, microenvironment.	[48]
ITGB4	Over expression	Cell motility, cell adhesion, microenvironment.	[48]
Claudin-7	Under expression	Cell motility, cell adhesion, microenvironment.	[49]
DNp63	Under expression	Cell motility, cell adhesion, microenvironment.	[50]
MMP-11	Over expression	Cell motility, cell adhesion, microenvironment.	[51]
ANO1	Over expression	Cell motility, cell adhesion, microenvironment.	[52]
uPAR PAI-1	Under expression	Cell motility, cell adhesion, microenvironment.	[53]
S100A4	Over expression	Cell motility, cell adhesion, microenvironment.	[54]
COX-2	Over expression	Cell motility, cell adhesion, microenvironment.	[55]
CYFRA 21-1	Over expression	Cell motility, cell adhesion, microenvironment.	[56]
CD68+ TAMs	Over expression	Cell motility, cell adhesion, microenvironment.	[57]
Claudin-1	Over expression	Cell motility, cell adhesion, microenvironment.	[58]
CMTM3	Over expression	Cell motility, cell adhesion, microenvironment.	[59]
E-cadherin	Under expression	Cell motility, cell adhesion, microenvironment.	[60]
EPOR	Over expression	Transcription factors, immune system, angiogenesis.	[61]
NKX3-1	Under expression	Transcription factors, immune system, angiogenesis.	[62]
NNMT	Over expression	Transcription factors, immune system, angiogenesis.	[63]
CNTN1	Over expression	Transcription factors, immune system, angiogenesis.	[64]
KLK13	Under expression	Transcription factors, immune system, angiogenesis.	[65]
TANGO	Over expression	Transcription factors, immune system, angiogenesis.	[66]
CD163	Over expression	Transcription factors, immune system, angiogenesis.	[67]
AEG-1	Over expression	Transcription factors, immune system, angiogenesis.	[68]
Lin28B	Over expression	Transcription factors, immune system, angiogenesis.	[69]
CAIX	Over expression	Transcription factors, immune system, angiogenesis.	[70]
GCS/P-gp	Over expression	Transcription factors, immune system, angiogenesis.	[71]
IL-37	Under expression	Transcription factors, immune system, angiogenesis.	[72]
KiSS-1	Under expression	Transcription factors, immune system, angiogenesis.	[73]
miR-483-5p	Under expression	Transcription factors, immune system, angiogenesis.	[74]

Indicated is also in what major pathway each of these markers have been reported to be involved in. Adapted from [75] with permission.

Changes in proteins levels in HNSCC are often due to underlying genetic alterations such as differences in gene copy numbers. However, different studies using such protein markers are hard to compare due to high variability of different tumor characteristics, the variation in immune-staining methods and scoring systems used to determine protein levels, as well as the subjective interpretation of the observers when assessing the protein expression [76]. Without a golden standard to determine tumor characteristics it is very difficult to directly compare protein and RNA expression between biomarker studies. And so far, the poor predictive value of such markers did not result in the incorporation into clinical modalities yet [77]–[79].

Invasion and metastasis are complicated multistep processes that rely on the deregulation of many different genes and pathways. This might be a reason why it is difficult to identify a single aberrantly expressed protein in the primary tumor that is predictive for tumor metastases [80], [81]. Fortunately, technical advances and novel methods that allow the simultaneous assessment of a high number of genes, have greatly improved biomarker discovery. Moreover, this progress has allowed for more elaborate profiling of tumors, enabling stratification of tumors which have similar histology or staging but are vastly different on a genetic level [82]. In order to investigate whether a gene expression profile could be identified that is predictive for the presence of lymph node metastasis in HNSCC in 2005, Roepman et al. performed a microarray study in HNSCC [83]. The microarray analysis allowed for testing of 21,329 genes simultaneously which led to the assembly of a 102-gene expression profile associated with nodal spread in HNSCC with a negative predictive value of 86%. The gene-panel was subsequently further validated and expanded to a 696-gene expression profile which had a negative predictive value of 89% [78]. This extensive gene panel shows the possible variation in involved genes and the complexity of metastasis as a whole and but also the difficulty of identify a universal metastatic profile in HNSCC. Especially considering similar studies employing even larger microarrays identify markers report on the single gene BMI1 as mostly predictive for HNSCC metastasis [84] while this gene is not included in the 696-gene expression profile reported [78]. And even this extensive study only covers genes, excluding emerging biomarkers like microRNA's such as miR-21, miR-16 and miR-30a-5p that have been associated with nodal metastasis in HNSCC cell lines [85]. To summarize, even the microarray expression studies did not result in the incorporation of biomarkers into clinical modalities. Great variance has been seen in different gene signatures for HNSCC and the application of these panels might not be suitable for clinical application yet [82].

Epigenetics as regulators of gene expression

A relatively recently discovered biological mechanism of gene regulation is epigenetics (Figure 1.2). Epigenetics means "above genetics" and is a collective term for modifications of the DNA other than structural changes in the DNA sequence that do impact gene expression (reviewed in [86]). The phenomenon of epigenetics can be summarized as the "development of phenotypes from genotypes" as defined by Waddington in 1942 [87]. In general, epigenetics consists of modifications of the DNA structure. The DNA double helix is extensively folded and packaged in an array of structures to fit in a human nucleus. Each human diploid cell contains a total of about 2 meters of DNA [88]. To facilitate the

storage of DNA into a human diploid cell, DNA is extensively packed and folded to fit into a cell nucleus. The degree of packaging and the resulting structure is not uniform in the complete genome. The chromosomes, the largest DNA structures in the human cells, for example are known to have structures depending on the degree of packaging referred to as chromatin [89]. Each structural chromatin form has a different density of the structure and a corresponding rate of gene expression as a result of the physical availability of the chromatin for gene expression [90]. Euchromatin refers to the less packed DNA structure and is associated with an increased level of gene expression while heterochromatin is more densely packed and genes in this structure are relatively less expressed. On a lower structural level, DNA is wrapped around nucleosomes (Figure 1.2). These nucleosomes consist of building blocks called histones containing long protein tails. These additional amino acid tails can be further reversibly modified by the addition of molecules such as methyl groups and ubiquitin groups to alter the chromatin structure [90]. Additionally, the chromatin structure is modified by interaction with RNA such as long noncoding RNA [91], [92]. These different modifications result in different structures of the nucleosomes, resulting in different genomic regulation of the DNA bound to these structural proteins. It is general accepted that these various epigenetic mechanisms have major contribution to the regulation of gene expression [89], [90].Through the complex regulation of gene expression through chromatin modification, the same set of genes, or genotype, can accommodate vastly different life stages or phenotype. For example, within bees the genome of the Queen genotype and the worker genotype are complete identical while through epigenetic regulation the phenotypes are vastly different (Figure 1.3)[93].
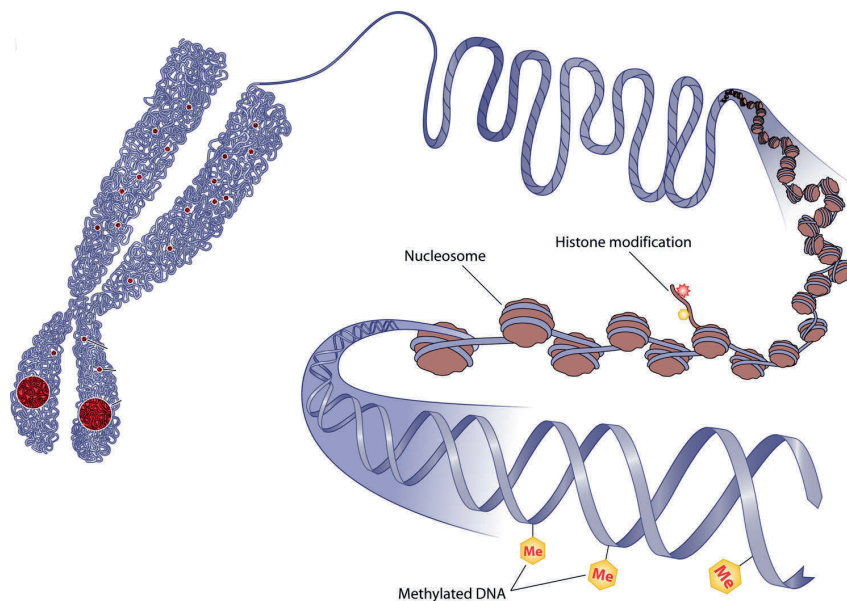
Figure 1.2. Different levels of DNA structures impacting inheritable epigenetics. On the left, the largest DNA structure, the chromosomes are depicted. The density of the structure of the chromosomes impacts gene expression. Next, DNA strands are wrapped around larger proteins, the nucleosomes. These larger structural proteins contain proteins tails that can have different molecular modifications that impact the DNA structure that impact gene expression. Finally, molecular modifications of the nucleotides such as the addition of methyl groups change the DNA structure and impact gene expression as well. Adapted from [86] with permission.

DNA methylation

The smallest building blocks of DNA are the nucleotides adenine, guanine, thymine and cytosine. These bases can be methylated. The addition of a methyl group to single nucleotide can affect all four bases but highly favors cytosines that precede guanine residuals. These preferred sites of methylation are known as CG dinucleotides and often referred to as CpG sites (Figure 1.4). Because the Methyl molecule is always added to the 5' position of a cytosine, a methylated cytosine molecule is referred to as 5'-methyl-cytosine or 5mC. DNA methylation affects the chromatin structure, contributing to gene expression regulation [89], [90], [94]. Additionally, a high percentage of nucleotide changes are caused by the spontaneous hydrolytic deamination of methylated cytosines which results in a thymine [95], [96].
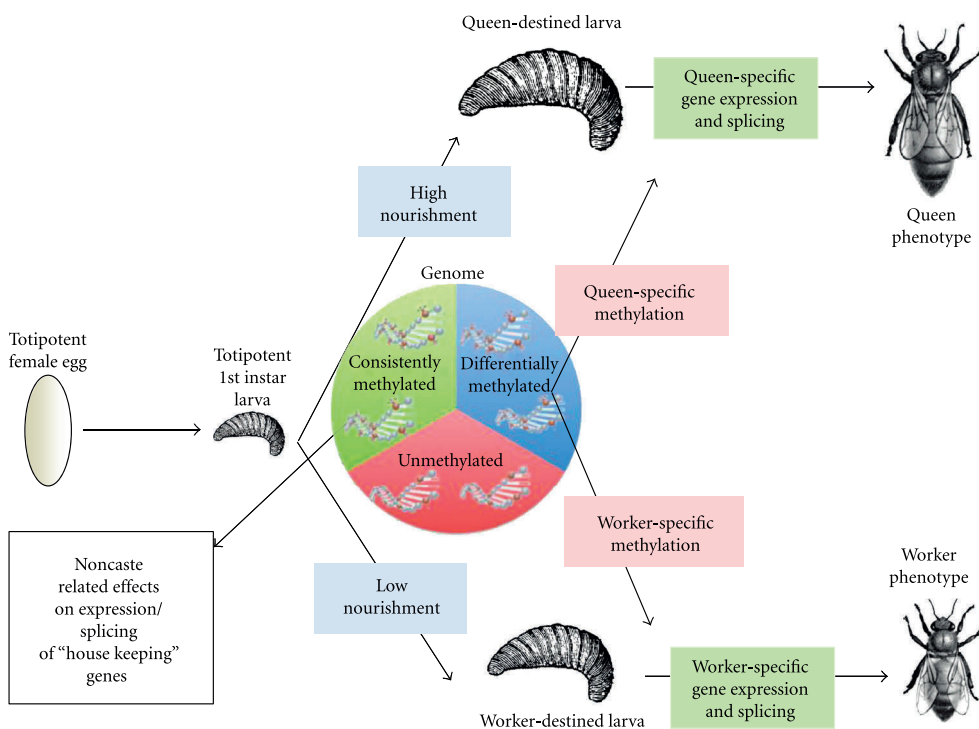
Figure 1.3. Schematic describing the influence epigenetic regulation by DNA methylation in phenotype development in honey bees. Each female egg begins in a totipotent state but develops different as a result of different DNA methylation caused by changes in nutrition. The differential methylation affects gene expression and gene splicing resulting in different growth, metabolism, and development that drive honey bee phenotype [93].

DNA methylation contributes to DNA changes is several ways. Accumulation of DNA methylation in the promoter region of genes is often associated with gene expression downregulation [94], [97], [98]. Generally, DNA methylation can lead to transcriptional repression in three ways [98]. The added methyl groups are capable of physical blocking the binding of transcription factors to gene regulatory regions. Additionally, methylated cytosines can attract methyl-binding proteins such as the methyl CpG binding proteins (MECP-1 and MEPC-2) chaperoning all (de)regulatory transcription factors to sites with dense

CpG-islands [99]. Consequently, accumulation of DNA methylation can result in a more condensed structure of the chromatin [89].
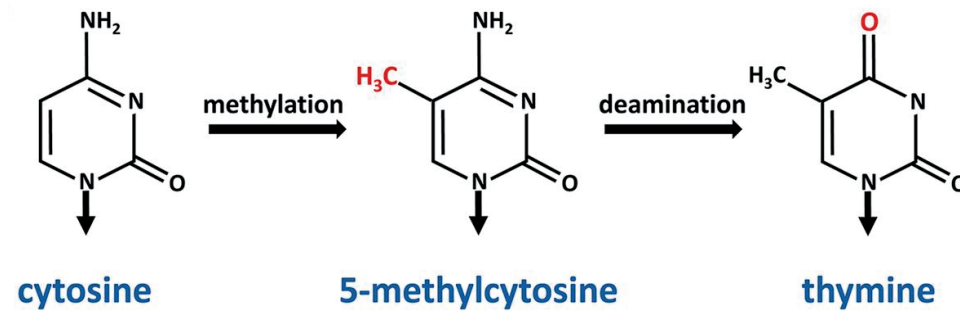
Figure 1.4. Schematic representation of the biochemical cytosine methylation and deamination of methylated cytosine to thymines. Adapted from [100] with permission.

DNA methylation can be enzymatically removed [101], therefore methylation status is a dynamic way of regulation of gene expression. However, DNA methylation is stable and does not reverse spontaneously. Consequently, DNA methylation status is copied during DNA replication, DNA methylation status is inheritable [102]. Even DNA released from their host cells retains any methylation of cytosines allowing for methylation detection in bodily fluids. CpG sites methylation status is maintained because CpG sites form genomic palindromes because cytosine and guanine nucleotides occur in base pairs, a CpG on either the forward or reverse strand of the DNA will have a reverse complimentary CpG on the other DNA strand. Since the methylation status of these CpGs is generally similar on both strands, DNA methylation status of these complimentary CpGs is transferred during cell division making DNA methylation inheritable to both daughter cells during mitosis [103]. About 1% of the human genome consists of CpG sites [104] of which 70 to 80% percent is methylated [105]. Additionally, these CpG sites are not evenly distributed [104]. Actually, the vast majority of the human DNA is void of CpG sites. The accumulation of CpG sites seems particularly high near gene regulatory regions that serve as important binding sites for the transcriptional complex near enhancers and Transcription Start Sites. Such CpG rich regions defined with a statistically significantly increased CpG density, are referred to as CpG islands [106].

DNA methylation levels are maintained by DNA Methylation Transferases (DNMTs) (Figure 1.5) [107]. There are three known DNMT proteins: DNMT1, DNMT3a and DNMT3b. DNMT1 maintains concordant DNA methylations status of opposite CpG sites on the different DNA strands. DNA methylation thus maintains tissue-specific DNA imprinting during cell-division [108], [109]. DNMT3a and DNMT3b facilitate the introduction of new DNA methylation of previously unmethylated CpG sites[107]. DNA methylation as a biomarker has several advantages compared to other tumor biomarkers such as mutations or RNA expression. The methylation status of a CpG sites can exist in only two states: methylated or unmethylated. In contrast, single nucleotide changes caused by mutations can change any base to any other base, creating a much larger amount of variability in outcome. In addition, changes of a single nucleotide do

not necessarily lead to changes in the encoded protein since several nucleotide triplets can encode for the same amino acid. And a single amino acid change does not always lead to functional changes in the tertiary structure of a protein. The binary state of the outcome of changes in DNA methylation, makes DNA methylation a much easier to study phenomenon than mutations. Moreover, changes in DNA methylation levels occur in higher frequency than mutations and allow for the detection of more tumor cells (Figure 1.6) [97]. And finally, changes in DNA methylation occur earlier in tumorigenesis and in general precede mutations (Figure 1.6) [97]. Combined with the reversibility of DNA methylation compared to mutations, DNA methylation has a very high potential of diagnostic applicability [97].
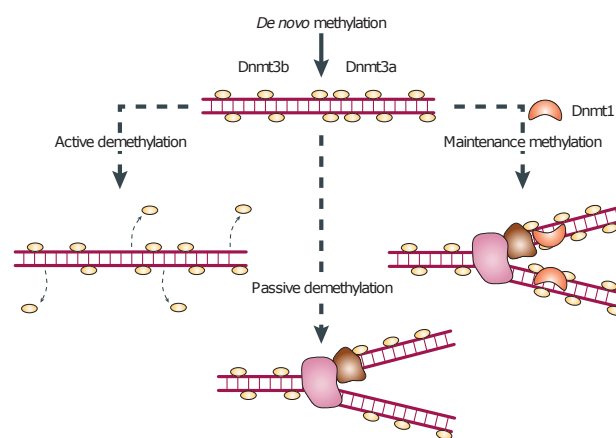
Figure 1.5. DNA methylation is maintained by the family of DNA Methylation Transferases proteins. DNMT1 copies the methylation status of the mother strand to the daughter strand during DNA replication, maintaining methylation during inheritance. DNMT3a and DNMT3b lead to de novo methylation of previously unmethylated CpG sites. Loss of methylation can be achieved due to the activity of a vast collection different compounds or lost by lack of DNMT1 methylation during DNA replication [110].

Interestingly, during carcinogenesis genome-wide overall hypomethylation is observed (Figure 1.6) [111]. This global loss of methylation contributes to tumor development through chromosomal instability as result of changes in chromatin structure, reactivation of transposable elements such as LINE-1, which are normally silenced by hypermethylation, and loss of imprinting causing expression of genes silenced in normal tissue [97]. Besides genome-wide hypomethylation, CpG islands tend to become hypermethylated during tumorigenesis which could lead to the repression of specific tumor suppression genes such as the DNA-repair gene BRCA1 in breast cancer (Figure 1.6) [112].

DNA methylation markers in the clinic

A classic example of a DNA methylation marker in the diagnostic clinical setting is methylation detection of the GSTP1 gene. A meta-analysis showed that hypermethylation of the GSTP1 promoter occurs in 82% of all prostate cancer and in only 5% of normal prostates. This establishes GSTP1 methylation status as a powerful diagnostic tool [113].
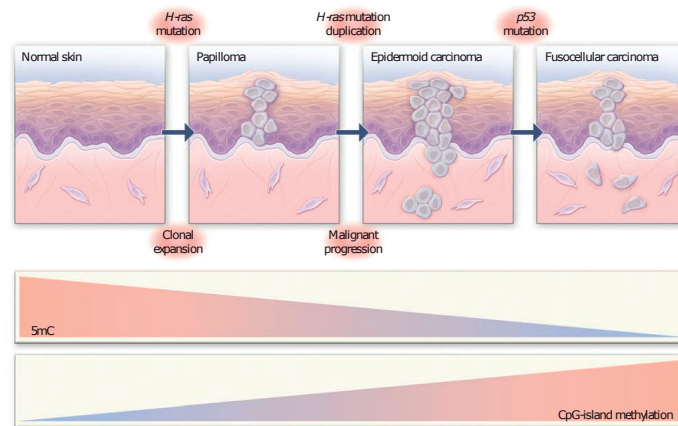
Figure 1.6. During carcinogenesis overall levels of methylated cytosines (or 5mC) decline while CpG Islands specifically are hypermethylated. These changes in genome wide levels of 5'-methyl-cytosines (5mC) and the hypermethylation of CpG Islands in particular occur in higher frequency and earlier in tumorigenesis than the accumulation of DNA mutations. Reproduced with permission from [97], Copyright Massachusetts Medical Society.

The methylation status of the MGMT promoter is being employed as a driver for decision making in treatment modalities in neuro-oncology [114]. In glioblastoma patients undergoing chemotherapy with alkylating agents, the methylation status of the MGMT promoter predicts a greater response to treatment with alkylating agent chemotherapy [115]. The median survival for patients with a methylated MGMT promoter was 21.7 months compared to 15.3 months for patients with an unmethylated MGMT promoter after treatment [115]. Patients with a hypermethylated MGMT promoter respond especially better to treatment with temozolomide than glioblastoma patients with a unmethylated promoter [114]–[116]. However, the relation between the MGMT promoter methylation status and patient survival is not consistent. In addition, the methylation status of MGMT is not persistent with MGMT mRNA levels. Patients have been identified with high MGMT expression but a methylated MGMT promoter as well as patients with unmethylated MGMT promoters but low MGMT expression [117]. This shows that expression levels are not only explained by methylation and other mechanisms are involved. This observation illustrates the complexity of mRNA regulation by DNA methylation and raises questions about where the transcriptional restricting threshold of DNA methylation clinically lies.

Progress has also been made in the diagnosis of tumor metastasis using DNA methylation markers. In breast cancer epigenetic disruption of expression of the cell-adhesion molecule E-cadherin enhances the metastatic potential of cancer cells [118]. DNA methylation-mediated down-regulation of adhesion molecules has also been observed in the lymph node metastases of melanoma and head and neck cancer [119]. More elaborate studies have identified a miRNA methylation signature in primary tumors that is associated with lymph node metastasis in HNSCC and other cancers [120]. More recently, using machine learning and whole-genome methylation data from The Cancer Genome Atlas, a DNA methylation profile has been developed that correctly identified 19 of 20 breast cancer metastases and 29 of 30 colorectal cancer metastases to the liver [121].

In OSCC specifically various studies reported genes that are frequently hypermethylated [122], [123] such as CDH1, CDKN2A, MGMT, DAPK1, RARB, and RASSF1, but only a few of these genes are predictive for LN metastasis [124], [125]. Several biomarkers reported in other cancers associated with cell migration and invasion in vitro [126] and with the presence of nodal metastasis [126], [127] have not been investigated in OSCC.

Body fluids can contain whole cells as well as partial DNA fragments originating from tumors. Due to the stability of the methylation cytosines, the DNA methylation in these bodily fluids (like saliva, sputum and plasma) can be used to determine the methylation status of the whole primary tumor [128]. Even when the location of the primary tumor is unknown, the DNA methylation in the body fluids can be used for cancer diagnosis. That is because body fluids can carry cell-free DNA that originated from the occult tumor that shares genetic aberrations that allow diagnosis as well as guidance in selecting therapeutic modalities [129]. For example, in 100% of the patients with squamous cell lung carcinoma hypermethylation of the CDKN2A and MGMT was found up to 3 years before clinical diagnosis [130]. Studies like this show the application of the detection of DNA methylation in body fluids is a promising new non-invasive early diagnostic tool.

Methylation sensitive endonuclease DNA digestion

Endonucleases are enzymes that cut the DNA in sequence specific locations. Some of these restriction enzymes are specific for sequences containing a CpG dinucleotide (reviewed by [131]). For some of these enzymes the activity is either blocked or enabled by the presence of a methylation residue on one of the nucleotides in the target sequence. Different endonucleases can target the same DNA sequence but may have a different DNA methylation sensitivity. This DNA methylation dependent activity can be used to measure the methylation status of target sequences of these endonucleases. Downside of these techniques is that high amounts of incomplete restriction enzyme digestion of target DNA sequences occur, as well as that most of these techniques rely on Southern-blotting which requires a lot of DNA of high molecular weight which can be problematic to acquire from tumors

DNA methylation detection

DNA methylation in vivo is maintained during DNA replication by DNA methyltransferases where thenewly synthesized DNA strand inherits the DNA methylation pattern of the mother DNA strand throughDNMT1 activity [108], [109](Figure 1.5). During PCR based methods DNMT1 is absent causing all newlysynthesized DNA strands to be complete unmethylated [131]. Therefore, the DNA methylation status ofthe original DNA template is obscured and cannot be studied using PCR-based methods. To overcomethis, techniques have been developed that translate CpG methylation status first into DNA changes thatare measurable by DNA-based methods [131]): methylation-sensitive/insensitive endonuclease digestion[132], affinity enrichment [133] and the most widely used bisulfite treatment [134].

1

Enrichment of methylated DNA

Methylated DNA fragments can be physically separated from non-methylated DNA fragments using molecules with an affinity for methylated cytosines. Known methylation-specific molecules include antibodies specific for methylated cytosines and proteins that bind methylated DNA such as the recombinant human Methyl-CpG-Binding-Domain 2 (MBD2) [135]. By binding methylated DNA fragments followed by a precipitation step methylated DNA can be measured in absence of unmethylated DNA [136]

Bisulfite treatment

Sodium Bisulfite is a chemical which has been found to cause the deamination of cytosines to an uracil which subsequently results in a thymine during PCR [134]. The principle of this method is that methylated cytosines are protected from this chemical conversion while unmethylated cytosines are not. Thus, the resulting bisulfite treated DNA sequence differs depending on the target DNA methylation-status (Figure 1.7). This process however always requires proper controls to determine the conversion rate of the sample DNA. In addition, the loss of high amounts of cytosines during bisulfite conversion results in very low complexity DNA as the difference between a high amount of cytosines and thymines are lost which makes it harder to distinguish between the resulting sequences. This makes it more difficult to design probes and primers specific for a bisulfite treated DNA sequence. However, progress in bioinformatic pipelines compensates for some of this lost complexity of sample DNA and thus bisulfite treatment is more and more frequently combined with next generation sequencing (NGS) (reviewed by [131]).
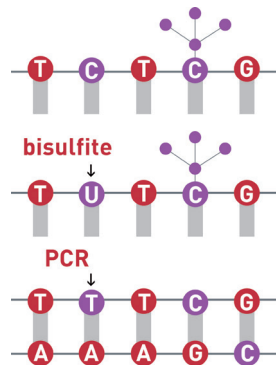
Figure 1.7. Bisulfite treatment of DNA results in different nucleotide sequences based on the methylation status of the cytosines in the target DNA. Unmethylated cytosines are converted into uracil as a result of deamination by the bisulfite treatment. PCR amplification of uracil is treated as if these nucleotides were a thymine resulting in adenines in the first amplicon during PCR. Methylated cytosines are resistant to this conversion by bisulfite. Amplification of these resistant cytosines results in guanines in the first amplicon during PCR where unmethylated DNA treated by bisulfite results in adenines. Picture used from Diagenode with permission: https://www.diagenode.com/en/applications/dna-bisulfite-conversion.

To measure DNA methylation different analysis methods have been developed. In summary there are two different kinds of methods for DNA methylation measurements: typing and profiling methods (Figure 1.8). Typing technologies are used to assess methylation in a limited amount of genomic locations and is generally suited to measure a lot of different samples at the same time for only a few markers (Figure 1.8).

These techniques rely on prior knowledge or selection of genomic regions. Most of these techniques rely on PCR steps such as Methylation Specific PCR, pyrosequencing and Sanger Sequencing.

Methylation Specific PCR (MS-PCR or MSP) is a variation of conventional PCR performed on bisulfite treated DNA. For each locus of interest primer pairs target either the product of unmethylated or methylated DNA after bisulfite treatment (Figure 1.7)[138]. During PCR the methylated or unmethylated products are amplified separately. Since this technique relies on the difference in target sequence of the primers, MS-PCR only interrogates the methylation status of one or two CpG sites in the primer annealing sites. A higher amount of CpG sites in the primers would facilitate more unspecific primer binding and therefore is not ideal [131], [137]. Best specificity is achieved when the CpG site is located at the 3' end of the forward primers. By combining MSP with fluorescent-labeled probes or a fluorescent dye specific for double-stranded DNA, MSP can be measure in a quantitative real-time manner called Quantitative MSP (Q-MSP) [139].
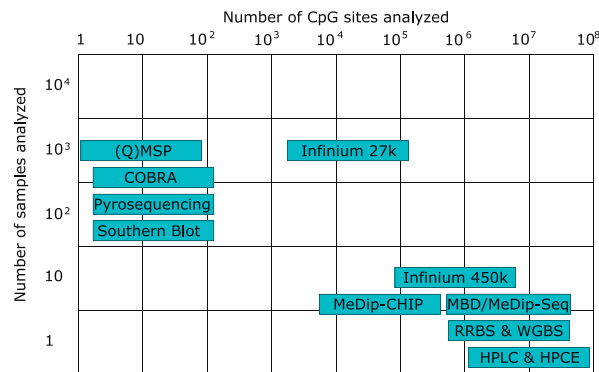
Figure 1.8. Examples of different methylation detection techniques. On the Y-axis the number of samples that can be simultaneously assessed by a technique are depicted. On the X-axis the amount of CpG sites that can be measured per single run is shown. Adapted from [131] with permission.

Like MSP, pyrosequencing relies on bisulfite treatment of DNA and PCR amplification [140]. However, for optimal pyrosequencing analysis unbiased amplification of methylated and unmethylated DNA using primers targeting CpG free sequences is needed. After an isolation step of complete amplicons using a biotinylated universal primer, up to 80 nucleotides of the amplicon are sequenced using a sequencing primer. The DNA sequence is determined using nucleotides with a pyrophosphate group which is released during polymerase incorporation of these nucleotides. The amount of released pyrophosphate is directly proportional to the number of incorporated nucleotides and is quantified by measuring light released by luciferase which is an enzyme driven by the released pyrophosphate. Pyrosequencing determines the ratio of methylated and unmethylated cytosines for each CpG residue separately. A disadvantage is that pyrosequencing is not always possible when no sufficient CpG-free flanking sequences are available.

COBRA, HPLC and methylation-sensitive restriction-enzymes

Combined bisulfite restriction analysis (COBRA) uses bisulfite treatment and methylation unspecific PCR amplification followed by DNA digestion by restriction enzymes [141], [142]. By using restriction enzymes that target a CpG site such as the CGCG targeting BstUI, CpG sites can be found that were originally methylated in the sample DNA. Because the target sequences recognized by these endonucleases are longer and can therefore contain multiple CpG, it can be impossible to target certain CpG sites, that are in proximity to other CpG sites, this is specifically problematic when using COBRA. In addition, the resolution of COBRA is lower compared to MSP and pyrosequencing when using an endonuclease that targets multiple CpG sites simultaneously.

By using methylation-sensitive and methylation-insensitive endonucleases that target the same CpG containing sequence, DNA methylation can be assessed by measuring the size of the DNA fragments after digestion using Southern Blotting[143], [144]. Southern blot is a classic technique that is traditionally used to semi-quantitatively measure the amount of DNA fragments of specific sizes [145]. The DNA is separated based on size using electrophoresis, then the DNA is transferred to a membrane and finally hybridized with a labelled DNA probe.

High performance capillary electrophoresis (HPLC) [146] and High-performance liquid chromatography (HPCE) [147] are used to separate DNA molecules by molecular weight, charge, size, hydrophobic potential and/or conformation. After the whole DNA sample is digested (for HPLC) or hydrolyzed (for HPCE) into single nucleotides, the ratio of normal cytosines and methylated cytosines can be determined. HPCE is cheaper and faster than HPLC but is more sensitive to improper DNA isolation due to the chemical separation of the nucleotides.

Methylation Profiling

Methods of broader DNA methylation measurement are called methylation profiling techniques. These methods rely less on predetermined regions of interest and allow a more unbiased methylation measurement on a much larger scale. Usually using different combinations of NGS with a high number of probes or primer, these profiling techniques allow for a more general discovery tool for differences in DNA methylation levels. Due to the increased scale and subsequent high-costs these techniques are less suited for the testing of extensive patient populations but are a great first step in pinpointing differential methylation levels in target groups (reviewed by [137]).

Whole-Genome Shotgun Bisulfite Sequencing (WGSBS) is basically the whole genome sequencing of bisulfite treated DNA. Sample DNA is sheared into small fragments and bisulfite treated (reviewed by [131]). The bisulfite converted DNA is then attached to adapters to allow PCR amplification and sequencing [148]. This combination of bisulfite treatment and whole genome sequencing gives the highest scale of coverage as well as the highest possible resolution of a single CpG site. However, because of the reduced DNA complexity due to bisulfite treatment, the mapping of sequencing results to the reference genome becomes difficult. Overall WGBS is very expensive to perform and requires very high amounts of DNA input which can be problematic when working with precious and low-quality sources of DNA such as formalin-fixed paraffin-embedded (FFPE) samples [131].

Reduced representation bisulfite sequencing (RRBS) combines WGBS with restriction enzymes to enrich for CpG high loci [149]. Restriction enzymes are used that cut the DNA at the end of CpG sites which is then used as a target site for adapter ligation. Next DNA fragments are size selected, bisulfite treated and finally sequenced. RRBS is efficient in capturing the most CpG rich regions such as CpG islands and promoters. But RRBS requires very little DNA input. However, regulatory regions with lower CpG content such as CpG island shores are missed as well as CpG rich regions that do not contain the target sequence of the used endonuclease [131], [137].

Methyl-CpG-binding domain sequencing (MethylCap-Seq) is a method of enriching DNA for loci with high amounts of methylated CpG sites followed by NGS [150]. The methylated DNA fragments are bound by Methyl-CpG-binding domain (MBD) proteins and separated from the unbound DNA fragments and subsequently sequenced [136], [151]. Unlike WGBS and RRBS, MethylCap-seq does not require enzymatic or bisulfite treatment which removes the limitation to target sequences that contain a particular restriction enzyme target site. However, due to the lack of a bisulfite conversion MethylCap-Seq does not contain the single base resolution of bisulfite-dependent techniques [137]. The DNA enrichment step makes MethylCap-Seq a relative cheap method. Although MethylCap-Seq enrichment is not limited to particular CpG rich regions such as CpG islands, DNA fragments need to contain several spatially-close methylated CpG sites before MBD binds optimal to methylated DNA fragments. This method is especially useful as a discovery tool as it covers methylated fragments of the whole genome.

Methylated DNA immunoprecipitation (MeDIP) is technique that combines DNA sequencing with methylated DNA enrichment comparable to MethylCap-Seq [152]. In contrast to MethylCap-Seq, MeDIP-Seq uses an antibody for methylated cytosines. While this antibody-dependent approach introduces a smaller dependency on a certain threshold of methylated CpG sites that is present in MethylCap-Seq [153], MeDIP-Seq cannot distinguish between a methylated cytosine in a CpG dinucleotide and a single methylated cytosine [137].

The Infinium 450k is a large-scale microarray containing over 485000 probes [154]. These oligos cover up to 96% of all CpG islands and more than 99% of all known promoters [154]. After whole genome amplification a DNA sample is bisulfite treated and applied to the microarray chip. These chips contain two oligo-probes for each locus, one for the bisulfite converted product of the methylated target sequence and one for the bisulfite converted product of the unmethylated DNA target sequence. All array oligos end before the cytosine of a CpG site. After the hybridization of target DNA to the array oligos, a single nucleotide extension is performed with labelled nucleotides to determine the result of the bisulfite-converted nucleotide in the target CpG of the array oligo [155]. Main advantages of the 450k microarray are that this technique is relatively cheap, requires little DNA input and contains previously identified regions of differential methylation [131], [137]. Compared to the other sequencing techniques such as WGBS and RRBS, methylation assessment is limited to the predetermined and designed microarray oligo's [137]. To increase the coverage of the DNA with the Infinium microarray approach an increase in the amount of array oligo's is necessary, which will make the Infinium platform more expensive.

MeDIP-CHIP uses methylated DNA enrichment with an antibody specific for methylated cytosines in combination with a microarray chip[156]. After isolation of methylated DNA comparable to MeDIP-Seq, the DNA is hybridized to a microarray. The chips available for application with MeDIP enrichment are limited and contain a maximum of 200000 probes allowing for a very limited view of the methylome [137].

In conclusion, the most commonly used techniques are based on restriction enzyme digestion, affinity enrichment and bisulfite treatment, combined with microarrays or NGS [157]. The vast majority of loci specific techniques rely on PCR based amplification and are easily adapted to commercial platforms. These platforms are being employed in many clinical labs with high sensitivity and specificity [158]. More specifically, MSP is the most widely used locus specific bisulfite based DNA methylation analysis and has been validated in a large number of clinical samples [137]. To select the most suited technique depends on the research questions, the costs, the quality and volume of sample DNA as well as the degree and nature of the expected of DNA methylation levels (Figure 1.9) [159].

The wide variaty of available techniques reflects the complexity of DNA methylation detection and the various stages necessary for making DNA methylation a biomarker suited for the clinic. While, for example, most progress in metastasis detection in OSCC has been made with an extensive gene expression signature [78], application in the clinic is not feaseble (yet). Employement of the techniques requires fresh or frozen tumor samples and are expensive [160].

For the discovery of new methylation marker large datasets of properly defined patient cohorts are required. It often takes years to accumulate such datasets for proper retrospective studies. This puts great limitations on the available samples. DNA quality and volume are often an important consideration in choosing the right technique and differs highly between platforms.

Recent developments in DNA methylation has elucidated the role of a cytosine methylation intermediate: hydroxymethylated cytosines [161], [162]. The role of hydroxymethylation is biologically different from methylation and some techniques are incapable of distinguishing between the two modified cytosines [137].

Additionally, since the relation between hypermethylation of gene promoters and gene expression is not always liniar or black and white, some genes require quantitative analysis while in other cases qualitative analyses is sufficient.

To properly establish DNA methylation markers for use in the clinic wide-spread validation by a range of institutes is required. To have such broad support, the method used to asses the methylation status of the biomarker needs to preferablly be commercially available and applicable by a wide-range of labs [159]. Next generation Sequencing based methods are quickly getting more traction as prices of equipment and runs drop. However, for the fast, easy and cheap validation of datasets uncovered by NGS, techiniques such as MSP will still be required due to their ease and low cost [137], [159].

SCOPE OF THIS THESIS

The goal of this thesis is to generate and explore an **Oral Squamous Cell Carcinoma-specific Methylome. Using this OSCC-Methylome**, new DNA methylation markers will be identified either associated with the presence of nodal metastases in pre-treatment primary tumor samples or with the presence of tumor cells in saliva during treatment of patients with OSCC.

In **chapter 2** previously reported DNA methylation markers associated with N-status in HNSCC and other malignancies were selected from literature searches to investigate the predictive value of these markers for the presence of nodal metastases in a cohort of 70 early-stage OSCC selected from a well-established database of patients treated in the University Medical Center Groningen between 1997-2008 [163]. For this purpose, we compared methylation status detected by MSP of pretreatment biopsies of OSCC patients without lymph node metastases (the N0-group) defined as patients without local-regional metastases with post-operative follow-up of at least 5 years versus those with lymph node metastases (the N+ group) defined by the presence of lymph node metastases at the time of or within 2 years after surgery.

To identify novel DNA methylation tumor markers predictive for the presence of nodal metastases in OSCC, we used MethylCap-Seq, an innovative high-resolution technology to uncover DNA-methylation in a genome-wide manner. To identify markers that are associated with lymph node status, a methylome will be generated from 6 pN0 and 6 pN+ OSCC. Candidate differentially methylated markers will be identified and statistically ranked. The highest differentially methylated markers will be validated on a large and well-established patient cohort using different typing techniques including MSP, pyrosequencing and immunohistochemistry of involved functional genes. In addition, several public datasets for external validation and confirmation of the predictive value and biological impact of selected DNA methylation markers will be used.

In **chapter 3**, the initial construction of the OSCC methylome using MethylCap-seq on the cohort of 6 pN0 and 6 pN+ OSCC will be described. With this methylome of metastatic OSCC, we aim to identify epigenetic markers that are associated with lymph node metastases in OSCC using both statistical analyses as well as experimental validation of the epigenetic regulation of lymph node metastasis in OSCC. Hypomethylation of the WISP1 gene was further evaluated as an epigenetic marker for pN+ status in OSCC.

Because methylation status of genes is related to gene expression, in **chapter 4** the MethylCap-seq data were combined with a gene expression signature predictive for pN+ status in OSCC and OPSCC [78]. Genes with increased methylation status and mRNA down-regulation in pN+ OOSCC will be validated using an independent OSCC cohort by both immunohistochemistry and pyrosequencing. Results will be confirmed on data retrieved from The Cancer Genome Atlas (TCGA). RAB25 was validated in greater detail to validate the clinical value for the detection of lymph node metastases.

In **chapter 5**, further data analysis was performed of the MethylCap-Seq data by combining available differentially methylated methylation markers with extensive validation of public datasets of TCGA. A new

algorithm was developed to select a marker that is down-regulated by hypermethylation in pN+ OSCC to identify a biomarker detectable by MSP and suitable as a therapeutic target. The three most relevant candidate markers identified by this novel algorithm (KCNAB1, LAMP3 and S100A9) will be further validated by pyrosequencing and subsequently immunohistochemistry as potential new markers associated with lymph node status.

OSCC are characterized by a relative high frequency of local recurrences of 20-30 % compared to other cancers [6][164]. In addition, these local recurrences have a low 30% diagnosis rate at an early localized clinical stage [6]. The early detection of tumor cells in saliva of patients with OSCC during follow-up after surgery might be of relevance for appropriate therapy-management. To increase the sensitivity of the detection of tumor cells in saliva, molecular detection using OSCC specific methylation markers was reported previously [165]–[167]. In **chapter 6**, the combined methylome of all 12 OSCC cases was used to identify new OSCC specific methylation markers for the early detection of tumor cells in saliva. For this purpose, our OSCC-methylome will be compared to methylation data of two pools of leukocytes and other malignancies and healthy tissues as methylation negative controls using the Map of the Human Methylome database [168]. New candidate methylation markers will be validated in saliva collected from 10 OSCC patients. Saliva from five orthognathic surgery patients and five implantology patients served as healthy controls. The selected biomarkers will be compared to three reported methylation markers for an improved sensitivity and specificity to detect tumor cells in patients with OSCC.

Chapter 2

Identification of methylation markers
for the prediction of nodal metastasis in oral and
oropharyngeal squamous cell carcinoma

L.J. Melchers[1,2] M.J.A.M. Clausen[1,2], ,M.F. Mastik[1], L. Slagter-Menkema[1, 3],
J.E. Van der Wal[1], G.B. A. Wisman[5], J.L.N. Roodenburg[2]*, E. Schuuring[1]*

* Both authors contributed equally to this work.
[1] Departments of Pathology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[2] Departments of Oral and Maxillofacial Surgery, University of Groningen,
University Medical Center Groningen, Groningen, the Netherlands.
[3] Departments of Otorhinolaryngology/Head & Neck Surgery, University of Groningen, University
Medical Center Groningen, Groningen, the Netherlands.
[4] Departments of Gynecologic Oncology, University of Groningen,
University Medical Center Groningen, Groningen, the Netherlands.

ABSTRACT

Hypermethylation is an important mechanism for the dynamic regulation of gene expression, necessary for metastasizing tumor cells. Our aim is to identify methylation tumor markers that have a predictive value for the presence of regional lymph node metastases in patients with oral and oropharyngeal squamous cell carcinoma (OOSCC).

Materials and Methods:

Significantly differentially expressed genes were retrieved from four reported microarray expression profiles comparing pN0 and pN+ head-neck tumors, and one expression array identifying functionally hypermethylated genes. Additional metastasis-associated genes were included from the literature. Thus, genes were selected that influence the development of nodal metastases and might be regulated by methylation. Methylation-specific PCR (MSP) primers were designed and tested on 8 head-neck squamous cell carcinoma cell lines and technically validated on 10 formalin-fixed paraffin-embedded (FFPE) OOSCC cases. Predictive value was assessed in a clinical series of 70 FFPE OOSCC with pathologically determined nodal status.

Results:

Five out of 28 methylation markers (OCLN, CDKN2A, MGMT, MLH1 and DAPK1) were frequently differentially methylated in OOSCC. Of these, MGMT methylation was associated with pN0 status (p = 0.02) and with lower immunoexpression (p = 0.02). DAPK1 methylation was associated with pN+ status (p = 0.008) but did not associate with protein expression.

Discussion:

In conclusion, out of 28 candidate genes, two (7%) showed a predictive value for the pN status. Both genes, DAPK1 and MGMT, have predictive value for nodal metastasis in a clinical group of OOSCC. Therefore, DNA methylation markers are capable of contributing to diagnosis and treatment selection in OOSCC. To efficiently identify additional new methylation markers, genome-wide methods are needed.

Keywords: biomarker, DAPK1, expression, head and neck cancer, lymph node metastasis, methylation, MGMT, oral cancer

INTRODUCTION

Oral and oropharyngeal squamous cell carcinomas (OOSCC) compose the largest subgroup of head and neck cancer, and are estimated to have caused over 42,000 new cases in the United States in 2014[164]. OOSCC are characterized by regional metastatic spread to the lymph nodes of the neck in an early stage. Patients with regional lymph node metastases are generally treated with curative intent. When regional metastases are not adequately treated, distant spread results, which is considered as incurable disease. Therefore, it is essential to make an accurate assessment of the nodal (N) status of the neck to adequately treat patients with OOSCC[169]. However, current imaging methods to assess the presence of metastases in the palpation-negative neck showed a sensitivity of 60–70%.[170] Sentinel lymph node biopsy, when performed intra-operatively on frozen sections, has a comparable sensitivity of 50–70%[171], [172]. DNA hypermethylation is an important mechanism for the regulation of gene expression, in both physiological and pathological conditions[94]. DNA hypermethylation is a form of epigenetic regulation, in which the genetic sequence is not altered, but CH3-groups are added to the cytosine of CpG dinucleotides which, when present in the promoter region of a gene, leads to transcriptional repression of the associated protein. This process is reversible, and hypomethylation leads to reactivation of gene transcription[173]. Thus, hypermethylation of tumor suppressor genes and hypomethylation of oncogenes may contribute to carcinogenesis and cancer progression[174]. Because of its dynamic nature, methylation is a possible candidate mechanism for the dynamic regulation of gene expression during metastatic progression of OOSCC cells [175]. Moreover, several demethylating drugs have been developed and show that treatment results in re-expression of formerly hypermethylated genes. Decitabine and Azacytidine are therapeutic demethylating agents and have already been used in treatment of specific hematological malignancies[176]. Therefore, methylation can also be therapeutically targeted[177]. Methylation-specific PCR (MSP) is one of the most widely used methylation detection methods, because of its cost-effectiveness and high sensitivity[178]. The availability of such a sensitive detection method may allow methylation to become a prognostic or diagnostic tool in the clinical setting. For example, hypermethylation of MGMT in gliomas has been shown to predict patient response to alkylating chemotherapy[179]. Various studies have identified several genes that are frequently hypermethylated in OOSCC [122], [123] such as CDH1, CDKN2A, O-6-methylguanine-DNA methyltransferase (MGMT), death-associated protein kinase 1 (DAPK1), RARB, and RASSF1, but only few of those have been associated with metastasis[124], [125]. In other cancers, various methylation markers have been associated with cell migration and invasion in vitro[126], [180] and the presence of nodal metastasis[126], [127]. In this study, we set out to identify novel methylation markers that are associated with the presence of lymph node metastases in patients with OOSCC. We selected candidate genes with a CpG island from the most differentially expressed genes, as reported in 4 published metastasis-associated gene profiles[83], [181]–[183] and the genes from these 4 profiles that were functionally methylated (showing increased expression after demethylating treatment), as determined in a previous study performed in our lab[142]. Additionally, we selected several genes that were reported to be associated with metastasis in previous studies in squamous cell carcinomas. These methylation markers were tested by MSP in a clinical series of OOSCC with pathologically determined N status for their predictive value for the presence of lymph node metastases.

MATERIALS AND METHODS

Selection of candidate genes

To select candidate genes that are regulated by methylation and associated with lymph node metastasis, we used reported microarray data from four independent studies in HNSCC[83], [181]−[183]. All selected candidate genes should have a CpG island present in the promoter region of the gene, and a negative correlation with nodal metastases, as methylated genes have an associated downregulation on mRNA level. From these lists of genes we selected (Figure 2.1): (1) all genes found in more than one of the four expression profiles[83], [181]−[183]; (2) the 5 highest ranking genes from the two studies that performed genome-wide arrays[83], [181]−[183]; (3) candidate genes that were reported in the 4 HNSCC expression profiles [83], [181]−[183] and showed functional methylation (increased expression after treatment with 5-aza-20-deoxycytidine (DAC)/trichostatin A (TSA) in vitro and an association with lymph node metastasis in cervical squamous cell carcinoma, in a previous study performed in our lab [142], [184]. Furthermore, four genes were selected that have been described to be associated with invasion and metastasis in squamous cell carcinoma: GJB6 [185], OCLN [186], TJP1 [187], and CD44 [188]. In this way, a total of 24 genes were selected that were not reported to be methylated in OOSCC and, consequently, were potential new candidate metastasis-associated genes whose expression might be regulated by methylation. Four genes (MLH1, MGMT, CDKN2A, and DAPK1) were included that showed frequent methylation in HNSCC in the literature[189]−[192] (Figure 2.1).

MSP primer design

For optimal MSP primer design in a region with the highest chance of finding differentially methylated regions[193], all candidate genes were checked for the presence of a CpG island in a range of -500 to +500 bp relative to the TSS, and primers were designed in this region using Methyl Primer Design software (Applied Biosystems, Foster City, CA, USA). Primers that were selected generally had three CGs in their sequence. Maximum product size was set at 160 bp, due to working with DNA isolated from FFPE tissue. For MGMT, CDKN2A, and DAPK1, primer sequences from literature were used[138], [194], [195] (Supplemental Table 2.1).

Candidate gene testing strategy

Selected candidate genes were tested for optimal annealing temperature and MgCl2 concentration on a set of 8 HNSCC cell lines (UMSCC-1, UMSCC-2, UMSCC-8, UMSCC-11a, UMSCC-14a, vuSCC-40, vuSCC78, vuSCC-96) and 2 normal tonsil FFPE samples. After optimization, MSPs were performed on a set of 5 N0 and 5 N+ tumors. All markers that showed methylation in 2 or more tumor samples were further tested on our total patient series (n = 70: 32 pN0 and 38 pN+; Fig. 1). All tumor samples were tested twice in separate experiments. Samples with discordant results were tested for a third time.

DNA isolation

From the FFPE blocks of the tumors, 2 10-mm thick sections were cut and used for DNA extraction. Subsequently, a 3-mm thick section was cut and HE-stained to check if tumor load was sufficient through the sections (preferably >60%). After deparaffinization, DNA isolation was performed, using standard salt-chloroform extraction and ethanol precipitation[163]. For quality control, genomic DNA was amplified in a multiplex PCR containing a control gene primer set resulting in products of 100, 200 300, 400, and 600 bp, according to the BIOMED-2 protocol[196]. Only cases with products 200 bp were included for further analysis.
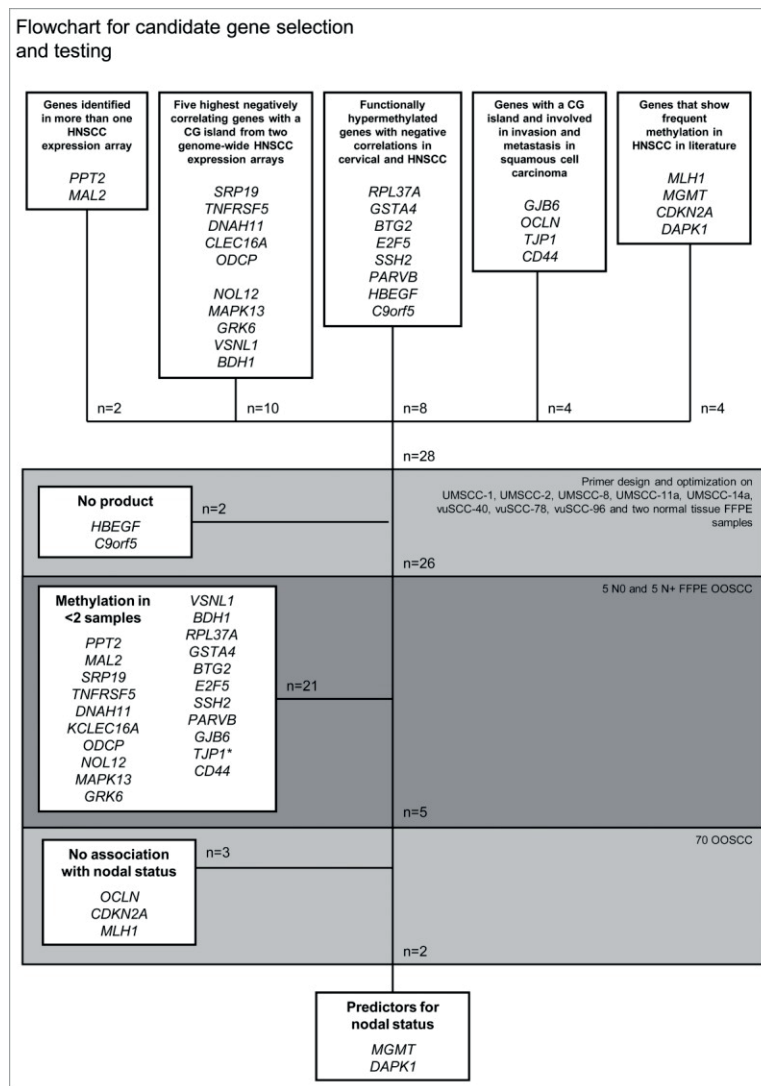
Figure 2.1. Flowchart for candidate gene selection and testing
* TJP1 showed methylation in all samples and was therefore excluded.

Bisulfite treatment and methylation-specific PCR (MSP)

Bisulfite-converted DNA (bisDNA) was made using the EZ DNA methylation kit according to the manufacturer's protocol (Zymo Research, Irvine, CA, USA). Methylation specific PCR (MSP) was performed using 20 ng bisDNA. All MSPs were run as follows: 10 min 95°C, 40 times (1 min 95°C, 1 min Tannealing, 1 min 72°C), 10 min 72°C, ∞ 4°C. Controls consisted of leukocyte DNA that was in vitro methylated by SssI methyltransferase (methylated control) or untreated leukocyte DNA (unmethylated control). Adequate bisulfite conversion was checked by b-actin MSP (Forward: 5' TAGGGAGTATAT AGGTTGGGGAAGTT 3'; Reverse: 5' AACACACAATAACA AACACAAATTCAC 3'). A sample was considered methylated when the methylated product of the right size was visible. It was considered unmethylated when the unmethylated product of the right size was visible and no methylated product was visible. A sample was considered not assessable, when no unmethylated and methylated products of the right size were present. Methylation- and unmethylation-specific PCRs were performed in parallel, and performed at the same annealing temperature (Tannealing), on the same plate.

Immunohistochemistry

TMA sections were deparaffinized in xylene and rehydrated in a graded alcohol series. Antigen retrieval was performed by heating in a microwave oven for 15 min in either Tris/EDTA pH = 9.0 (for MGMT) or EDTA pH = 8.0 (for DAPK1). After antigen retrieval endogenous peroxide was blocked by incubating the slide in 0.3% peroxide solution. After one-hour incubation with anti-MGMT 1:100 (MT3.1, Millipore, Billerica, MA, USA) or anti-DAPK1 1:200 (D1319, Sigma-Aldrich, St. Louis MO, USA), a horseradish peroxidase-conjugated secondary antibody was used, followed by a horseradish conjugated tertiary antibody. Slides were developed with di-aminobenzidene chromogen solution, followed by hematoxylin counterstaining. In addition to the control tissues included on the TMA slide, full sections of the control tissue, specific for each staining, were also included (normal liver for MGMT[197]; normal duodenum for DAPK2 according to manufacturer's protocol).

Analysis of immunohistochemistry

Cases were semi-quantitatively scored, assessing percentage of tumor cells stained and the intensity of staining (0, no staining; 1, weak; 2, moderate; 3, strong). Staining was scored by 2 observers, independently. Discordant results were discussed until consensus was reached. High MGMT expression was defined as moderate to strong nuclear expression in 10% of tumor cells, as reported previously[198]–[200].For DAPK1, scores were given to cell proportion: 0: staining in 50% of tumor cells. Intensity was then scored as 0: negative; 1: weak; 2: moderate; and 3: strong. The final score (ranging 0–9) was obtained by multiplying the cell proportion by the intensity. A final score of <4 was considered to indicate low expression, and 4 was considered high expression[201], [202].

Statistical analysis

Statistical analysis was performed with SPSS version 20. Categorical data were compared using the Chi-square test, or Fisher's exact test, when appropriate. Univariate logistic regression was used to assess the relationship between predictor variables and the dichotomous pN status. All predictor variables with $p < 0.10$ in univariate logistic regression were entered in multiple logistic regression. All tests were performed 2-tailed. Results were considered significant when $p < 0.05$.

Table 2.1. Clinicopathological characteristics

	Total	pN0	pN+
Total patients	70 (100)		
Total tumors	70 (100)	32	38
Sex			
Male	39	19	20
Female	31	13	18
Age at diagnosis (y)			
Median	63.5	64	63.5
Range	25–94	25–89	25–94
Site			
Tongue	26	15	11
Floor of mouth	22	12	10
Oropharynx	9	1	8
Other	13	4	9
cN status			
0	48	31	17
+	22	1	21
pT status			
1–2	50	27	23
3–4	20	5	15
Extranodal spread (only pN+)			
No	21		21
Yes	17		17
Perineural invasion			
No	53	28	25
Yes	14	2	12
Unknown	3	2	1
Lymphovascular invasion			
No	48	25	23
Yes	12	5	7
Unknown	10	2	8
Histological differentiation			
Well	14	13	1
Moderate	42	16	26
Poor	9	1	8
Unknown	5	2	3
Infiltration depth (mm) (n = 65)			
Median	8	5.7	10
Range	0.52–30.0	0.52–25.0	1.90–30.0
High-risk HPV status			
Negative	61	30	31
Positive	3	1	2
Unknown	6	1	5

Table 2.2. Selected candidate genes

(1)   Genes identified in more than one HNSCC expression array

Gene	GenBank ID	Study	Correlation
PPT2	NM_005155	[83], [182], [183]	−0.417
MAL2	NM_052886	[83], [182]	−0.544

(2)   Five highest negatively correlating genes with a CpG island from two genome-wide HNSCC expression arrays

Gene	GenBank ID	Study	Correlation or p-value
SRP19	NM_003135	[83]	Correlation: −0.814
TNFRSF5 (=CD40)	NM_001250	[83]	Correlation: −0.802
DNAH11	NM_003777	[83]	Correlation: −0.776
KIAA0350 (=CLEC16A)	NM_015226	[83]	Correlation: −0.760
ODCP	NM_052998	[83]	Correlation: −0.741
NOL12	NM_024313	[183]	P-value: 0.0001
MAPK13	NM_002754	[183]	P-value: 0.0003
GRK6	NM_001004106	[183]	P-value: 0.0009
VSNL1	NM_003385	[183]	P-value: 0.0013
BDH1	NM_004051	[183]	P-value: 0.002

(3)   Functionally hypermethylated genes with negative correlations in cervical and HNSCC

Gene	Affymetrix ID	Study	Correlation or Z-score
RPL37A	213459_at	[83]	Correlation: −0.162
GSTA4	202967_at	[182]	Z-score: −3.91
BTG2	201236_s_at	[182]	Z-score: −4.58
E2F5	221586_s_at	[83]	Correlation: −0.356
SSH2	230970_at	[83]	Correlation: −0.475
PARVB	37966_at	[83]	Correlation: −0.286
HBEGF	38037_at	[182]	Z-score: −4.11
C9orf5	230764_at	[182]	Z-score: −0.075

(4)   Genes with a CpG island and involved in invasion and metastasis in squamous cell carcinoma

Gene	GenBank ID	Study
GJB6	NM_001110221	[185]
OCLN	NM_002538	[186]
TJP1	NM_003257	[187]
CD44	NM_000610	[188]

(5)   Genes that show frequent methylation in HNSCC

Gene	GenBank ID	Study
MLH1	NM_001258271	[191], [192]
MGMT	NM_002412	[189], [191], [203]
CDKN2A	NM_000077	[189], [191], [203]
DAPK1	NM_004938	[189], [203]

RESULTS

Candidate gene selection and initial testing

Using the strategy outlined in Figure 2.1, 28 candidate genes were selected for analysis (Table 2.2). Two markers did not show any product during the optimization phase and were excluded. Of the 26 markers tested on the initial set of 5 pN0 and 5 pN+ formalin-fixed, paraffin-embedded (FFPE) OOSCC samples, 17 markers were methylated in none of the 10 OOSCC samples, three markers (PPT2, BTG2 and CAV1) were methylated in only one sample, and one marker (TJP1) was methylated in all samples. Five markers showed methylation in two or more tumor samples and were considered eligible for further analysis (OCLN, CDKN2A, MGMT, MLH1, and DAPK1).

OCLN, CDKN2A, MGMT, MLH1, and DAPK1 were tested on 32 pN0 and 38 pN+ cases (Table 2.3). MGMT was methylated in 13/32 (41%) of pN0 and 6/38 (16%) of pN+ cases and showed a significant association with nodal status (p = 0.02). DAPK1 methylation was also significantly associated with nodal status (p = 0.008); however, in contrast to MGMT, DAPK1 was more frequently methylated in pN+ (10/38, 26%) than in pN0 cases (1/32, 3%). OCLN, CDKN2A, and MLH1 showed more methylation in pN+ tumors also, but the difference was not statistically significant (Table 2.3). MGMT had a predictive value of OR = 0.28 (95% confidence interval (CI): 0.09–0.84) and DAPK1 had an OR = 11.1 (95% CI: 1.33–92.1) for the pN status (Table 2.4). The wide 95% CI is probably attributable to the relatively small patient sample (n = 70) used in this study. Multivariate regression analysis revealed that both markers were not independent from currently used clinicopathological predictors, reflected in the cN status. However, the predictive values of MGMT and DAPK1 were independent from each other (Table 2.5A). The combined regression model of MGMT and DAPK1 had a negative predictive value for the pN status of 76% (Table 2.5B).

To assess if methylation of the two predictive markers MGMT and DAPK1 was associated with decreased expression, we performed immunohistochemistry on the available tumor tissue of the same cases that had been used to assess the predictive values of methylation. Because MGMT[204] and DAPK1[205] in particular, are known to be heterogeneously expressed within the tumor, we investigated expression in the tumor center and tumor front separately in 66 OOSCC cases that were present on the tissue microarrays (Figure 2.2). MGMT methylation was associated with low expression both in the tumor front (12% expression in methylated vs. 43% in unmethylated cases) and in the tumor center (26% in methylated vs. 36% in unmethylated cases), but this was only statistically significant in the tumor front (p = 0.02; Table 2.6; Figure 2.3). For DAPK1 methylation, no associations were found with expression in tumor front (p = 1.0) or center (p = 0.14; Table 2.6).

Table 2.3. Cross table analyses of the five genes eligible for testing on the patient series.

Gene		pN0	pN+	P-value
OCLN	U	14	16	0.67
	M	2	4	
CDKN2A	U	27	27	0.19
	M	5	11	
MGMT	U	19	32	0.02
	M	13	6	
MLH1	U	32	36	1
	M	0	1	
DAPK1	U	31	28	0.008
	M	1	10	

M: Methylated
U: Unmethylated
Predictor gene identification

Table 2.4. Univariate and multiple logistic regression with pN status. All assessed with univariate logistic regression. Infiltration depth is continuous (per millimeter).

Variable		Univariate logistic regression		Multiple logistic regression	
		OR	95% CI	OR	95%CI
cN status	0	1		1	
	+	38.3	4.7–310	38.5	3.5–422
pT status	1	1			
	2	3.5	1.11–11.2		
Perineural invasion	No	1			
	Yes	6.7	1.4–33.0		
Lymphovascular invasion	No	NS			
	Yes				
Histological differentiation	Well	1		1	
	Moderate-poor	26	3.1–215	25.9	1.9–351
Infiltration depth	(per mm)	1.1	1.0–1.3		
HR-HPV status	Negative Positive	NS			
MGMT	U	1			
	M	0.28	0.09–0.84		
DAPK1	U	1			
	M	11.1	1.33–92.1		

CI: Confidence Interval
M: Methylated
U: Unmethylated
Immunohistochemistry

Table 2.5. (A) Multiple logistic regression of DAPK1 and MGMT for pN status. (B) Cross table for the DAPK1 and MGMT test combined vs. pN status.

(A)		Multivariate logistic regression	
Variable		OR	95%CI
DAPK1 methylation	U	0	
	M	11.1	1.28−96.7
MGMT methylation	U	0	
	M	0.27	0.08−0.90
(B)	pN status		
Column1		0	+
DAPK1 M or MGMT U	No	13	4
	Yes	19	34

P = 0.003; sensitivity = 89%; specificity = 41%; positive predictive value (PPV) = 64%; negative predictive value (NPV) = 76%.

M: Methylated
U: Unmethylated

Table 2.6. Associations between methylation and expression for MGMT and DAPK1.

MGMT expression		MGMT methylation U	M	P-value
Front	Low	28	15	0.02
	High	21	2	
Center	Low	28	14	0.44
	High	16	5	

DAPK1 expression		DAPK1 methylation U	M	P-value
Front	Low	4	1	1
	High	51	10	
Center	Low	9	0	0.14
	High	44	11	

MGMT expression was not assessable in the tumor center for 3 cases.

DAPK1 expression was not assessable in the tumor center for 2 cases.

M: Methylated
U: Unmethylated

DISCUSSION

The goal of our study was to identify novel methylation markers for the prediction of nodal metastasis. We selected 28 candidate genes, of which two (7%) showed a predictive value for the nodal (N) status. Both genes, DAPK1 and MGMT, have been described as frequently methylated in OOSCC[124], [125] and other cancers[206]. Most candidate genes (12/28) were selected from the most differentially expressed genes in independent microarray studies of N0 versus N+ HNSCC. We hypothesized that gene-specific promoter methylation lead to the observed gene silencing in N+ cases. However, none of these selected genes showed any methylation, indicating that other mechanisms are responsible for their downregulation. One explanation for the finding that the most differentially expressed genes lack promoter methylation is that our selection might have caused a bias toward genes downregulated by other mechanisms because methylation rarely causes complete transcriptional repression. We also selected eight genes that had predictive value in the metastatic gene profiles [83], [181]–[184] and showed upregulation after demethylating treatment in cell lines[142]. However, the functional regulation of these genes by methylation in vitro might not apply to clinical tumor samples, due to (in vivo) intra-tumor heterogeneous methylation[207]. Additionally, genes selected from metastatic profiles reported in microarray studies do not accurately reflect the metastatic genotype, because these signatures are largely platform and analysis related and composition of predictive profiles varies enormously between different studies[208]. In fact, comparing the four microarray studies, shows that no single gene was reported in all four profiles [83], [181]–[183]. This demonstrates that using expression profiles to identify new metastasis-specific OOSCC methylation markers is not effective. Differentially hypermethylated regions (DMRs) in cancer are frequently found in or overlapping CpG islands (»40% of hypermethylated DMRs). Another 30% of hypermethylated DMRs are located in a region of 500 bp flanking the CpG islands[193].33 Our MSP primers were designed in the conventional areas (in CpG islands within -500 to +500 bp from the transcription start site (TSS)), which include 40–70% of the DMRs. However, it is possible that the regions most responsible for transcriptional regulation are located in specific regions outside these areas (CpG island shores)[193]. The CpG island shores are not CG-rich and consequently not useful for optimal MSP primer design. Because we restricted our analysis to the CpG-rich regions close to the TSS to enable optimal MSP design, we cannot exclude that the differentially expressed genes are regulated by DNA methylation in other regions, such as CpG island shores, which contain »15% of the
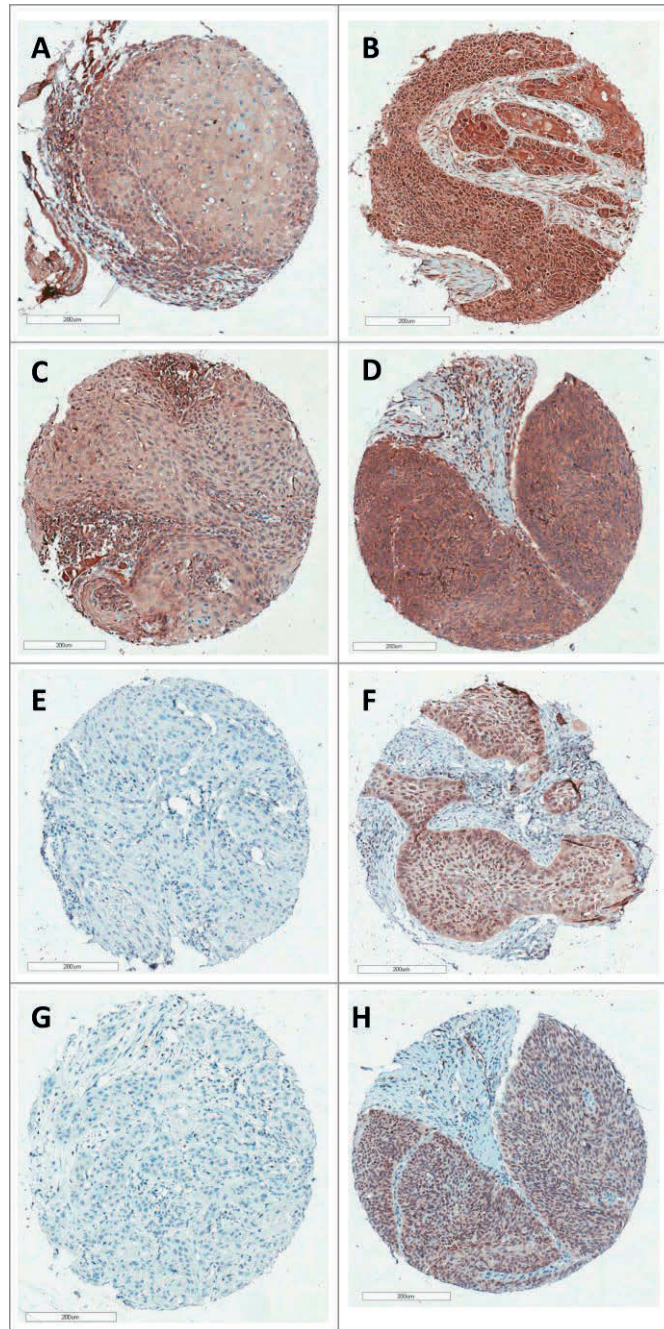
Figure 2.2. Representative examples of immunohistochemical staining. (A) DAPK1 low expression core, tumor center; (B) DAPK1 high expression, core tumor center; (C) DAPK1 low expression core, tumor front; (D)DAPK1 high expression core, tumor front; (E) MGMT low expression core, tumor center; (F) MGMT high expression core, tumor center; (G) MGMT low expression core, tumor front; (H) MGMT high expression core, tumor front.
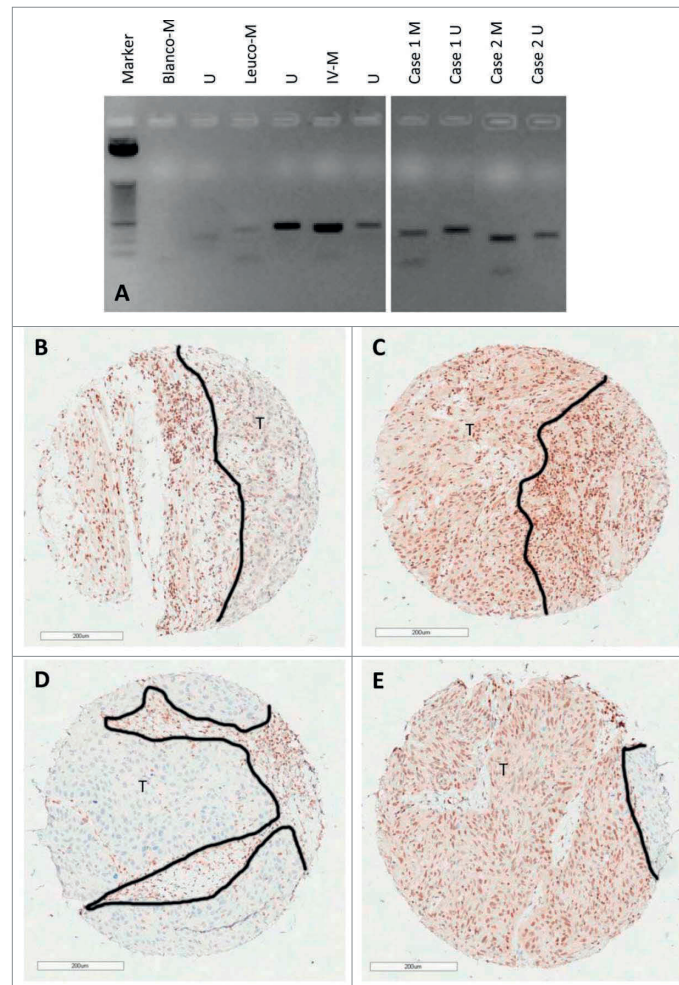
Figure 2.3. Examples of two cases that showed MGMT methylation, associated with low expression in the invasive tumor front, but high expression in the tumor center. (A) MGMT methylation controls [pure water, leucocytes, and IV (in vitro SssI methylated leucocytes)] and two cases. (B) Low MGMT expression in the tumor invasive front (Case 1). (C) High MGMT expression in the tumor center (Case 1). (D) Low MGMT expression in the tumor invasive front (Case 2). (E) High MGMT expression in the tumor center (Case 2). The border of the tumor area is indicated by a black line.
U: unmethylated; M: methylated;
Blanco: pure water control;
Leuco: leucocytes; I
V: in vitro Sss I methylated leucocytes.
T: tumor tissue.

hypermethylated DMRs. The selection of four genes that show frequent methylation in HNSCC produced the two methylation markers that were ultimately found to have predictive value for the presence of lymph node metastases (DAPK1 and MGMT). DAPK1 is one of the most widely studied methylated genes. DAPK1 methylation is frequently found in a wide array of over 20 tumor types[209]. DAPK1 is a tumor suppressor gene, and methylation of this gene has been associated with shorter disease-free survival in surgically treated Stage I lung tumors[209] and with metastasis in several tumor types including, head and neck tumors[210]. This latter study, which used similar primers, found comparable rates of DAPK1 methylation, 15/79 (19%) overall (compared to 16% in our study), and a significant association with N status (27% methylation in N+ group, compared to 26% in our study), confirming the results found in our study. In contrast to the studies in leiomyosarcoma and urothelial carcinoma that utilized the same immunohistochemical scoring method and found associations with methylation status[201], [202], we did not find an association between DAPK1 methylation and protein expression. Because this scoring method might not be reliable in OOSCC, we also analyzed high- and low-expression compared to the median (percentage of tumor cells having moderate or strong expression), and associated this with DAPK1 methylation. Again, no significant associations were found. DAPK1 is a serine/threonine kinase involved in several mechanisms linked to cell death and autophagy. It has pro-apoptotic activity by suppressing integrin-mediated survival signals, thus inducing a specific form of apoptosis, called anoikis. Tumor cells that have loss of anoikis by inactivated DAPK1 are more likely to survive during migration and, therefore, more likely to cause metastases[211]. Furthermore, DAPK1 has an antimigratory effect by blocking integrin-mediated cell polarization[212]. Therefore, DAPK1 downregulation by hypermethylation increases metastasis and tumor cell survival. MGMT is a DNA repair enzyme. MGMT methylation is mostly known for being predictive for better response to alkylating chemotherapy in glioblastoma and, to a lesser extent, to radiotherapy[213]. In OOSCC, several studies assessing MGMT methylation using various techniques did not find associations with N status[189], [203]. However, in a large study of >200 laryngeal and hypopharyngeal tumors, MGMT methylation was significantly associated with N0 status[190]. In that study, the same primers were used and a comparable MGMT methylation rate of 27% was found (also 27% in our study). How the higher methylation rates in pN0 cases affect the metastatic potential of OOSCC is not clear. Loss of the repair function of MGMT may increase the accumulation of mutations, especially in smoking-induced tumors, such as OOSCC. Because smoking is associated with higher methylation rates in general [214] and methylation of MGMT specifically[215], MGMT methylation might be a pseudo marker for smoking-induced tumors, rather than for HPV-associated tumors, which are more frequently pN+, according to some authors[216]. However, MGMT methylation was not associated with HPV status in our study (data not shown), nor in another study with more HPV-positive cases[217]. In our series, we show for the first time that in OOSCC, MGMT methylation is associated with a decreased expression in the invasive tumor front, but not in the tumor center (Figure 2.3). This is in line with the reported heterogeneity of methylation markers and their associated proteins [204], [207] and with the fact that methylation is associated with heterogeneous rather than with overall low expression[218]. The negative predictive value (NPV) of the combined model of DAPK1 and MGMT methylation of 76% in the current study is even slightly better than the 72% found in a 696-gene expression signature[78]. However, a NPV of over 80% is needed to outperform current clinical nodal staging techniques [193], including sentinel lymph node biopsy[219]. Obviously, further validation of the methylation markers, especially on the

clinically most relevant subgroup of pT1-2cN0 cases, is needed. In the current study, both DAPK1 and MGMT were non-significant predictors in the pT1-2cN0 subgroup (n = 37; data not shown). Treatment of OOSCC patients using demethylating drugs may not be effective, as our study shows that demethylation of DAPK1 might be beneficial, but demethylation of MGMT might result in nodal disease. MSP is not a quantitative technique. Although quantitative MSP for DAPK1 and MGMT enables specific cut-off values, thus customizing sensitivity and specificity, MSP is a more suitable technique for assessing a set of markers because it is a quick, low-cost and sensitive technique, able to detect a single methylated allele in a background of 1,000 unmethylated alleles[138]. However, selecting and testing of various possible methylation markers proved to be an inefficient method to identify new predictive markers. To improve marker selection efficiency, genome-wide methods are needed[220]. In conclusion, we analyzed 28 candidate methylation markers for their predictive value for N status by MSP on a large clinical group of OOSCC. MGMT and DAPK1 were identified as predictors of nodal metastasis in OOSCC with a high predictive value and specificity and sensitivity comparable to other markers previously reported. In addition, we showed for the first time that MGMT methylation is associated with a decreased expression in the invasive tumor front. This confirms the predictive value of methylation markers and the biological impact of methylation on the metastatic potential of OOSCC. In the future, DAPK1 and MGMT might be included in a panel of methylation markers that aid the clinician in the assessment of the N status, improving patient diagnosis and treatment selection.

Supplemental table 2.1. Primer sequences and optimized MSP conditions.

Primer	Forward sequence (location relative to TSS)	Reverse sequence (location relative to TSS)	Tannealing (ºC)	[MgCl$_2$] (mM)	Expected product size (bp)
PPT2 U	5' TTTTATTGGTTTAAATGGATTGTTTT 3' (-42 - -17)	5' AAACTTTCTCTAACAACCACACAA 3' (+65 - +88)	60	3	131
PPT2 M	5' TTGGTTTAAATGGATCGTTTC 3' (-37 - -17)	5' AACTTTCTCTAACAACCGCG 3' (+68 - +87)	60	2	125
MAL2 U	5' GGGGTGTGTATGGAGATGTTT 3' (-226 - -206)	5' ACTCACAATCACATACACAAATACAACT 3' (-128 – -101)	60	2	126
MAL2 M	5' GGTGCGTATGGAGACGTTC 3' (-224 - -206)	5' GCGGATCACGTACACAAATACG 3' (-125 - -105)	60	2	120
SRP19 U	5' TTGTTAGAGATTAGAGATTTTGGTGTT 3' (-10 - +16)	5' ACCCAACTCTAAATTTCCAAAAC 3' (+105 - +127)	60	2	138
SRP19 M	5' GAGATTAGAGATTTTGGCGTC 3' (-4 - +16)	5' CCGACTCTAAATTTCCGAAAC 3' (+105 - +125)	60	2	130
CD40 U	5' TTTGTTTTTTTGATAGGTGGATTGT 3' (-49 - -25)	5' CCACTAAACACCCAAACAAAACC 3' (+43 - +65)	60	2	115
CD40 M	5' TTTTTCGATAGGTGGATCGGC 3' (-44 - -25)	5' ACTAAACGCCCGAACGAA 3' (+46 - +63)	60	2	108
DNAH11 U	5' TTTTTGTTTAATTTTGGGGGTT 3' (+85 - +106)	5' ACCCACACATCTTAAACAAAACTC 3' (+184 - +207)	60	3	123
DNAH11 M	5' CGTTTAATTTCGGGGGTC 3' (+89 - +106)	5' CGGCGTCTTAAACGAAACTC 3' (+184 - +202)	60	3	114
CLEC16A U	5' GTATTTTTTGTTTGTGTTATTGTTGT 3' (+138 - +163)	5' AAACAACCAAACATATCAACAACC 3' (+221 - +244)	60	1	107
CLEC16A M	5' TCGTTTGTGTTATCGTCGC 3' (+145 - +163)	5' AACGACCAAACATATCGACG 3' (+224 - +243)	60	1.5	99
ODCP U	5' TGGGGTTATATAAGTTAGTGGTGGGT 3' (+5 - +30)	5' AAAATAAATCAAATCCTCAACACCT 3' (+115 - +139)	60	1	135
ODCP M	5' TATATAAGTTAGCGGCGGGC 3' (+11 - +30)	5' ATAAATCGAATCCTCGACGC 3' (+117 - +136)	60	1	126
NOL12 U	5' TTGGTGTGATGTTAAAGTGTGTGTTT 3' (-33 - -8)	5' AACTAAAAACAAACCTCAACCACC 3' (+97 - +120)	60	1.5	154
NOL12 M	5' CGACGTTAAAGTGTGCGTTC 3' (-26 - -8)	5' CTAAAAACGAACCTCGACCG 3' (+99 - +118)	60	1.5	146
MAPK13 U	5' GAATGTAGTTGTTATGTTGGGGTT 3' (+56 - +79)	5' ACTACCAACATACATCAAAAACACATA 3' (+172 - +198)	60	2	143
MAPK13 M	5' GTAGTCGTTACGTTGGGGTC 3' (+60 - +79)	5' ACGTACGTCGAAAACACGTA 3' (+172 - +191)	60	2	132
GRK6 U	5' TTGTGTTGATTGTTATTTCGTTTTT 3' (+76 - +99)	5' ACCATATTCACTACAATATTCTCAAA 3' (+167 - +192)	60	2.5	117
GRK6 M	5' TCGATCGTTATTCGGTTTC 3' (+81 - +99)	5' TTCGCTACGATATTCTCGAA 3' (+167 - +192)	60	2.5	106
VSNL1 U	5' GTGTGGTGAGTTTGGGTAATTT 3' (-173 - -152)	5' AAAAACTCAAAATTTCCACAAATAAAT 3' (-49 - -23)	60	1.5	151
VSNL1 M	5' GGCGAGTTCGGGTAATTC 3' (-169 - -152)	5' CGAAATTTCCGCGAATAAAT 3' (-49 - -30)	60	1.5	140
BDH1 U	5' GAGATGGTTGTATTGGGAGTTTAGT 3' (-132 - -108)	5' AACAAAACTCACAACAACATAACTATCA 3' (-36 - -9)	60	1	124
BDH1 M	5' GGTCGTATCGGGAGTTTAGC 3' (-127 - -108)	5' TCACGACGACGTAACTATCG 3' (-36 - -17)	60	2	111
RPL37A U	5' ATTTTTTAGGAGGTTGTTTGAAAAT 3' (-68 - -44)	5' CACAATACACAAACACAATATTAAACAA 3' (+13 - +40)	60	2.5	109
RPL37A M	5' TTTTAGGAGGTCGTTTGAAAAC 3' (-65 - -44)	5' GCAAACGCGATATTAAACGA 3' (+13 - +32)	60	3	98
GSTA4 U	5' GTGAGGTTGTTTTGGAGTTTT 3' (+14 - +34)	5' CACTCAAAAACCTAAAACCACA 3' (+103 - +124)	60	2	111
GSTA4 M	5' AGGTCGTTTCGGAGTTTC 3' (+17 - +34)	5' CACTCGAAAACCTAAAACCG 3' (+105 - +124)	60	2	108
BTG2 U	5' TAGAGTTTGAGTAGTGGTTAGGGTAAT 3' (-17 - +9)	5' ACAACAATCTCCAAAAACATATCAA 3' (+91 - +115)	60	2	133
BTG2 M	5' TCGAGTAGCGGTTAGGGTAAC 3' (-11 - +9)	5' CGATCTCCGAAAACATATCG 3' (+92 - +111)	60	2	123
E2F5 U	5' GGAGTTGATTTGGTAGGTGGTT 3' (-44 - -23)	5' CACCTACTAACCCAAACTCACAA 3' (+54 - +76)	60	2	121
E2F5 M	5' GTCGATTCGGTAGGTGGTC 3' (-41 - -23)	5' CTACTAACCCGAACTCGCG 3' (+55 - +73)	60	2	115
SSH2 U	5' GATGGTTTTGGTTATGGTTTAGT 3' (+228 - +250)	5' CCACTAAAACAAAACAAAACCAC 3' (+333 - +355)	59	2.5	128
SSH2 M	5' GGTTTTGGTTACGGTTTAGC 3' (+231 - +250)	5' TAAAACAAAACGAAACCGC 3' (+333 - +351)	60	1.5	121
PARVB U	5' GGGATTTGTTTGGTGGTGTTT 3' (+207 - +227)	5' AATCCCAACCATTATTTACAAATCC 3' (+333 - +357)	60	2	151
PARVB M	5' ATTTGTTCGGCGGTGTTC 3' (+210 - +227)	5' TCCCGACCGTTATTTACGAA 3' (+336 - +355)	60	3	146
GJB6 U	5' TTTTTATTTGAAATTTGATGAGAGTTT 3' (+78 - +104)	5' CCTACTCTACAACCAACAACCC 3' (+182 - +203)	60	1.5	126
GJB6 M	5' TCGAAATTCGACGAGAGTTC 3' (+85 - +104)	5' CCTACTCTACGACCGACGAC 3' (+184 - +203)	60	1.5	119
OCLN U	5' GGTTTTATTTGAAGTAGGTGGAGTATT 3' (+25 - +51)	5' CAACATTACAACCCAAAAAAACAA 3' (+124 - +146)	60	1.5	122
OCLN M	5' ATTCGAAGTAGGCCGGAGTATC 3' (+31 - +51)	5' CGTTACGACCCGAAAAAC 3' (+126 - +143)	60	2.5	113
TJP1 U	5' GTGTTGGTTGAGTTAGTGGATGTT 3' (+54 - +77)	5' CACCCATAACCTCCCAACATCT 3' (+136 - +157)	60	1.5	104
TJP1 M	5' GGTTGAGTTAGCGGACGTC 3' (+59 - +77)	5' CGTAACCTCCCGACGTCT 3' (+136 - +153)	60	2	95
CD44 U	5' TGTTTGGGTGTGTTTTTTGTTT 3' (+210 - +231)	5' ATAACAAACCAAACCTAACAAAAA 3' (+324 - +347)	60	1.5	138
CD44 M	5' TTGGGTGTGTTTTTCGTTC 3' (+213 - +231)	5' AACGAACCGAACCTAACAAA 3' (+326 - +345)	60	1.5	133
MLH1 U	5' AGGTTATGGGTAAGTTGTTTTGATG 3' (-539 - -515)	5' CCACTACAAACTAAACACAAATACTACAA 3' (-468 - -439)	60	1.5	101
MLH1 M	5' TACGGGTAAGTCGTTTTGACG 3' (-535 - -515)	5' ACGAAACTAAACACGAATACTACGA 3' (-468 - -444)	60	1.5	90
MGMT U[1]	5' TTTGTGTTTTGATGTTTGTAGGTTTTTTGT 3' (+57 – 85)	5' AACTCCACACTCTTCCAAAAACAAAACA 3' (+122 - +149)	60	1.5	93
MGMT M[1]	5' TTTCGACGTTCGTAGGTTTTCGC 3' (+63 - +85)	5' GCACTCTTCCGAAAACGAAACG 3' (+122 - +143)	60	1.5	81
CDKN2A U[2]	5' TTATTAGAGGGTGGGGTGGATTGT 3' (+227 - +250)	5' CAACCCCAAACCACAACCATAA 3' (+356 - +377)	60	1.5	151
CDKN2A M[2]	5' TTATTAGAGGGTGGGGCGGATCGC 3' (+227 - +250)	5' GACCCCGAACCGCGACCGTAA 3' (+356 - +376)	60	1.5	150
DAPK1 U[3]*	5' GGAGGATAGTTGGATTGAGTTAATGTT 3' (+201 - +227)	5' CCCTCCCAAACACCAACC 3' (+284 - +301)	60	1.5	101
DAPK1 M[3]	5' GGATAGTCGGATCGAGTTAACGTC 3' (+204 - +227)	5' CCCTCCCAAACGCCGA 3' (+286 - +301)	60	1.5	98

[1] [195], [2] [194], [3] [138]. *Primer was adapted from reference.

TSS: Transcription Start Site
M: Methylated
U: Unmethylated

Chapter 3

Identification and validation of WISP1 as an epigenetic regulator of metastasis in oral squamous cell carcinoma

M.J.A.M. Clausen[1,2], L.J. Melchers[1,2,] M.F. Mastik[1], L. Slagter-Menkema[1, 3],
H.J.M. Groen[4], B.F.A.M. van der Laan[3], W. van Criekinge[6], T. de Meyer[6], S. Denil[6],
G.B. A. Wisman[5], J.L.N. Roodenburg[2], E. Schuuring[1]

[1] Departments of Pathology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[2] Departments of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[3] Departments of Otorhinolaryngology/Head & Neck Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[4] Departments of Pulmonology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[5] Departments of Gynecologic Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[6] Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium

ABSTRACT

Lymph node (LN) metastasis is the most important prognostic factor in oral squamous cell carcinoma (OSCC) patients. However, in approximately one third of OSCC patients' nodal metastases remain undetected, and thus are not adequately treated. Therefore, clinical assessment of LN metastasis needs to be improved. The purpose of this study is to identify DNA methylation biomarkers to predict LN metastases in OSCC.

Method:

Genome wide methylation assessment was performed on six OSCC with (N+) and six without LN metastases (N0). Differentially methylated sequences were selected based on the likelihood of differential methylation and validated using an independent OSCC cohort as well as OSCC from The Cancer Genome Atlas (TCGA). Expression of WISP1 using immunohistochemistry was analyzed on a large OSCC cohort (n=204).

Results:

MethylCap-Seq analysis revealed 268 differentially methylated markers. WISP1 was the highest-ranking annotated gene that showed hypomethylation in the N+ group. Bisulfite pyrosequencing confirmed significant hypomethylation within the WISP1 promoter region in N+ OSCC (p = 0.03) and showed an association between WISP1 hypomethylation and high WISP1 expression (p = 0.01). Both these results were confirmed using 148 OSCC retrieved from the TCGA database. In a large OSCC cohort high WISP1 expression was associated with LN metastasis (p = 0.05), disease-specific survival (p = 0.022) and regional disease-free survival (p = 0.027).

Conclusion:

These data suggest that WISP1 expression is regulated by DNA methylation and that WISP1 hypomethylation contributes to LN metastasis in OSCC. WISP1 protein and WISP1 DNA methylation levels are potential biomarkers for identifying OSCC patients who require neck dissection treatment.

INTRODUCTION

Oral Squamous Cell Carcinoma (OSCC) is the most common subtype of Head and Neck Squamous Cell Carcinomas (HNSCC). These OSCC are characterized by a low 5-year survival of 48% [1]. OSCC tends to metastasize to the lymph nodes (LN) before distant metastasis occurs. The presence of these LN metastases has a major impact on OSCC patient survival and is therefore the most important prognostic factor [221]–[223]. Hence, status of the lymph node (N-status) significantly contributes to the selection of current treatment options. Patients with clinically positive N-status (cN+) are generally treated by a neck dissection, which causes mutilation and severe co-morbidity. Clinically negative N-status (cN0) patients are subjected to either an elective neck dissection or a "watchful waiting policy". Therefore, accurate assessment for the presence of lymph node metastases is essential for appropriate treatment management. However, clinical N-status assessment by palpation is inaccurate with a rate of occult LN metastases of 17-30% [49], [224]. In addition, current imaging modalities to determine clinical N-status in palpation-negative necks have a sensitivity of only 60-70% [18], [170]. Thus, to avoid under- and overtreatment, novel prognostic tumor markers are needed.

DNA methylation has been established as important regulator of tumorigenesis and affects cancer progression, metastatic potential, therapy response and patient survival (reviewed in [225]). Through the physical alteration of the cytosine nucleotide by methylation, changes occur in the DNA structure influencing gene transcription (reviewed in [94]). Patterns in genome-wide DNA methylation are cell specific, heritable and influence phenotypes allowing for the prediction of biological behavior of cancer cells (reviewed in [97], [226]). In the clinic, DNA methylation of genes, such as MGMT [227], can be used to predict treatment response, clinical outcome and clinical tumor characteristics including LN metastasis [228]. For instance, LN metastasis in HNSCC has been shown to be associated with methylation of TWIST1 [229], IGF2 [230], CDKN2A, MGMT, MLH1 and DAPK [231]. However, no improvement in the clinical assessment of N-status in OSCC has been made so far with these markers.

To identify new differentially methylated genes and pathways, various global methylation screening approaches have been reported including Infinium BeadArrays and WGBS (reviewed in [131], [136]). More recently, MethylCap-Seq was reported as an innovative new high-resolution technology to uncover DNA-methylation [232] in a genome-wide manner (reviewed in [136]). The approach is based on the identification of DNA CpG methylation by capturing DNA fragments with the Methyl Binding Domain of proteins as MeCP2 and MBD2 followed by next-generation nucleotide sequence analysis on e.g. an Illumina GA platform. Recently, we applied this assay to assess global methylation patterns in OSCC patients with histologically confirmed metastasis positive N-status (pN+) and compared those to OSCC with histopathologically confirmed negative N-status (pN0) or cN0 status for at least two year (Clausen et al., manuscript in preparation). In total 268 regions of differential methylation called Methylation Cores (MC) were identified as potential predictors of N-status. The majority of MC were hypermethylated in pN+ OSCC and only few hypomethylated loci were identified (17%) (Clausen et al., manuscript in preparation). In the present paper, we report on the detailed characterization of the WISP1 (WNT1-inducible-signaling pathway protein 1) gene that we identified as the most significantly hypomethylated annotated gene in

3

pN+ OSCC and which might act as a new potential diagnostic marker to identify OSCC with LN metastasis. Expression of WISP1 in malignancies other than OSCC was previously reported, but here we describe that high WISP1 expression in a large cohort of primary OSCC is a predictor for the presence of lymph node metastasis. In addition, to our knowledge we describe for the first time that WISP1 promoter methylation levels are associated with pN+ status and WISP1 expression levels in OSCC.

MATERIALS AND METHODS

Patient selection

All patients and carcinomas used in this study were selected from a large cohort described previously [18], [49]. Briefly, Netherlands Cancer Registry records and patient characteristics were collected of all patients with OSCC treated in the University Medical Center Groningen (UMCG) between 1997 and 2008. All patients had no history of prior treatment for HNSCC or other tumors. The histopathological diagnoses were revised for all cases by an experienced head and neck pathologist using the original haematoxylin and eosin (HE)-slides of the formalin-fixed, paraffin embedded (FFPE) tissue blocks. Patient and tumor characteristics are presented in Table 3.1. All cases were treated by primary tumor resection and a neck dissection. To ensure that pN0 cases do not contain occult LN metastases, we included only tumors with histologically confirmed pN0 status and cN0 status after > 2 years of follow-up. For the immunohistochemical study, we used 227 OSCC assembled in triplicate on in total five tissue-microarrays (TMA) as described previously [49]. Each TMA contains seven different normal tissues used as controls as well as for TMA orientation and recognition. All OSCC used in this study were tested for active high-risk HPV16 according to the algorithm of Smeets and colleagues [233]. In total five patients tested positive [49]. For the validation of clinical outcome, only patients with HPV-negative OSCC were included. For the MethylCap-Seq study, six pN+ cases and six pN0 cases were selected from the total cohort. Cases were matched for age and primary tumor site. Leukocytes were acquired from healthy women and served as controls for endogenous methylation and methylation background estimation of tumor samples [234], [235]. All patient tissues were coded. This study was performed according to the Code of Conduct for proper secondary use of human tissue in the Netherlands (www.federa.org), as well as to the relevant institutional and national guidelines.

Table 3.1. Patients characteristics of the UMCG and TCGA patient cohorts

N (%)	UMCG	TCGA
Total tumours	204 (100)	147 (100)
Total patients	204 (100)	147 (100)
Gender		
Male	127 (62)	100 (68)
Female	77 (38)	47 (32)
Age at diagnosis (yrs)		
Median	63	61
Range	35-94	19-87
Site		
Tongue	58 (28)	80 (54)
Floor of mouth	69 (34)	26 (18)
Cheek mucosa	7 (3)	0 (0)
Gum	24 (12)	41 (28)
Retromolar area	14 (7)	0 (0)
Oropharynx	27 (13)	0 (0)
Other	5 (3)	0 (0)
cN status		
0	125 (61)	73 (50)
+	79 (39)	73 (50)
pT status		
01-02	129 (63)	17 (12)
03-04	75 (37)	42 (28)
pN status		
pN0	104 (51)	61 (42)
pN+	100 (49)	86 (58)
Extranodal spread (only pN+)		
No	64 (57)	40 (47)
Yes	48 (43)	26 (30)
Perineural invasion		
No	133 (72)	51 (35)
Yes	51 (28)	69 (47)
Lymphovascular invasion		
No	144 (85)	83 (57)
Yes	26 (15)	31 (21)
Histological differentiation		
Well	50 (23)	18 (12)
Moderate	130 (61)	102 (69)
Poor	34 (16)	27 (18)
HPV16 status		
Negative	191 (97)	28 (97)
Positive	5 (3)	1 (3)
Infiltration depth (mm) (n = 181)		
Median	9.3	Not available
Range	0.07 – 40	Not available

3

DNA isolation

DNA isolation was performed as reported previously [236]. Briefly, two 10 μm thick FFPE sections were deparaffinized in xylene and incubated overnight in 300 μl 1% SDS-proteinase K at 60ºC. Subsequently, DNA was extracted using phenol-chloroform extraction and ethanol precipitation. DNA pellets were washed with 70% ethanol, dissolved in 50 μl TE^{-4} (10 mM Tris/HCL; 0.1 mM EDTA, pH 8.0) and stored at 4ºC. Genomic DNA was amplified in a multiplex PCR according to the BIOMED-2 protocol to check the DNA's structural integrity [196]. Only cases with products ≥200 bp were included for further analyses. For the MethylCap-Seq samples, DNA was extracted from snap frozen material. Then DNA quantity was measured using Quant-iT™ PicoGreen® dsDNA Assay Kit according to manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). For OSCC in the validation cohort, DNA concentrations and 260/280 ratios were measured using the Nanodrop ND-1000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA). A 260/280 ratio of >1.8 was required for all samples. A 3 μm thick section was HE-stained to check for tumor load. Only cases with at least 60% tumor cells were included in this study.

MethylCap-Seq analysis

MethylCap-Seq was performed to assess genome-wide methylation of 12 OSCC and two pools of leukocytes by capturing fragmented DNA with the Methyl Binding Domain protein MeCP2 followed by paired-end next generation sequencing on the Illumina GA II as reported previously [151], [232]. In summary, 500 ng DNA was fragmented using Covaris S2 (Covaris, Woburn, MA, USA) and methylated DNA fragments were captured with the MethylCap kit (Diagenode, Belgium) according to the manufacturer's protocol. Subsequently, the captured fragments were paired-end-sequenced using the Illumina Genome Analyzer II (Illumina, San Diego, CA, USA). The sequenced paired-ends were mapped to the human reference genome (NCBI build 37.3) using BOWTIE software [237]. Reads were included when the paired-end fragments were mapped to a unique locus and the distance between paired ends after mapping was within 400 bp. Exactly overlapping reads were discarded because these reads are most likely amplifications of the same captured DNA fragment. Mapped reads were summarized using the "Map of the Human Methylome", an in house developed overview of possibly methylated regions, called "Methylation Cores" [168], [382].

All Methylation Cores (MCs) located 2000 bp upstream to 500 bp downstream of the Transcription Start Site (TSS) or in the first exon of an Ensemble (v65) gene were statistically compared using R [238] with R-package Bayseq [239]. MC were ranked according the likelihood of differential methylation and we also calculated an approximate false discovery rate (FDR). The 5000 MC with the lowest FDR were used for further analysis. For all annotated MCs p-values were calculated using the two-sided independent student-t test. Subsequently, the following criteria were applied for further MCs selection: significant p-value (p <0.05); the lowest read count in the relatively hypermethylated group is equal or higher than the highest read count in the relatively hypomethylated group of the pN0 and pN+ OSCC.

Verification and validation of MethylCap-sequencing data of WISP1 by bisulfite pyrosequence analysis.

Methylation levels within the WISP1 gene promoter were determined using bisulfite pyrosequence analysis. Sodium bisulfite treatment of isolated genomic DNA (1 μg/sample) was performed according to the recommendations of the EZ DNA methylation kit (Zymo Research, Corp, Irvine, CA). The efficiency of the cytosine to uracil conversion by bisulfite treatment of each DNA sample was checked by Methylation Specific PCR (MSP) for beta-actin and DAPK as reported previously [139]. All primer sequences and PCR conditions are described in Table 3.2. Pyrosequencing primers for WISP1 (Table 3.2) were designed using Pyromark Assay design version 2.0.1.15 (Qiagen, Hilden, Germany). Bisulfite treated DNA was amplified using the Pyromark PCR kit according to the company's protocol (Qiagen). Each reaction was performed with 12.5 μl PCR master mix 2x, 200 nmol of the forward, 20 nmol of the reverse primer and 180 nmol of a universal biotinylated primer. The reverse primer contained a universal 23 bp DNA tag 5'-GACGGGACACCGCTGATCGTTTA-3' that is recognized by a biotinylated primer as described [140] The PCR was performed as following: 15 min 95°C, 50 cycles of (30 sec 94°C, 30 sec 56°C, 30 sec 72°C), 10 min 72°C. Purification of the PCR product was performed using the Q24 Vacuum Workstation (Qiagen) according to the manufacturer's protocol. The biotinylated PCR products were captured using 1 μl Streptavidin-coated Sepharose High Performace beads (GE Healthcare, Little Chalfont, UK). The immobilized products were washed with 70% alcohol, denatured with PyroMark Denaturation Solution (Qiagen) and washed with PyroMark Wash Buffer (Qiagen). The purified PCR product was then added to 25 μl PyroMark Annealing Buffer (Qiagen) containing 0.3 μM Sequening Primer specific for the WISP1 amplicon. Finally, pyrosequencing was performed using the Pyromark Q24 (Qiagen). Methylation percentages of all measured CpGs were analyzed using the provided Pyromark Q24 software version 2.0.6 (Qiagen). Average methylation of all measured CpG's as well as all individual CpG's were compared between groups. Pyrosequencing measurements considered failed by the Pyromark software were excluded. Incomplete bisulfite conversion threshold was 5%.

Leukocyte DNA from healthy volunteers was used as control for normal/endogeneous methylation levels, in vitro methylated (by SssI enzyme) leukocyte DNA as a positive control for hypermethylation and Whole Genome Amplified (WGA) leukocyte DNA using the illustra Ready-To-Go GenomiPhi HY DNA Amplification Kit (GE Healthcare) as a control for unmethylated DNA. All three controls were included in each bisulfite pyrosequencing run to check for differences between runs.

Immunohistochemistry

Immunohistochemical staining was performed as described previously [49]. Briefly, 3-μm thick sections of FFPE tumor tissue were deparaffinized and rehydrated. Antigen retrieval was performed by citrate buffer in a microwave oven for 15 min at 400 W as previously reported [240]. Endogenous peroxidase was blocked in a 0.3% H_2O_2 solution for 30 min at room temperature (RT). Slides were incubated overnight at 4°C with rabbit polyclonal antibody to human WISP1 (H-55: sc-25441) (Santa Cruz, CA, USA) diluted 1:50 in PBS [240] with 1% Bovine Serum Albumin. Subsequently, the sections were incubated with Envision+ (Dako, Glostrup, Denmark) horseradish peroxidise for 30 min at RT, developed with 3,3 -diaminobenzidine

solution (Dako) containing 0.03% H_2O_2 and counterstained with haematoxylin for 2 min. Fallopian tube was used as a positive control for WISP1 expression and pancreas, gallbladder and colon tissue were used as negative controls [241].

Staining intensity was semi-quantitatively scored as reported [240]. The percentage of tumor cells stained and the intensity of staining (0, no staining; 1, weak; 2, moderate; 3, strong). Each staining was scored by two observers independently. Discordant results were discussed until consensus was reached. ROC curve analysis was performed to determine the optimal cut-off between pN0 and pN+ OSCC. Cases with strong staining (3) in > 7.5 % of the tumor cells were considered to have high WISP1 expression.

Table 3.2. Primer Sequences and optimized PCR conditions

Primer	Sequence 5'-3'	Tannealing (°C)
ACTB MSP forward	TAGGGAGTATATAGGTTGGGGAAGTT	57
ACTB MSP reverse	AACACACAATAACAAACACAAATTCAC	57
DAPK1 MSP meth forward	GGATAGTCGGATCGAGTTAACGTC	60
DAPK1 MSP meth reverse	CCCTCCCAAACGCCGA	60
DAPK1 MSP unmeth forward	GGAGGATAGTTGGATTGAGTTAATGTT	60
DAPK1 MSP unmeth reverse	CCCTCCCAAACACCAACC	60
WISP1 pyroseq forward	TTAGTGGTAGTAGTGTAATAAGGGTATAG	54
WISP1 pyroseq reverse	GACGGGACACCGCTGATCGTTTAACTCAAATTACAACATCACCTTCATAAC	54
WISP1 pyroseq sequencing	GTGGGGATAGTTTTAGTATT	54

TCGA data analysis

All cases (n = 148) from The Cancer Genome Atlas (TCGA) database (the TCGA Research Network 2014) were selected with the following criteria; tumor located in either "Floor of Mouth", "Oral Cavity" or "Oral Tongue"; available pN-status and available Level 3 Methylation (Illumina Infinium 450k) data. Additional annotation for the Infinium 450k probe was acquired from the Gene Expression Omnibus (GEO) accession number GSE42409 including distance to TSS, associated CpG island and position [242]. All 450k probes associated with the WISP1 TSS were extracted for further analysis (n = 14). Subsequently, beta-values were quantile normalized by using R (version 3.0.3) to apply the normalizeBetweenArrays function from the R package preprocessCore from Bioconductor [243]. With R and the Lumi package the normalized WISP1 450k probe beta values were converted to M-values using the beta2m function and statistically compared between the pN0 and pN+ OSCC using the eBayes function of the Limma package [244].

Statistical Analysis

Statistical analysis was performed using SPSS version 22.0.1 software package. Associations between WISP1 expression and clinico-pathological characteristics were tested using the ^{2} test. The WISP1 IHC cut-off was optimized using a ROC-curve analysis. Survival was defined as the number of days between the first treatment and disease-specific death (DSS) or disease recurrence (DFS) and analyzed by Kaplan-Meier curves and log rank test. All tests were performed two-tailed and a p-value ≤ 0.05 was considered

statistically significant.

RESULTS

Identification of WISP1 as the most significantly hypomethylated gene in pN+ OSCC.

To assess the genome-wide methylation status of metastatic and non-metastatic OSCC, MethylCap-Seq was performed on six pN0 and six pN+ OSCC primary tumors. On average 10.2 million methylated DNA fragments were captured, sequenced and mapped to the human genome for each of these 12 OSCC DNA samples. The R package BaySeq was used to identify and rank all MC that were differentially methylated between pN0 and pN+ OSCC and located -2000 to 500 bp from a TSS or in the first exon of an Ensemble gene. Using the highest ranking 5000 MC with the lowest FDR, 1609 MC tested significantly different using a two-sided independent student-t test ($p ≤ 0.05$). From these 1609 MC, 355 MC were selected for which the lowest read count in the relatively hypermethylated group was higher or equal to the highest read count in the relatively hypomethylated group. Finally, for 268 MC there was an annotated function in the UniProtKB/Swiss-Prot database (supplemental Table 3.1).

The majority of the MC were hypermethylated in pN+ OSCC (84%, 226/268), only 42 of the 268 (16%) selected MC were hypomethylated in pN+ OSCC (supplemental Table 3.1). Table 3.3 shows that of the 15 highest ranking differentially methylated MC, 13 showed higher levels of methylation in the pN+ group, whereas SLC7A10 and WISP1 were higher in the pN0 OSCC. Of these two genes, WISP1 was the highest ranking hypomethylated annotated gene in pN+ in comparison to both pN0 OSCC as well as normal DNA (Table 3.3). The annotated WISP1 MC was located between position 134,202,288 and 134,202,631 on chromosome 8 (GRCh37/hg19), 681 to 1025 bp upstream of the WISP1 TSS [245] (Figure 3.1). According to the GSE42409 "Additional Annotation Information" of the Infinium 450k probes [242], this region contains a CpG island (chr8: 134,202,271 − 134,202,560 bp) (Figure 3.1) [246].

Verification of the association between WISP1 promoter methylation and lymph node status
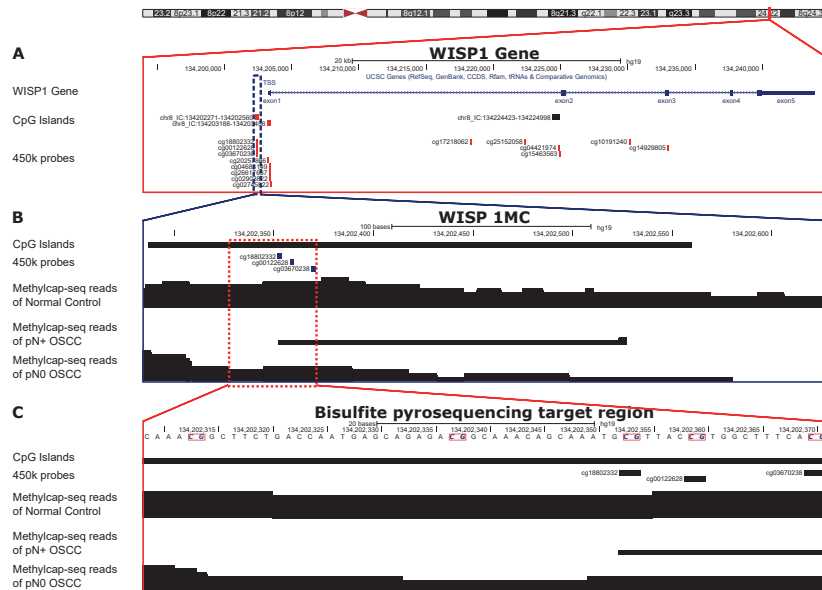
Figure 3.1. The WISP1 differentially methylated region annotated by MethylCap-Seq and the location of the bisulfite pyrosequenced CpGs. A) Schematic representation of the genomic region around the WISP1 gene (chr8:134,201,000 - 134,246,000) as extracted from the UCSC browser (GRCh37/hg19). The Transcription Start Site (TSS) is located at position 134,203,282. B) The WISP1 MC located 134,202,288 - 134,202,631, which is 681 to 1025 bp upstream of the WISP1 TSS, as retrieved from the Map of the Human Methylome [168], [382], the reads retrieved by MethylCap-seq analysis comparing 6 pN+ and 6 pN0 OSCC in this region, the known Infinium 450k probes and CpG Island location as retrieved from the GSE42409 database. C) The genomic region within the WISP1 MC as sequenced by bisulfite pyrosequencing, the reads retrieved by MethylCap-seq analysis comparing 6 pN+ and 6 pN0 OSCC in this region, the known Infinium 450k probes and CpG Island locations as retrieved from the GSE42409 database.

in OSCC

To validate whether the levels of DNA methylation of the selected WISP1 MC is associated with the lymph node status of OSCC, a bisulfite pyrosequencing assay was designed to quantify the methylation of five CpG sites in this WISP1 MC (Figure 3.1). Methylation levels were quantified on an independent validation cohort of 19 OSCC (Figure 3.2). Average (± SEM) methylation levels of all five CpG sites in the 10 pN0 (64 ± 3) and nine pN+ cases (49 ± 6) were both significantly lower than in leukocyte samples (86 ± 1) ($p < 0.01$). Moreover, the average methylation levels of the five CpG sites in pN+ cases were significantly lower than in the pN0 OSCC cases ($p = 0.02$) (Figure 3.2). When we compared the average methylation levels of each CpG site separately, the third CpG of the five pyrosequenced WISP1 MC CpGs showed the most significant difference between pN+ and pN0 ($p < 0.05$) (Figure 3.3 A). In summary, bisulfite pyrosequencing of the WISP1 promoter region in an independent cohort of 19 OSCC confirmed the lower methylation levels in pN+ OSCC as detected from the MethylCap-Seq marker discovery analysis.

Table 3.3. The fifteen highest annotated genes after statistical analysis of the enriched and sequenced reads of MethylCap-Seq.

Rank by False Discovery Rate	Gene name	Average distance to TSS (bp)	MC size (bp)	Average readcount pN0 OSCC	Average readcount pN+ OSCC	Hypermethylated in	False Discovery Rate	P-value student-test
1	ARHGEF4	0	327	1	5	pN+	0.26	0
2	U6	208	404	3	13	pN+	0.34	0.04
3	WISP1	-853	343	6	3	pN0	0.41	0
4	EMX2	-955	194	1	5	pN+	0.43	0
5	KCNIP1	0	298	1	5	pN+	0.44	0.02
6	SOBP	0	384	3	10	pN+	0.48	0.01
7	snoU13	70	286	2	6	pN+	0.48	0
8	GPD2	507	108	0	3	pN+	0.5	0
9	EDNRB	0	212	2	7	pN+	0.56	0.01
10	RAB13	-2092	270	3	8	pN+	0.57	0
11	SLC7A10	0	211	6	3	pN0	0.57	0
12	TAPT1	-2065	444	2	6	pN+	0.6	0
13	C1orf212	-1133	426	3	9	pN+	0.61	0
14	AL078621.5	-341	88	0	3	pN+	0.61	0
15	TMEM75	-591	138	1	4	pN+	0.63	0

All MC were ranked according to False Discovery rate (see supplemental table 3.1). After this ranking for each annotated gene the statistical difference between pN0 and pN+ were calculated by two-sided student t-test. Subsequently all MC were selected for which: the lowest read count in the relatively hypermethylated group is equal or higher than the highest read count in the relatively hypomethylated group of the pN0 and pN+ OSCC and there was an annotated description in the UniProtKB/Swiss-Prot database.
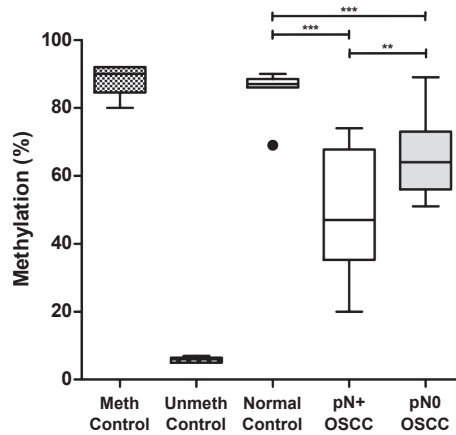
Figure 3.2. Methylation levels of the WISP1 MC are lower in pN+ OSCC compared to pN0 OSCC. The average methylation of the 5 WISP1 CpG sites were determined in 10 pN+ and 9 pN0 OSCC by bisulfite pyrosequence analysis. DNA from leukocytes from healthy controls was used as control for normal/endogenous methylation levels (Normal Control), in vitro methylated leukocyte DNA as a positive control for DNA methylation (Meth Control) and Whole Genome Amplified leukocyte DNA as an unmethylated DNA control (Unmeth Control). (* = p-value ≤ 0.1, ** = p-value ≤ 0.05, *** = p-value ≤ 0.01).
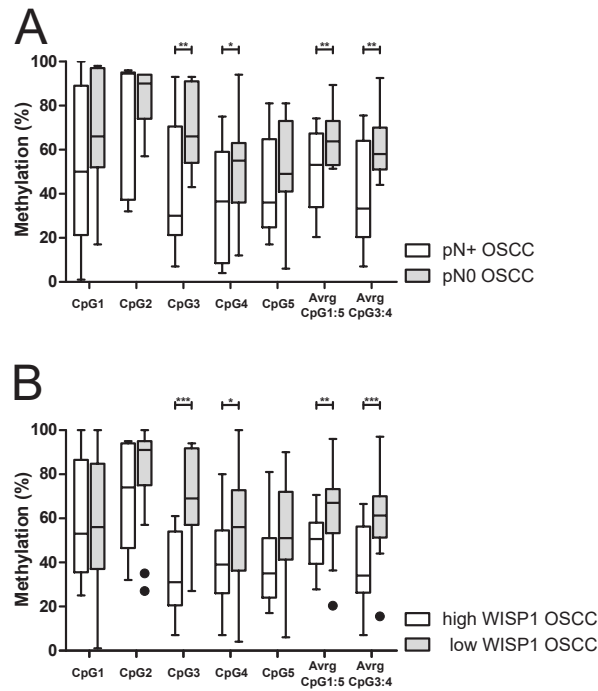
Figure 3.3. Methylation levels of the WISP1 promoter differ between pN0 and pN+ OSCC as well as between low and high WISP expressing OSCC. A) The methylation status of 5 CpG sites in the WISP1 MC was determined in 10 pN+ and 9 pN0 OSCC by bisulfite pyrosequencing. The methylation percentages of each individual CpG sites were compared between groups in addition to the average methylation percentage of all 5 CpG sites and the average of CpG sites 3 and 4. B) The methylation status of the same 5 CpG sites in the WISP1 MC was determined in 8 high WISP1 expressing OSCC and 16 low WISP1 expressing cases. The methylation percentages of each individual CpG sites were compared between groups in addition to the average methylation percentage of all 5 CpG sites and the average of CpG sites 3 and 4. (* = p-value ≤ 0.1, ** = p-value ≤ 0.05, *** = p-value ≤ 0.01).

To confirm the association between methylation of the WISP1 promoter region and lymph node status in OSCC, we used OSCC present in the TCGA database as independent validation cohort. For this purpose, 148 patients were selected with OSCC for which both the pN-status as well as well methylation data were available. In total 14 probes were identified that are associated with the WISP1 TSS, located -6996 to 7689 bp from the WISP1 TSS (Figure 3.1). 3 of these probes overlapped with the WISP1 MC and the CpGs in the WISP1 MC pyrosequenced region (Figure 3.1). All 14 probes were found to be significantly differentially methylated between pN0 (n = 61) and pN+ (n = 87) (supplemental table 3.2). Moreover, all three probes (cg18802332, cg00122628, cg03670238) overlapping with the bisulfite pyrosequencing assay were hypomethylated in the pN+ OSCC compared to the pN0 OSCC (supplemental table 3.2), in agreement with the bisulfite pyrosequence data. The analysis of the TCGA data confirms that hypomethylation of the WISP1 region is characteristic for pN+ OSCC.
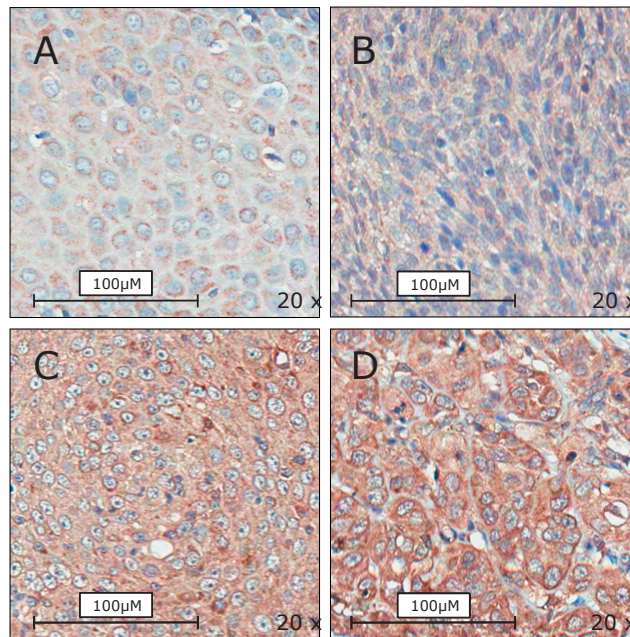
Figure 3.4. Representative examples of the different WISP1 intensity in 4 OSCC using immunohistochemistry. Tissues were scored for both immunoreactivity intensity ((A) no staining, (B) weak, (C) moderate; (D) strong staining and percentage of positive neoplastic cells. Cases with strong staining in >7.5 % of the neoplastic cells were considered as high WISP1 expressers and all other patterns as negative/low.

WISP1 promoter methylation in association with WISP1 expression

To validate whether WISP1 methylation levels are associated with WISP1 expression levels, bisulfite pyrosequencing of the WISP1 promoter region was performed on 21 OSCC cases with very high (score 2-3 in 100% of tumor cells) or low (negative or score 0-1) WISP1 protein expression as determined by immunohistochemistry. Average methylation levels of the five CpG sites were significantly lower ($p < 0.05$) in eight high WISP1 expressing OSCC compared to 16 low WISP1 expressing cases (Figure 3.3 B). Analysis of the separate CpG sites revealed that methylation levels of CpG3 ($p < 0.05$) and CpG3-4 ($p < 0.01$) were significantly lower in high WISP1 expressing OSCC ($p < 0.05$) (Figure 3.3 B).

High WISP1 expression in primary OSCC is a predictor for lymph node metastasis and clinical outcome

To validate whether WISP1 expression is associated with clinical outcome in OSCC, immunohistochemistry was performed on a cohort of 227 pretreatment biopsies of patients with OSCC of which clinic-pathological and follow-up data are available. Because 23 cores were lost during immunostaining or cores did not contain enough tumor cells anymore, WISP1 immunostaining could be assessed on 204 OSCC cases (see examples in Figure 3.4). High WISP1 expression was observed in 49 of 204 OSCC (24%) and found to be significantly associated with pN+ status ($p = 0.05$) (Table 3.4). Expression of WISP1 was also correlated with poor tumor differentiation ($p = 0.041$) but not with any of the other clinico-pathological features (Table 3.4).

Table 3.4. Clinical characteristics of OSCC patients with low or high WISP1 protein levels measured by IHC.

	Low WISP1 OSCC N (%)	High WISP1 OSCC N (%)	P-value
Total tumours	155 (76)	49 (24)	
Total patients	155 (76)	49 (24)	
Gender			
Male	97 (76)	30 (24)	0.864
Female	58 (75)	19 (25)	
Age at diagnosis (yrs)			
Median	63	62	0.199
Range	35-94	36-89	
Site			
Tongue	46 (79)	12 (21)	0.483
Gum	21 (88)	3 (12)	0.16
Retromolar area	11 (79)	3 (21)	0.814
Cheek mucosa	6 (86)	1 (14)	0.54
Floor of mouth	52 (75)	17 (25)	0.883
Oropharynx	15 (56)	12 (44)	0.056
Other	4 (80)	1 (20)	0.831
cN status			
0	99 (79)	26 (21)	0.176
+	56 (71)	23 (29)	
pT status			
01-02	98 (76)	31 (24)	0.996
03-04	57 (61)	37 (39)	
pN status			
0	85 (82)	19 (18)	0.05
+	70 (70)	30 (30)	
Extranodal spread (only pN+)			
No	37 (67)	18 (33)	0.511
Yes	33 (73)	12 (27)	
Perineural invasion			
No	106 (80)	27 (20)	0.188
Yes	36 (71)	15 (29)	
Lymphovascular invasion			
No	115 (80)	29 (20)	0.227
Yes	18 (69)	8 (31)	
Histological differentiation			
Well	41 (87)	6 (13)	0.041
Moderate or Poor	106 (73)	40 (27)	
HPV16 status			
Negative	146 (76)	45 (26)	0.061
Positive	2 (40)	3 (60)	
Infiltration depth (mm) (n = 181)			
Median	7	9	0.114
Range	0.07 - 30	2.10 - 40	
Infiltration depth (mm) (n = 200)			
<4 mm	25 (83)	5 (17)	0.318
>4 mm	113 (75)	38 (25)	

In order to assess the association between WISP1 expression and clinical outcome Kaplan-Meier log-rank analysis was performed on our OSCC cohort. To correct for different survival between HPV-positive and HPV-negative OSCC, only OSCC patients that were tested HPV-negative were included in this analysis. Kaplan-Meier log-rank analysis revealed a significant correlation between high WISP expression and worse disease-specific survival (Figure 3.5A, p = 0.022) as well as worse regional disease-free survival (Figure 3.5B, p = 0.027).
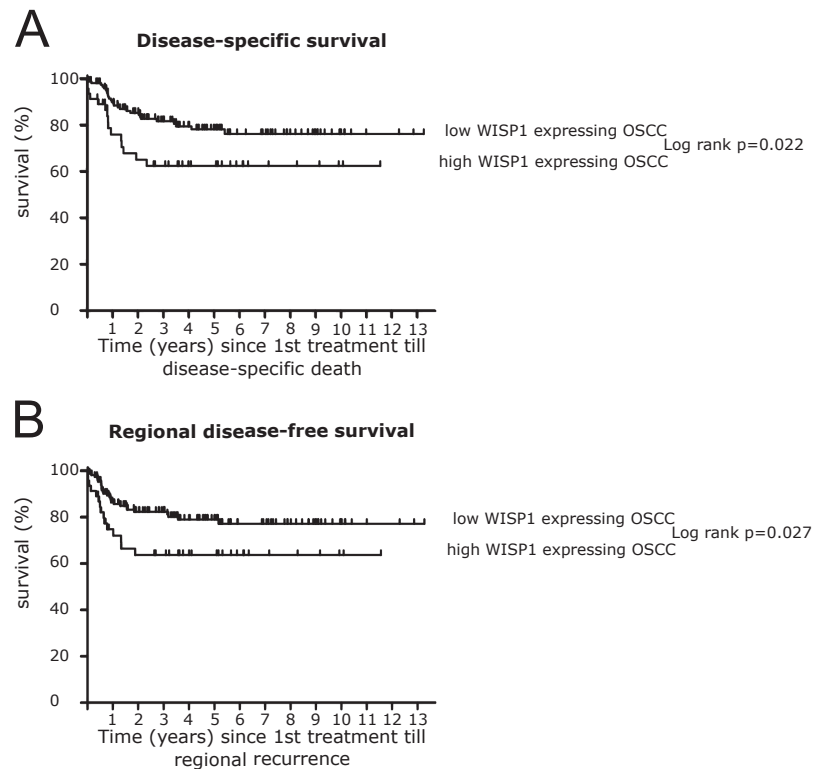
Figure 3.5. Kaplan–Meier curves of (A) Disease-specific survival stratified according to WISP1 expression for all HPV16 negative OSCC; and (B) Regional disease-free survival stratified according to WISP1 expression for all HPV16 negative OSCC. P-values of Log rank analysis.

DISCUSSION

During carcinogenesis the expression of cancer-associated genes is altered by changes in promoter methylation. These changes are thought to be an important event in tumor progression, therapy response, invasion and metastasis [225]. Therefore, the generation of a methylome specific for certain stages of disease in OSCC might be very helpful to gain new insight in OSCC carcinogenesis and to develop new tools for diagnosis and prognosis. Various studies have reported methylation markers in head and neck cancer, but today only few genes have been associated with lymph node status [247]–[249], clinical outcome and treatment response [250], [251]. Moreover, none of these markers showed sufficient predictive value for detecting the presence of lymph node metastases in early OSCC and clinical application [228], [252]. To identify new methylation markers associated with pN-status in OSCC, we performed a genome-wide methylation analysis using MethylCap-Seq as described recently [232]. This analysis revealed that of all differentially methylated markers, the WISP1 gene was the highest annotated hypomethylated gene in the OSCC with lymph node metastases (pN+) when compared to both OSCC without lymph node metastases (pN0) and leukocytes from healthy controls. In the present paper, we report on the identification and characterization of the association between DNA hypomethylation of the WISP1 gene and the presence of lymph node metastases in OSCC. We showed that WISP1 expression was correlated with DNA methylation of its promoter, and that decreased methylation levels were associated with increased WISP1 expression. In addition, using a large OSCC cohort (n = 204) high WISP1 expression was significantly associated with lymph node metastasis (p = 0.05) in 204 OSCC as well as with DFS (p = 0.022) and DSS (p = 0.027) in 191 HPV negative OSCC. Our data suggest that WISP1 is a new prognostic marker to predict OSCC metastasis to the regional lymph nodes.

WISP1 expression was observed before in other cancers [240], [253]–[259]. However, we describe for the first-time decreased DNA methylation levels of WISP1 in primary OSCC of patients with lymph node metastases. WISP1 DNA methylation was found 681 to 1025 bp upstream of the WISP1 TSS in a CpG island [193], [242]. We also showed that decreased WISP1 promoter DNA methylation was associated with high WISP1 expression. These findings suggest that hypomethylation of the WISP1 promoter in pN+ OSCC could be responsible for the up regulation of the WISP1 gene in metastatic OSCC. Using bisulfite pyrosequencing, we confirmed the increased WISP1 DNA methylation levels in OSCC cases with low WISP1 expression, whereas in pN+ OSCC decreased WISP1 DNA methylation levels were associated with a high WISP1 expression. Therefore, our data imply that WISP1 promoter methylation is as a potential new mechanism to regulate WISP1 expression in tumor cells. To confirm WISP1 DNA hypomethylation in pN+ OSCC on a larger independent cohort, we selected 148 OSCC cases present in the TCGA database from which also pN and methylation data were available. All 14 probes in the Infinium 450k platform located in the WISP1 promoter region were significantly differentially methylated between pN0 and pN+ OSCC. In particular, the three probes in the Infinium 450k platform located in the WISP1 MC identified by MethylCap-Seq analysis were also significantly hypomethylated in pN+ OSCC in agreement with our MethylCap-Seq and bisulfite pyrosequencing data.

It is now generally accepted that during cancer progression overall methylation decreases while gene promoter methylation increases especially of tumor suppressor genes [97]. Consequently, hypermethylation is thought to be correlated with tumor progression and metastasis [252]. In contrast to these tumor suppressor genes, we now found that WISP1 was hypomethylated in progressed disease as defined by the presence of lymph node-metastases. WISP1 DNA methylation levels were not only lower compared to pN0 OSCC but also to normal leucocytes suggesting that during progression of disease, WISP1 becomes actively hypomethylated. We also showed that the majority of the highest ranking differentially methylated MC other than WISP1 (see supplemental Table 3.2) were hypermethylated in pN+ OSCC (84%,226/268), whereas hypomethylation in the pN+ group was found only sporadically (16%, 42/268). Also in other studies, demethylation during cancer progression [260] in HNSCC has been reported for few other genes (MAGEB2, CSPG4 and ALK [247], [261], [262]), imprinted genes like Insuline-like growth factor 2 [263], repetitive elements Alu [264] and LINE-1 [265]. In addition, decreased DNA methylation levels associated with worse clinical outcome or response to chemoradiation has been described for a number of genes including TGM2, ASS1, TP73 and RASSF1A (reviewed in [174]). Based on increased DNA methylation levels of promoter sequences of tumor suppressor genes and their association with progressive disease in general[97], demethylating agents like 5-aza-2'-deoxycytidine might be a powerful new treatment strategy and in fact has already been implicated in e.g. acute myeloid leukemia[266]. However, our observations now strongly suggest that clinical treatment strategies using demethylating agents might be harmful for patients with OSCC, since upon treatment WISP1 might be de-methylated and re-expressed resulting in an increased metastatic potential.

WISP1, or WNT1-inducible-signaling pathway protein 1, belongs to the CCN protein family of six homologous, cysteine-rich secreted proteins induced by the Wnt-pathway. All CCN proteins are known to be involved in cancer related processes like cell adhesion, cell migration, proliferation and cell survival [267]. Moreover, WISP1 expression has been connected to the cancer promoting Notch pathway [256] and the Wnt-pathway [259] which are known to be abnormal in HNSCC [268] and is thought to contribute to lymph node metastasis in OSCC [269]. WISP1 has also been reported to influence P53 mediated-apoptosis by activating AKT [270] which is downstream in the ALK pathway [271].Subsequently, WISP1 over expression has been correlated with cancer progression, poor survival and metastasis in breast cancer [258], colorectal cancer [254], rectal cancer [257], esophageal squamous cell carcinoma [240] and NSCLC [253]. Interestingly, other published methylation markers that predict lymph node metastasis in OSCC are ALK, which is linked with WISP1 activity through AKT [247] and two genes which interact with WISP1 through the Wnt-pathway: RUNX3 and WIF1 [249]. In summary, our data suggest that WISP1 expression is regulated by DNA methylation and that WISP1 de-methylation contributes to lymph node metastasis in patients with OSCC. WISP1 DNA methylation and expression might contribute to a better selection of patients that might benefit from more optimal therapy. This will result in better patient survival and quality of life for OSCC patients. Therefore, WISP1 DNA methylation levels in primary OSCC could be used in deciding whether to treat patients shown to be cN0 by imaging with neck dissection.

3

Supplemental table 3.1. All 268 selected genes after statistical analysis of the enriched and sequenced reads of MethylCap-Seq.

All MC were ranked according to False Discovery rate. After this ranking for each annotated gene the statistical difference between pN0 and pN+ were calculated by two-sided student t-test. Subsequently all MC were selected for which: the lowest read count in the relatively hypermethylated group is equal or higher than the highest read count in the relatively hypomethylated group of the pN0 and pN+ OSCC and there was an annotated description in the UniProtKB/Swiss-Prot database.

Rank by False Discovery Rate	Gene name	Average distance to TSS (bp)	MC size (bp)	Average readcount pN0 OSCC	Average readcount pN+ OSCC	Hypermethylated in	False Discovery Rate	P-value student-test
1	ARHGEF4	0	327	1	5.3	pN+	0.26	0
2	U6	208	404	2.7	13.2	pN+	0.34	0.04
3	WISP1	-853	343	5.8	2.5	pN0	0.41	0
4	EMX2	-955	194	1	4.8	pN+	0.43	0
5	KCNIP1	0	298	0.5	4.7	pN+	0.44	0.02
6	SOBP	0	384	3	9.7	pN+	0.48	0.01
7	snoU13	70	286	1.5	6.2	pN+	0.48	0
8	GPD2	507	108	0	2.5	pN+	0.5	0
9	EDNRB	0	212	1.8	6.5	pN+	0.56	0.01
10	RAB13	-2092	270	2.5	7.8	pN+	0.57	0
11	SLC7A10	0	211	6.2	3	pN0	0.57	0
12	TAPT1	-2065	444	1.7	5.8	pN+	0.6	0
13	C1orf212	-1133	426	3.3	9.3	pN+	0.61	0
14	AL078621.5	-341	88	0.2	2.7	pN+	0.62	0
15	TMEM75	-591	138	0.8	4	pN+	0.63	0
16	hsa-mir-516a-2	-579	217	1.7	6	pN+	0.64	0.01
17	APH1A	0	57	2.2	0.3	pN0	0.64	0
18	Y_RNA	611	396	1	4.3	pN+	0.65	0.01
19	SMOC2	455	178	1	4.7	pN+	0.66	0.02
20	HHLA2	-1128	167	1.3	5	pN+	0.67	0.01
21	SGPL1	-1927	130	2	0.2	pN0	0.68	0.01
22	NFE2L1	-1806	86	0	2.2	pN+	0.68	0
23	TMEM117	-763	123	0.2	2.5	pN+	0.7	0
24	OR2B11	-404	188	1.7	5.5	pN+	0.71	0
25	EMR2	-1322	151	1	4.2	pN+	0.71	0.01
26	DLGAP4	-381	391	0.7	3.5	pN+	0.72	0.01
27	AC084125.1	-841	226	2.7	7.5	pN+	0.72	0
28	ACTA1	-756	273	0.7	3.7	pN+	0.73	0.02
29	DYRK1A	-1938	22	0.5	3	pN+	0.73	0
30	PHYHIP	-303	289	2.7	7.5	pN+	0.73	0
31	AAMP	0	539	4.7	11.5	pN+	0.74	0
32	PAX7	169	293	0.7	3.7	pN+	0.75	0.03
33	SEC31B	-684	302	2	6	pN+	0.76	0
34	TOR1A	-1925	127	1.8	0.2	pN0	0.76	0.01
35	EXOSC6	0	202	1.3	4.7	pN+	0.76	0
36	AC010645.1	-1721	296	3.3	8.7	pN+	0.76	0
37	TMEM110	0	169	2.5	7.2	pN+	0.77	0.01
38	DEDD	-1674	232	1.5	4.8	pN+	0.77	0
39	TINAG	0	81	0.5	3.2	pN+	0.78	0.02
40	EPM2AIP1	0	150	3	1.2	pN0	0.78	0.01
41	EGR4	611	636	1	4	pN+	0.79	0.01
42	TRABD	-1948	358	5.5	3	pN0	0.79	0.03
43	PCK1	0	47	1.5	5	pN+	0.79	0.01
44	ZDBF2	97	203	0	2	pN+	0.8	0
45	FAM107B	-370	123	1.3	4.5	pN+	0.81	0.01
46	RHOD	-620	72	0.2	2.2	pN+	0.81	0
47	GRIK5	90	199	1.7	5	pN+	0.81	0
48	LIN7B	-1501	197	0.5	2.8	pN+	0.81	0
49	snoU13	172	127	0.8	3.5	pN+	0.82	0.01
50	RPS6KB1	0	417	3.2	8	pN+	0.82	0
51	Y_RNA	-855	298	1.8	5.3	pN+	0.82	0
52	ASB3	-2017	164	1.7	5	pN+	0.83	0
53	AIM1	0	155	0.7	3.2	pN+	0.83	0.02
54	snoU13	404	293	2.7	7	pN+	0.83	0
55	VPS33A	-1192	233	1.8	5.3	pN+	0.83	0
56	SH3YL1	-274	227	2	5.7	pN+	0.83	0
57	ZDHHC8P	304	331	2	0.5	pN0	0.83	0.01

58	RAP1GAP	-384	536	0.2	2.2	pN+	0.83	0.01
59	DUS3L	-302	297	3.2	7.8	pN+	0.83	0
60	AB019438.11	591	221	0.8	3.3	pN+	0.84	0
61	SRP_euk_arch	-193	46	1	3.7	pN+	0.84	0.01
62	FAM65A	-1483	379	1.3	4.3	pN+	0.84	0
63	NDST3	-1270	345	3.8	9.2	pN+	0.84	0
64	LNX1	-1767	219	0.2	2.2	pN+	0.84	0.01
65	RNF212	239	528	3.7	8.8	pN+	0.84	0
66	SNORA74	-1681	241	3.5	1.7	pN0	0.84	0.02
67	MAPK11	-2018	386	5	2.8	pN0	0.84	0
68	Y_RNA	358	136	0.5	3	pN+	0.84	0.03
69	TRIP6	-1802	269	2.2	5.8	pN+	0.85	0.01
70	SLC7A7	629	477	1.2	4	pN+	0.85	0.01
71	FBRS	0	274	0.8	3.2	pN+	0.85	0
72	LRRC15	-707	26	2.2	0.7	pN0	0.85	0.02
73	DRGX	587	365	0.5	2.8	pN+	0.85	0.03
74	C19orf59	-1618	261	1.7	4.8	pN+	0.85	0
75	MYL2	-1098	122	0.5	2.7	pN+	0.85	0
76	PRX	-402	170	3	1.3	pN0	0.86	0.01
77	U4	37	233	4.5	2.5	pN0	0.86	0.01
78	DNAJC1	-931	180	1	3.7	pN+	0.86	0.02
79	NAPSB	488	67	0	1.8	pN+	0.86	0
80	PSG5	-1783	131	0.5	2.7	pN+	0.86	0
81	VNN1	-612	215	1.5	4.5	pN+	0.86	0
82	GTF2H3	-1896	98	2	5.5	pN+	0.86	0
83	DDA1	-1311	219	1.8	5.2	pN+	0.86	0
84	C16orf63	-1870	37	1.3	0	pN0	0.87	0
85	TJP2	50	288	1.5	4.5	pN+	0.87	0
86	SPATS2	-747	223	0.5	2.5	pN+	0.87	0
87	HS3ST5	540	115	0.2	2	pN+	0.87	0
88	AC135457.3	-1667	468	2	5.5	pN+	0.87	0.01
89	U6	-1218	313	1.3	4.3	pN+	0.87	0.03
90	METTL1	-1833	99	0.5	2.5	pN+	0.87	0
91	DKK4	433	164	4.7	2.7	pN0	0.87	0.03
92	U6	-1337	169	0.5	2.5	pN+	0.87	0
93	IZUMO1	269	496	2.7	6.7	pN+	0.87	0.01
94	HADHA	-1793	234	1.3	4.2	pN+	0.87	0
95	FOXRED2	-1075	343	3.7	8.5	pN+	0.87	0
96	AC120053.2	-891	118	1.7	4.8	pN+	0.87	0.01
97	DIRC3	19	263	2.7	6.7	pN+	0.88	0
98	Y_RNA	135	30	0.3	2.3	pN+	0.88	0.02
99	C17orf87	-1723	272	2.5	6.3	pN+	0.88	0
100	PKNOX1	0	281	2.2	5.7	pN+	0.88	0
101	RNF11	-1696	258	0.8	3.2	pN+	0.88	0
102	FTSJD1	-1907	103	1.3	0	pN0	0.88	0
103	PI16	-1349	177	1.8	5	pN+	0.88	0
104	SULT1A2	145	323	3.5	1.8	pN0	0.88	0
105	C20orf191	-2109	284	2	5.3	pN+	0.88	0.01
106	BPIL3	-1232	387	4.3	2.5	pN0	0.88	0.03
107	GTF3C2	-1891	232	1.8	5	pN+	0.89	0.01
108	SLC9A7	0	16	0.5	2.5	pN+	0.89	0
109	SCFD1	-1508	236	3	7.2	pN+	0.89	0.01
110	C1QB	-729	380	2.3	5.8	pN+	0.89	0
111	SRP_euk_arch	-1368	134	0.7	2.8	pN+	0.89	0.02
112	ZBTB34	-806	97	1.7	0.3	pN0	0.89	0.03
113	U6	-1268	88	0.5	2.5	pN+	0.89	0.01
114	C1orf198	-1498	76	1.5	4.3	pN+	0.89	0.02
115	SAMM50	0	358	2.5	6.2	pN+	0.89	0
116	MED31	-375	255	2	0.7	pN0	0.89	0.02
117	GLS	-1270	315	1.8	5	pN+	0.89	0.03
118	MYL2	-534	300	2.7	6.5	pN+	0.89	0.01
119	LY6H	-1414	85	2	0.7	pN0	0.9	0.02
120	RNH1	-1377	189	2.7	6.5	pN+	0.9	0.01
121	C6orf222	221	30	1.8	0.5	pN0	0.9	0.03
122	OR5AR1	0	1	1.8	0.5	pN0	0.9	0.05
123	SNORD115	-631	14	0.3	2.2	pN+	0.9	0.01
124	U6	-1756	311	1.7	4.5	pN+	0.9	0.01
125	NOL10	-755	401	2	5.2	pN+	0.9	0
126	U6atac	-103	413	1.8	4.8	pN+	0.9	0.01
127	snoU13	-344	214	1	3.3	pN+	0.9	0.01
128	VPRBP	0	123	1.7	4.5	pN+	0.9	0.05
129	TBX2	-1493	249	1.8	4.8	pN+	0.9	0
130	TBC1D12	-1236	314	3.3	7.7	pN+	0.9	0
131	SNORA32	-930	171	1	3.3	pN+	0.9	0.02

3

132	SETD4	0	234	1.3	4	pN+	0.9	0.03
133	SYNJ1	-1360	277	3.3	7.7	pN+	0.9	0
134	LSM8	-1377	279	2.3	5.7	pN+	0.9	0
135	ACIN1	-1207	409	3.2	7.3	pN+	0.9	0
136	INSM2	0	504	1.3	4	pN+	0.9	0.03
137	TADA1L	-1946	139	1	3.3	pN+	0.91	0.02
138	NCK1	0	32	0.8	2.8	pN+	0.91	0
139	IFT74	0	146	0.7	2.7	pN+	0.91	0
140	AC067852.1	-1723	141	1.2	3.5	pN+	0.91	0
141	FAM125A	-1983	260	1.7	4.5	pN+	0.91	0
142	IMPA1	-1295	258	4.2	9.2	pN+	0.91	0
143	OR7A10	-451	94	0.3	2.2	pN+	0.91	0.02
144	GABARAPL2	-757	225	2.5	1.2	pN0	0.91	0.01
145	ZNF490	-636	615	2.8	6.7	pN+	0.91	0.01
146	PROM1	-37	58	0.7	2.7	pN+	0.91	0.01
147	NEUROD6	-86	112	0.5	2.3	pN+	0.91	0
148	KLHDC8B	-2109	309	2.2	5.3	pN+	0.91	0
149	RP3-470B24.1	0	534	5.5	11.8	pN+	0.91	0
150	WDR33	-396	279	0.5	2.3	pN+	0.91	0
151	ZNF785	-1061	426	3.7	8.2	pN+	0.91	0.01
152	COBLL1	-1247	191	1.2	3.5	pN+	0.91	0.01
153	ALDOC	22	439	0.5	2.5	pN+	0.91	0.05
154	PARP12	301	397	1.3	3.8	pN+	0.91	0.01
155	OR4M1	522	228	1.2	3.5	pN+	0.91	0.01
156	L1TD1	-109	227	3.2	7.2	pN+	0.91	0
157	SRP_euk_arch	417	192	0.8	2.8	pN+	0.91	0
158	RP11-413E6.2	482	85	0.7	2.7	pN+	0.91	0.02
159	HUNK	-1704	375	3	6.8	pN+	0.92	0.01
160	ADAM8	-1894	57	0	1.7	pN+	0.92	0.01
161	TIMP4	-499	375	2.2	5.3	pN+	0.92	0
162	PIN1	-1045	559	2.7	6.2	pN+	0.92	0.01
163	CRIP3	367	1	0	1.5	pN+	0.92	0
164	AE000659.4	-1519	186	1.3	3.8	pN+	0.92	0.02
165	LMX1A	681	417	0.5	2.3	pN+	0.92	0.02
166	EIF2AK1	-539	336	2.7	6.2	pN+	0.92	0
167	CLEC14A	6	401	2.8	6.5	pN+	0.92	0
168	WIF1	-1274	1	0	1.7	pN+	0.92	0.04
169	FAM171A1	-1954	1	1.3	0.2	pN0	0.92	0
170	Y_RNA	-1717	291	1.8	4.7	pN+	0.92	0.02
171	LILRB1	0	115	1.8	0.7	pN0	0.93	0.01
172	GLI1	-1573	120	1.8	0.7	pN0	0.93	0.01
173	AL512662.7	244	41	1.8	0.7	pN0	0.93	0.01
174	OR5AR1	0	1	1.7	0.5	pN0	0.93	0.03
175	C2orf21	0	204	0.8	2.8	pN+	0.93	0
176	B4GALNT2	123	445	0.8	2.8	pN+	0.93	0
177	NLE1	-632	203	0.8	2.8	pN+	0.93	0.01
178	snoU13	-2125	382	2.2	5.2	pN+	0.93	0
179	ZNF24	-837	91	0.3	2	pN+	0.93	0.01
180	U5	524	599	0.7	2.3	pN+	0.93	0.01
181	MGMT	-1916	174	1.8	0.7	pN0	0.93	0.05
182	KATNAL2	-2019	151	0	1.5	pN+	0.93	0.01
183	Y_RNA	-1529	187	1.3	3.7	pN+	0.93	0
184	AGPAT3	0	286	2.8	6.3	pN+	0.93	0
185	ETF1	-1550	221	2	4.8	pN+	0.93	0.02
186	hsa-mir-320b-1	-706	1	1.2	0	pN0	0.93	0.02
187	B3GALNT2	-1540	127	2	4.8	pN+	0.93	0.01
188	ELAVL2	0	290	2.8	6.3	pN+	0.93	0.01
189	MBD1	-419	181	1.7	4.2	pN+	0.93	0.01
190	SLC45A4	-1753	100	3.7	7.8	pN+	0.93	0
191	TTLL4	0	373	5.8	12.2	pN+	0.93	0
192	KIAA1609	-1623	225	0.2	1.7	pN+	0.93	0
193	TACSTD2	561	641	1.2	3.3	pN+	0.93	0.01
194	EFCAB1	563	267	1.8	4.5	pN+	0.93	0.01
195	AMBN	-1442	273	2.5	5.7	pN+	0.93	0
196	SRP_euk_arch	240	74	1.2	0	pN0	0.94	0.02
197	METTL13	-972	39	0.2	1.7	pN+	0.94	0
198	RLBP1L2	69	361	3.5	7.5	pN+	0.94	0.02
199	BSDC1	-1670	160	0.5	2.2	pN+	0.94	0.01
200	C5orf46	-1265	228	0.5	2.2	pN+	0.94	0
201	HMGN1	-1629	1	1.2	0	pN0	0.94	0.03
202	GRHL2	85	533	0.5	2.2	pN+	0.94	0.01
203	TTC15	0	358	5.3	11	pN+	0.94	0
204	MT1F	-1401	143	0.2	1.7	pN+	0.94	0.01
205	ID4	0	57	1.2	0	pN0	0.94	0.03

206	C4orf50	-959	280	1.7	4.2	pN+	0.94	0.02
207	POTEH	78	142	0.5	2.2	pN+	0.94	0.03
208	B4GALNT4	545	468	3	6.5	pN+	0.94	0
209	ZAN	-296	299	1.5	3.8	pN+	0.94	0.01
210	C22orf9	-772	1	1.2	0	pN0	0.94	0.03
211	IGF2BP3	0	176	2	4.7	pN+	0.94	0
212	ATG4B	-1602	272	2.2	4.8	pN+	0.94	0.02
213	BCAN	-749	130	0.5	2.2	pN+	0.94	0.04
214	CRISPLD2	-501	283	1.3	3.5	pN+	0.94	0.02
215	C3orf54	-992	135	0.3	1.8	pN+	0.94	0
216	ZFP2	502	165	0.3	1.8	pN+	0.94	0
217	SRP_euk_arch	-1152	274	3	6.3	pN+	0.94	0.01
218	UBR7	-891	276	1.3	3.5	pN+	0.94	0
219	PPP4R2	-439	188	3	6.3	pN+	0.95	0.02
220	SLC10A6	-1106	68	0.3	1.8	pN+	0.95	0.01
221	AC092139.3	-482	418	1.8	4.3	pN+	0.95	0
222	GUCY2C	-1755	208	1.2	3.2	pN+	0.95	0.01
223	PPP4R4	-1345	80	2	1	pN0	0.95	0.02
224	CRIP3	360	1	0	1.3	pN+	0.95	0
225	C21orf32	633	302	1	2.8	pN+	0.95	0.01
226	DNA2	-1095	185	1.2	3.2	pN+	0.95	0.01
227	snoU13	-1534	70	1.2	3.2	pN+	0.95	0.01
228	TRIP6	-2067	260	3	6.3	pN+	0.95	0.01
229	U6	-59	136	1	2.8	pN+	0.95	0
230	CEL	-965	211	1.5	0.5	pN0	0.95	0.01
231	DEFB137	239	95	0.7	2.3	pN+	0.95	0.01
232	SH3RF3	0	293	3	6.3	pN+	0.95	0
233	GPD2	387	23	0.3	1.8	pN+	0.95	0.02
234	CRIP3	448	48	0	1.3	pN+	0.95	0
235	IGSF6	-1207	101	0.7	2.3	pN+	0.95	0.02
236	PPARD	-2016	146	1	2.8	pN+	0.95	0
237	C14orf167	-1893	19	0	1.3	pN+	0.95	0
238	HMX1	-362	310	1.8	0.8	pN0	0.95	0.02
239	SAR1B	-275	323	2.2	4.8	pN+	0.95	0
240	STK32B	-518	299	1.3	0.3	pN0	0.95	0.03
241	MCM2	0	260	2.2	4.8	pN+	0.95	0
242	RPP38	-601	156	1.3	0.3	pN0	0.95	0.03
243	RASL11B	-611	265	2.8	5.8	pN+	0.95	0.01
244	AL512624.1	-1134	255	5.8	11.7	pN+	0.95	0
245	Y_RNA	-790	227	1	2.7	pN+	0.95	0.02
246	TPM2	0	132	1	0	pN0	0.95	0
247	AC008537.3	0	84	1	0	pN0	0.95	0
248	NOC3L	-629	379	2	4.5	pN+	0.95	0.01
249	AC079354.2	0	415	3.3	6.8	pN+	0.95	0
250	PIF1	-1671	702	3	6.2	pN+	0.95	0.01
251	ZNF296	459	671	0.5	2	pN+	0.95	0.01
252	SFRS4	-714	269	0.8	2.5	pN+	0.95	0
253	THBS2	-1189	213	0.8	2.5	pN+	0.95	0
254	CASP12	-1748	176	1.3	3.3	pN+	0.95	0
255	DLGAP1	-481	439	4.2	8.3	pN+	0.95	0
256	SPANXN3	288	24	0	1.3	pN+	0.95	0.01
257	HK1	0	437	1.3	3.3	pN+	0.95	0.01
258	BBS2	-1332	96	0.5	2	pN+	0.95	0.01
259	snoU13	-2034	259	3.3	6.7	pN+	0.95	0.01
260	LRFN2	523	333	0.5	2	pN+	0.96	0.02
261	C1orf113	-66	419	1.2	3	pN+	0.96	0.01
262	GPR97	-1379	66	0.2	1.5	pN+	0.96	0
263	PDXK	-1433	252	0.5	2	pN+	0.96	0.02
264	PIAS4	-1474	275	3.8	7.7	pN+	0.96	0
265	RPS10L	-1511	253	0.2	1.5	pN+	0.96	0
266	SRP_euk_arch	-2258	683	1.8	4.2	pN+	0.96	0.01
267	SYPL2	425	157	0	1.3	pN+	0.96	0.02
268	TMEM98	-1560	16	0.5	2	pN+	0.96	0.05

3

Supplemental table 3.2. Statistical results and position information Infinium 450k for all WISP1 probes from the TCGA database.

Probe name	Chromosome	CpG location	CpG region	CpG Island location	Distance to WISP1 TSS	False Discovery Rate	Hypermethylated in
cg17218062	8	134218279	Open Sea	Not applicable	-6996	<0.00	pN+
cg25152058	8	134222322	CpG shore	Not applicable	-2953	<0.00	pN0
cg18802332	**8**	**134202353**	**CpG Island**	**134202271-134202560**	**-928**	**<0.00**	**pN0**
cg00122628	**8**	**134202359**	**CpG Island**	**134202271-134202560**	**-922**	**<0.00**	**pN0**
cg03670238	**8**	**134202370**	**CpG Island**	**134202271-134202560**	**-911**	**<0.00**	**pN0**
cg04421974	8	134224814	CpG Island	134224423-134224998	-461	<0.05	pN+
cg15463563	8	134224890	CpG Island	134224423-134224998	-385	<0.00	pN+
cg20257866	8	134203235	CpG Island	134203188-134203458	-46	<0.00	pN0
cg04683149	8	134203304	CpG Island	134203188-134203458	23	<0.00	pN+
cg26617637	8	134203339	CpG Island	134203188-134203458	58	<0.00	pN0
cg02903822	8	134203379	CpG Island	134203188-134203458	98	<0.05	pN0
cg02745822	8	134203435	CpG Island	134203188-134203458	154	<0.00	pN0
cg10191240	8	134230101	Open Sea	Not applicable	4826	<0.00	pN0
cg14929805	8	134232964	CpG Island	134232724-134233195	7689	<0.00	pN0

False Discovery Rate (Benjamini Heinberg) and direction of the differential methylation between the TCGA pN0 (n=61) and pN+ (n=87) OSCC for all WISP1 TSS associated Infinium 450k probes and all probe location info according to the GSE42409 database [242]. The probes that overlap with the MC identified by MethylCap-Seq are marked in bold (see Figure 3.1).

3

Chapter 4

RAB25 expression is epigenetically down-regulated in oral and oropharyngeal squamous cell carcinoma with lymph node metastasis

M.J.A.M. Clausen[1,2], L.J. Melchers[1,2], M.F. Mastik[1], L. Slagter-Menkema[1 3], H.J.M. Groen[4], B.F.A.M. van der Laan[3], W. van Criekinge[6], T. de Meyer[6], S. Denil[6], B. van der Vegt[1], G.B.A. Wisman[5], J.L.N. Roodenburg[2], E. Schuuring[1]

[1] Departments of Pathology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[2] Departments of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[3] Departments of Otorhinolaryngology/Head & Neck Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[4] Departments of Pulmonology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[5] Departments of Gynecologic Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[6] Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium

ABSTRACT

Oral and oropharyngeal Squamous Cell Carcinoma (OOSCC) have a low survival rate, mainly due to metastasis to the regional lymph nodes. For optimal treatment of these metastases a neck dissection is required but inaccurate detection methods result in under- and overtreatment. New DNA prognostic methylation biomarkers might improve lymph node metastases detection.

Materials and Methods:

To identify epigenetically regulated genes associated with lymph node metastases genome-wide methylation analysis was performed on 6 OOSCC with (pN+) and 6 OOSCC without lymph node (pN0) metastases and combined with a gene expression signature predictive for pN+ status in OOSCC. Selected genes were validated using an independent OOSCC cohort by immunohistochemistry and pyrosequencing, and on data retrieved from The Cancer Genome Atlas (TCGA).

Results:

A two-step statistical selection of differentially methylated sequences revealed 14 genes with increased methylation status and mRNA down-regulation in pN+ OOSCC. RAB25, a known tumor suppressor gene, was the highest-ranking gene in the discovery set. In the validation sets, both RAB25 mRNA (P = 0.015) and protein levels (P = 0.012) were lower in pN+ OOSCC. RAB25 mRNA levels were negatively correlated with RAB25 methylation levels (P < 0.001) but RAB25 protein expression was not.

Discussion:

Our data revealed that promoter methylation is a mechanism resulting in down-regulation of RAB25 expression in pN+ OOSCC and decreased expression is associated with lymph node metastasis. RAB25 methylation detection might contribute to lymph node metastasis diagnosis and serve as a potential new therapeutic target in OOSCC.

INTRODUCTION

Oral and oropharyngeal Squamous Cell Carcinoma (OOSCC) are the most common subtypes of Head and Neck Squamous Cell Carcinomas (HNSCC) and are characterized by an overall 5-year survival below 50% [1]. This low survival rate it greatly impacted by the presence of lymph node (LN) [224]. Patients with metastases in the regional lymph nodes of the neck have a 5-year survival half of those who do not have regional metastases [272], [273]. Therefore, for treatment decision making it is of importance to accurately detect the presence lymph node (LN) metastasis. Currently, diagnosis consists only of clinical examination and imaging, which are known to have low sensitivity and low specificity for LN metastasis detection [170], [274]–[276]. When LN metastases are detected a neck dissection is required but this surgical procedure is accompanied by neck and shoulder morbidity. As a result, under- and overtreatment of OOSCC patients occurs frequently [18], [170]. Currently, there is a major lack of appropriate clinical and tumor biomarkers that predict the presence of LN metastasis.

DNA methylation is a mechanism of epigenetic modification that impacts cellular phenotypes by regulating gene expression and is known to affect carcinogenesis by altering proliferation rates and DNA repair [94], [228]. As a result, DNA methylation screening has been used as a tool to predict clinical outcome and therapy response in cancer patients [94], [174]. Moreover, DNA methylation of several genes has been reported to have a predictive value for nodal metastasis in HNSCC including TWIST1 [229], IGF2 [230], [277], CDKN2A, MGMT, MLH1 and DAPK [231], [278]. However, these tumor markers have not resulted in improved clinical LN detection rate.

Recently, we have reported on the identification of new DNA methylation markers that predict LN status by MethylCap-Seq [279]. This combination of enrichment of methylated DNA fragments and next generation sequencing has been established as a true genome-wide assay compared to other DNA methylation screening techniques. Using a quantitative ranking of genomic loci by likelihood of differential methylation between OOSCC with metastasis negative LN (pN0) and OOSCC with metastasis positive LN (pN+), we identified WISP1 as a hypomethylation marker associated with pN+ OOSCC [279]. In the present study, we report on a new approach tailored towards identifying potentially epigenetically down-regulated genes in the metastatic OOSCC phenotype. Epigenetically down-regulated genes are more suitable for opening up new clinical options, because hypermethylation can be more easily detected in a unmethylated background and are more suited as therapeutic targets due to the emergence of epigenetic editing and demethylating agents [280].

For this purpose, we used 696 genes that were previously reported to be differentially expressed between 143 pN0 and 79 pN+ OOSCC. This gene signature has a validated negative predictive power of 89% for LN metastases [78], [83], [281]. We combined the expression levels of the genes in this predictive gene signature with DNA methylation data acquired by MethylCap-Seq analysis [279]. Using this approach, we found that 14 genes were simultaneously hypermethylated and down-regulated in pN+ OOSCC. In this manuscript, we report on the identification of RAB25 as the highest-ranking gene and the association between expression and methylation of RAB25 and the presence of LN metastases.

MATERIAL AND METHODS

Patient selection

All treatment naive patients with OOSCC who underwent a neck dissection for primary tumor resection resulting in free resection margins upon histopathological examination in the University Medical Center Groningen (UMCG) between 1997 and 2008 were selected. Pathological revision was performed of all the original hematoxylin and eosin (HE)-slides formalin-fixed, paraffin embedded (FFPE) tissue blocks. All pN0 tumors were histologically confirmed or had cN0 status with > 2 years of LN metastasis free follow-up. All patient and tumor characteristics are available in Supplemental Table 4.1. For the immunohistochemical study, 227 OOSCC tumors were used for 5 tissue-microarrays (TMA) in triplicate as described previously [49]. All TMA contained 7 different normal tissues that served as control. Human papilloma virus (HPV) status was tested by p16 immunohistochemistry followed by high-risk HPV PCR detection of OOSCC scored p16 IHC positive as previously reported [163]. For 197 OOSCC patients HPV16 status was available of which 5 patients were HPV16 positive. A total of 192 HPV negative patients (pN0 n= 102, pN+ n=90) were included for further analysis. For the MethylCap-Seq study, 6 pN+ and 6 pN0 tumors matched for age and primary tumor site were selected from the total cohort. Leukocytes were acquired from healthy women for endogenous methylation and methylation background estimation [235], [282]. This study was performed in accordance with the Code of Conduct for proper secondary use of human tissue in the Netherlands (www.federa.org), and relevant institutional and national guidelines.

DNA isolation

DNA isolation was performed as previously reported [279]. Briefly, two 10 µm thick FFPE sections were deparaffinized in xylene and incubated in 300 µl 1% SDS-proteinase K at 60º C overnight. DNA extraction was performed using phenol-chloroform and ethanol precipitation. The acquired DNA pellets were then washed with 70% ethanol, dissolved in 50 µl TE-4 (10 mM Tris/HCl; 0.1 mM EDTA, pH 8.0) and stored at 4ºC. To check the DNA's structural integrity, genomic DNA was amplified by multiplex PCR according to the BIOMED-2 protocol [196]. Cases with products ≥ 200 bp were selected for further analyses. DNA used for MethylCap-Seq samples was measured by Quant-iT™ PicoGreen® dsDNA Assay Kit according to manufacturer's protocol (Invitrogen). The DNA used for pyrosequencing was measured using the Nanodrop ND-1000 Spectrophotometer (Thermo Scientific). Only samples with an absorbance ratio at 260 nm and 280 nm of > 1.8 were selected for further testing. The number of tumor cell required for this study was set at 60% as estimated with HE-staining of 3 µm thickness.

MethylCap-Seq

MethylCap-Seq analysis was performed as reported previously [279]. Briefly, genome-wide methylation was assessed for 500 ng of DNA fragmented by Covaris S2 (Covaris) from 6 pN0 OOSCC, 6 pN+ OOSCC and 2 pools of leukocytes using methylated DNA enrichment by the methyl binding domain protein MeCP2 (MethylCap-kit, Diagenode) followed by paired-end next generation sequencing on the Illumina

GA II (Illumina). Subsequently, the enriched, captured and sequenced reads were mapped to the human reference genome (NCBI build 37.3) using the BOWTIE software [283]. Only the reads were included that were mapped to a unique locus. Exactly overlapping reads were excluded as identical reads are most likely the result of amplification of the same DNA fragment. Additionally, the mapped distanced between the paired-ends could not be longer than 400 bp. Finally, all the mapped reads were compared to the "Map of the Human Methylome" build 2 [168], [382]. This is an in house developed summary of all experimentally assessed genomic sites of potential differential methylation. These regions are called "Methylation Cores" (MC)

Table 4.1. Epigenetically down-regulated genes in pN+ OSCC.

Gene	Methylation Core data			DNA Methylation data				mRNA in pN+	Epigenetic reg.
	Chr	to TSS (bp)	size (bp)	P-Value	Hypermeth	Pos. Pred.	Neg. Pred.		mRNA & Meth corr.
RAB25	1	-108	233	0.02	pN+	100%	86%	-0.15	↓ Negative
COBLL1	2	-1247	191	0.02	pN+	100%	75%	-0.14	↓ Negative
GFRA1	10	-809	120	0.04	pN+	100%	67%	-0.11	↓ Negative
S100A9	1	490	125	0.04	pN+	100%	60%	-0.1	↓ Negative
LAMP3	3	0	284	0.05	pN+	80%	71%	-0.09	↓ Negative
ACTA1	1	-756	273	0.01	pN+	100%	67%	-0.08	↓ Negative
KRT17	17	-296	1	0.02	pN+	100%	55%	-0.08	↓ Negative
MAST4	5	-271	57	0.03	pN+	0%	50%	-0.06	↓ Negative
IL22RA1	1	114	229	0.05	pN+	75%	63%	-0.04	↓ Negative
BRUNOL4	18	-1543	24	0.03	pN+	100%	67%	-0.03	↓ Negative
NDUFA10	2	-1155	9	0.02	pN+	100%	55%	-0.01	↓ Negative
MALL	2	413	152	0.03	pN+	100%	55%	-0.01	↓ Negative
WDR13	X	0	54	0.05	pN+	100%	55%	-0.01	↓ Negative
H2AFY	5	-1065	90	0.03	pN+	100%	55%	-0.01	↓ Negative

All 14 potentially epigenetically down regulated genes in pN+ OSCC compared to pN0 OSCC after cross-reference of expression microarray and MethylCap-Seq data (see Figure 4.1). The positive and negative predictive value of the reads for pN+ status, associated hypermethylated, the read distribution between pN0 and pN+ OSCC and the predictive value of the methylation data are illustrated. P-value for the differential DNA methylation was calculated using the Mann-Whitney-U test. Positive and negative predictive value for the methylation status of all MC were calculated as follows: OOSCC with a read count of ≥ 3 reads were considered true positives and OOSCC with a count read < 3 were considered true negatives. Subsequently, the positive predictive value was then calculated as: (true positive pN+ OOSCC) / (true positive pN+ OOSCC + false positive pN0 OOSCC). Finally, the negative predictive value was calculated as: (true negative pN0 OOSCC) / (true negative pN0 OOSCC + false negative pN+ OOSCC).

To identify a candidate set of genomic regions differentially methylated between pN0 and pN+ OOSCC, all MC located 2000 bp upstream to 500 bp downstream of the transcription start site (TSS) or in the first exon of an Ensemble (v65) gene were statistically compared using R with R-package Bayseq [239]. The sequencing experiment proved to be underpowered in terms of sequencing depth and number of biological replicates, precluding any definite conclusions. Therefore, we focused on the identification of the most interesting set of putatively differentially methylated regions which could be validated in a subsequent setup. This led to the following two-step MC selection method. In the first step, the number of samples methylated was determined for both groups (pN+ and pN0). A sample was called

unmethylated if there were no reads and methylated if there were one or more reads. A Fisher exact test was performed to rank the MCs for differential number of methylated samples between both groups. Ties in p-values, due to the limited number of samples, were broken by secondary ranking on log fold change methylation between groups (average methylation was incremented with 1 in both groups). In contrast to our previous quantitative ranking on basis of differential methylation [279] this pre-selection is unaffected by the variability of the signal in the methylated group. In the second step, the Mann-Whitney-U test was applied to the 5000 highest ranked MCs from the first step. MCs with a P-value < 0.05 (n=1709) were retained for further consideration. Finally, only the MCs associated with genes that have an annotated function in the UniProtKB/Swiss-Prot database were selected for further analyses.

Positive and negative predictive value for the methylation status of all MC was calculated. For each MC all OOSCC with a read count of ≥ 3 reads were considered as methylated and OOSCC with a read count < 3 reads were considered as unmethylated. The positive predictive value was then calculated as: (true positive pN+ OOSCC) / (true positive pN+ OOSCC + false positive pN0 OOSCC). The negative predictive value was calculated as: (true negative pN0 OOSCC) / (true negative pN0 OOSCC + false negative pN+ OOSCC).

Gene selection

To identify epigenetically down-regulated genes in pN+ OOSCC, a validated gene signature predictive for pN-status in OOSCC published by Hooff et al. [78] was combined with MethylCap-Seq data (Figure 4.1). This gene signature is based on a diagnostic microarray consisting of 696 genes and was validated on 222 OOSCC from 8 different medical centers in the Netherlands [78], [83], [281]. Genes that were found by MethylCap-Seq to be hypermethylated in pN+ OOSCC and found to be down-regulated in pN+ OOSCC by microarray were selected for further analyses.

The Cancer Genome Atlas data analysis

The Cancer Genome Atlas (TCGA) validation was performed as reported previously [279]. All clinical data (n=423) for all HNSCC patients was downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) on April 7th 2013. All patients with a tumor located in either "Floor of Mouth", "Oral Cavity" or "Oral Tongue", known pathological N-status, available methylation and mRNA data were selected (n=147). All patient and tumor characteristics of the selected TCGA cases are depicted in Supplemental Table 4.1. All pathological N-statuses were dichotomized for further analyses.

For methylation analysis level 3 methylation Infinium 450k data was downloaded for the previously selected oral SCC (OSCC) patients from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) on April 7th 2013. Additional Infinium 450k probe information was acquired from the gene expression omnibus (GEO) accession number GSE42409 including: distance to TSS; associated CpG island; chromosomal localization. All probes located up to 2000 upstream and 500 bp downstream of a TSS were selected for further analyses. R (version 3.0.3), Rstudio (RStudio, Inc) and the Lumi package[284] were used to convert the 450k probe beta values to M-values using the beta2m function. Subsequently, all M-values

were quantile-normalized by the normalizeBetweenArrays function of R package Limma [285]. Using the eBayes function of the Lumi R package all 450k probes located 2000 bp upstream to 500 bp downstream of the RAB25 TSS (n=3) were statistically compared between pN0 OSCC (n=61) and pN+ OSCC (n=86) [284].
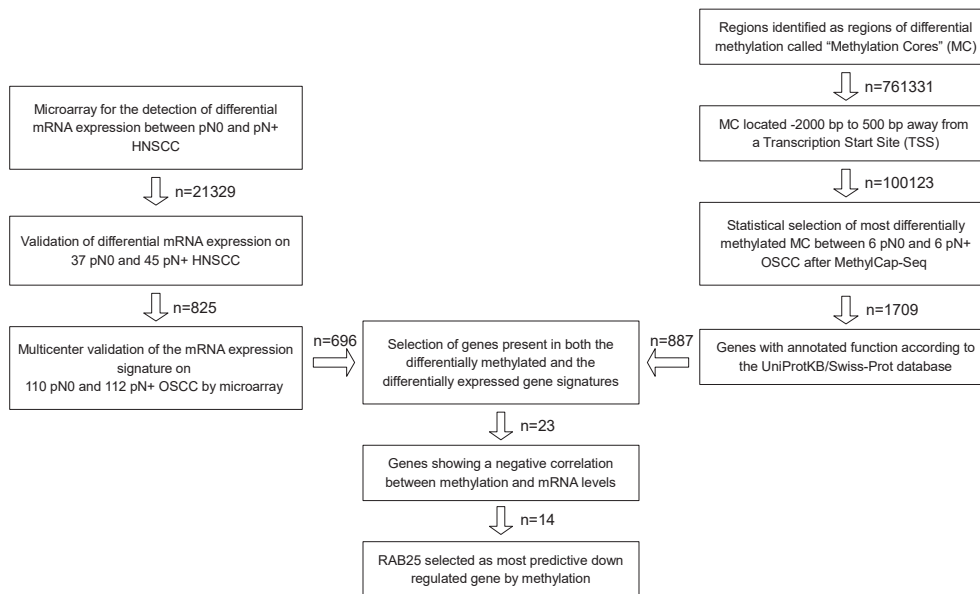
Figure 4.1. Strategy to identify epigenetically down-regulated genes in pN+ OSCC. On the left: published gene signatures predictive of pN-status in OSCC were used to identified significantly down-regulated genes in pN+ OSCC [78], [83], [281]. On the right: MethylCap-Seq was performed on 6 pN0 OSCC and pN+ OSCC [279]. All reads in MC in gene promoter regions were ranked according the likelihood of differential methylation and an approximate FDR. The 5000 MC with the lowest FDR were further tested by Mann-Whitney-U. The MC associated with genes without annotated gene functions were excluded. In the middle: the gene signature and methylation data were compared to select epigenetically regulated genes in pN+ OSCC (n=23). From these 23 genes, the epigenetically down-regulated genes in pN+ OSCC were selected. Based on the amount of mRNA down-regulation; statistical differences in methylation between pN0 and pN+ OSCC; positive and negative predictive value; RAB25 was selected as the most significantly epigenetically down-regulated gene in pN+ OSCC compared to pN0 OSCC.

For expression analysis, all mRNA Expression z-scores (RNA Seq V2 RSEM) from the HNSCC TCGA, "provisional cancer study" were downloaded from the cBioportal public portal (http://www.cbioportal.org/public-portal/)[286], [287] on April 30th 2014 and statistically compared between pN0 and pN+ OSCC by Mann-Whitney test using R. The optimal cutoff value for RAB25 mRNA levels between pN0 and pN+ OSCC was determined to be z-score of -0.4250 by ROC-curve analysis using SPSS version 22.0.1 (IBM). For copy number and mutation analyses all RAB25 mutation and GISTIC data from the HNSCC TCGA "provisional cancer study" was downloaded from cBioportal public portal on April 6th 2015. TCGA survival data was incomplete and varied between test labs and were therefore not analyzed. Spearman rank correlations between RAB25 mRNA z-scores and normalized M-values of all RAB25 probes were calculated by the basic R function cor.test [288]. Putative RAB25 regulating miRNAs were identified using the miRDB database (http://mirdb.org/miRDB/) on April 6th 2015 (n=12) [289]. Subsequently, for

all miRNAs with available data (n=6) all RNA Seq V2 RSEM (z-score Threshold ± 2), Mutation and gene copy number data for the miRNA and RAB25 were downloaded from the cBioportal public portal (http://www.cbioportal.org/public-portal/) [286], [287] on April 17th 2015. In total, 5 different types of gene copy number alterations were distinguished; -2, homozygous deletion; -1, hemizygous deletion; 0, no gene copy number alterations; 1, gain; 2, high level amplification.

Bisulfite pyrosequencing

Extracted genomic DNA (1 µg/sample) was sodium bisulfite treated EZ DNA methylation kit (Zymo, BaseClear, Leiden, the Netherlands) according to the manufacturer's protocol. RAB25 bisulfite pyrosequencing PCR and sequencing primers were designed using Pyromark Assay design version 2.0.1.15 (Qiagen). All primer sequences and PCR conditions are available in Supplemental Table 4.2. Bisulfite treated DNA was amplified using the Pyromark PCR kit according to the company protocol (Qiagen). Each reaction was performed with 12.5 µl PCR master mix 2x, 200 nmol of the forward primer and 200 nmol of the reverse primer. The PCR was performed as following: 15 min 95°C, 50 cycles of (30 sec 94°C, 30 sec 59°C,30 sec 72°C), 10 min 72°C. PCR products were checked on a 2% agarose gel with 15 µl ethidium bromide before 15 µl biotinylated PCR product was captured using 1 µl Streptavidin-coated Sepharose High Performance beads (GE Healthcare). The captured amplicons were then purified using the Q24 Vacuum Workstation (Qiagen) according to the manufacturer's protocol, washed with 70% alcohol, denatured with PyroMark Denaturation Solution (Qiagen) and washed with PyroMark Wash Buffer (Qiagen). The purified PCR product was then added to 25µl 0.3 µM RAB25 sequence primers followed by bisulfite pyrosequencing analysis using the Pyromark Q24 (Qiagen). The pyrosequencing results were analyzed using the provided Pyromark Q24 software version 2.0.6 (Qiagen). Each pyrosequencing run included 3 control samples; leukocyte DNA from healthy controls as controls for normal/endogeneous methylation levels; in vitro methylated (by SssI enzyme) leukocyte DNA as hypermethylation control and whole genome amplified (WGA) leukocyte DNA using the Illustra Ready-To-Go GenomiPhi HY DNA Amplification Kit (GE Healthcare) as a control for unmethylated DNA.

Immunohistochemistry

FFPE tumor tissue sections of 3 µm thickness were deparaffinized in xylol and rehydrated using decreasing ethanol concentrations (100%, 96%, 80%, 70%, and 50%). Antigen retrieval was performed using a citrate buffer (10mM Citric Acid, 0.05% Tween 20, pH 6.0) and heated in a microwave oven for 15 min at 300 W. Endogenous peroxidase was blocked with a 0.3% H_2O_2 solution for 30 min at room temperature followed by incubation overnight at 4°C with a mouse monoclonal antibody to human RAB25 clone 3F12F3 (Santa Cruz), diluted 1:50 in PBS with 1% bovine serum albumin. Subsequently, primary antibody detection was achieved by incubation with Envision+ (Dako) horseradish peroxidase for 30 min at room temperature and developed with 3,3-diaminobenzidine solution (Dako) H_2O_2 containing 0.03% and counterstained with hematoxylin for 2 min. Mammary epithelial cells were used as a control for positive RAB25 expression [290]. The percentage of positive tumor cells was scored as reported [291], [292] as well as three RAB25 immunoreactivity intensity (0, no staining; 1, moderate; 2, strong). Each staining was scored by 2 blinded

observers independently (MJAMC and MFM). Discordant results were discussed until consensus was reached or decided by an experienced HNSCC pathologist (BvdV). The optimal cut-off between high or low RAB25 positive tumors was determined by ROC curve analysis to be 33% RAB25 positive tumor cells. 178 Out of the 192 HPV negative tested HNSCC were evaluable for RAB25 immunoreactivity analysis.

Statistical Analysis

Statistical analysis was performed using SPSS (IBM) and R (version 3.0.3). Associations between RAB25 expression and clinico-pathological characteristics were tested using the 2 test. Survival was defined as the number of days between the first treatment and disease specific death (DSS) or disease recurrence (DFS) and analyzed by Kaplan-Meier curves and log rank test. All tests were performed two-tailed and P-value < 0.05 was considered statistically significant.

RESULTS

RAB25 is the highest ranking differentially methylated and expressed gene in pN+ OOSCC.

To identify genes whose expression is regulated by methylation, a validated gene expression signature and methylation data were combined using a stepwise selection approach as outlined in Figure 4.1. After combining the gene signature and methylation data, 23 genes were found to be present in both the differentially methylated gene panel and the differentially expressed gene panel (Supplemental Table 4. 3).

Out of these 23 potentially epigenetically regulated genes, 20 genes were hypermethylated in the pN+ OOSCC of the UMCG panel by MethylCap-Seq. Finally, 14 of these 20 genes (ACTA1, BRUNOL4, COBLL1, GFRA1, H2AFY, IL22RA1, KRT17, LAMP3, MALL, MAST4, NDUFA10, RAB25, S100A9 and WDR13) showed both promoter hypermethylation as well as expression down-regulation in pN+ OOSCC (Table 4.1). Of these 14 genes, RAB25 showed the highest down-regulation of expression and concomitant highest rate of hypermethylation in pN+ OOSCC (Table 4.1). Moreover, the RAB25 read count distribution between pN0 and pN+ OOSCC showed the highest positive and negative predictive value for pN-status (Table 4.1 and Supplemental Table 4.3). Therefore, RAB25 was studied in more detail as an epigenetically down-regulated gene in pN+ OOSCC.

Validation of epigenetic regulation of RAB25 in the independent TCGA cohort

Our data revealed a strong association between decreased mRNA expression and increased methylation of the RAB25 gene in pN+ OOSCC compared to pN0 OOSCC. To confirm this association, we selected all 147 OSCC available in the public TCGA database with available RAB25 mRNA levels, RAB25 methylation and pN-status data. Amongst the Illumina Infinium 450k probes, 5 probes were associated with the RAB25 gene (Supplemental Table 4.4). In total, 3 probes (cg15896939, cg09243900 and cg19580810) were located in the RAB25 promoter region (Supplemental Figure 4.1). Methylation status of these 3 RAB25 promoter

probes (cg15896939; P = 0.003, cg09243900; P = 0.023 and cg19580810; P < 0.001) was significantly higher in the OSCC with low RAB25 mRNA levels (Figure 4.2A). Additionally, methylation levels of all 3 RAB25 probes showed a significant negative correlation with RAB25 mRNA levels (cg15896939: R = -0.230, P = 0.005; cg09243900: R = -0.162, P = 0.049; cg19580810: R = -0.390, P < 0.001, Figure 4.2B). The analysis of the TGCA database confirmed that methylation of RAB25 is associated with decreased expression levels. Additionally, the location of 2 of these 3 probes (cg15896939 and cg09243900) overlapped with the RAB25 MC annotated by MethylCap-Seq (Supplemental Figure 4.1).

Association between RAB25 methylation and lymph node status

To determine whether RAB25 promoter methylation is associated with pN-status in OSCC, we analyzed the methylation levels of the 3 RAB25 promoter probes (cg09243900, cg15896939 and cg19580810) in 61 pN0 and 86 pN+ OSCC in the TCGA database. No significantly different methylation was found for any of the 3 RAB25 promoter probes between pN0 and pN+ OSCC (Supplemental Figure 4.2A). Additionally, RAB25 methylation was measured in an independent UMCG OOSCC cohort (n=47) using 3 different bisulfite pyrosequencing assays of the promoter region containing the annotated RAB25 MCs (bisulfite primer locations are shown in Supplemental Figure 4.1). No significant differences in RAB25 methylation levels were found between pN0 and pN+ OOSCC for any of the 9 CpG sites (Supplemental Figure 4.2B). These data suggest that DNA methylation of RAB25 promoter region is not directly related to LN metastasis in OOSCC.

Association between RAB25 expression and lymph node status

To determine the association between RAB25 expression and LN status in OOSCC, we analyzed RAB25 mRNA levels in OSCC using data available in the public TCGA database. Analyses of RAB25 mRNA levels in 147 OSCC revealed significantly lower (P = 0.015) RAB25 expression in pN+ (n=86) compared to pN0 OSCC (n=61) (Figure 4.3A). High RAB25 mRNA expression was found to be significantly associated with pN0-status (P = 0.006) (Table 4.2A). High RAB25 mRNA expression was also associated with decreased lympho-vascular invasion (P = 0.029) (Table 4. 2A).

To validate whether also RAB25 protein expression was associated with lymph node status in our UMCG OSCC cohort, immunohistochemistry was performed on 192 HPV negative-tested OOSCC. For 178 OOSCC RAB25 immunoreactivity could be scored. RAB25 immunohistochemistry (example in Figure 4.4) revealed a significant lower number of neoplastic cells with RAB25 protein expression in the pN+ OOSCC (p = 0.012; Figure 4.3B). Using a cut-off of 33% RAB25 positive neoplastic cells to define low and high expression, low RAB25 expression was significantly associated with pN+ OOSCC (p=0.002; Table 4.2B). The association between low RAB25 expression and pN+ status is in good agreement with the TCGA analysis (Table 4. 2 and Figure 4.3).

The amount of RAB25 protein expression was not associated with other clinical characteristics (Table 4.2B), DSS (P = 0.232) and DFS-survival (P = 0.260). These data support an anti-invasive function of RAB25 expression in OOSCC. Analysis of RAB25 protein levels and RAB25 MC levels revealed no associations between RAB25 methylation and RAB25 protein expression in the UMCG cohort (data not shown).
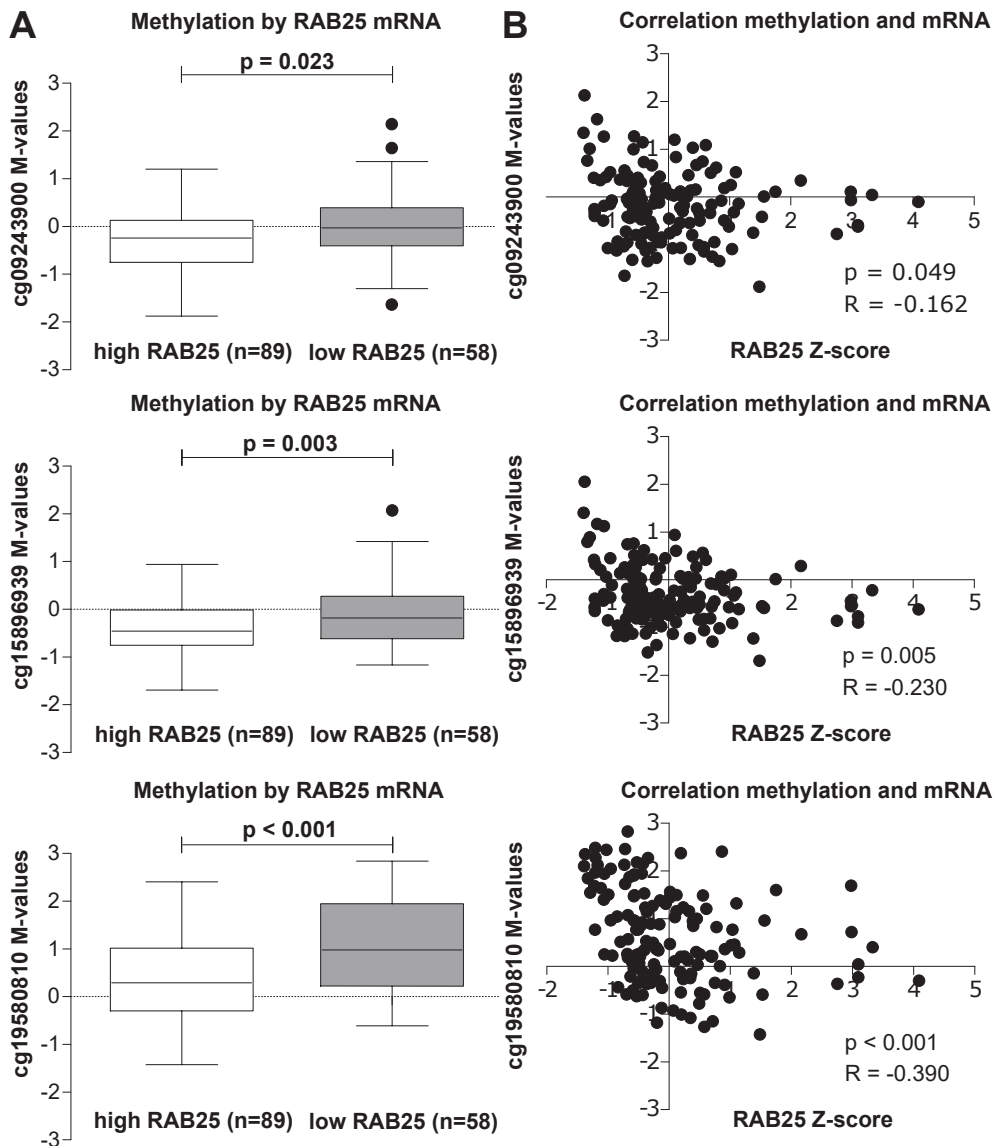
Figure 4.2. RAB25 mRNA levels in relation with the 3 RAB25 TSS 450k probes (cg09243900, cg15896939, and cg19580810) methylation levels in the TCGA OSCC cohort. A) RAB25 methylation levels compared between OSCC with high RAB25 mRNA levels and OSCC with low RAB25 mRNA levels. The M-values of the 3 RAB25 Infinium 450k promoter probes were significantly higher in OSCC with low RAB25 mRNA z-scores compared to OSCC with high RAB25 mRNA z-scores. B) Spearman correlations between RAB25 methylation and RAB25 mRNA levels. All 3 RAB25 promoter probes showed a significant negative correlation between RAB25 promoter probe M-values and RAB25 mRNA z-scores.

RAB25 gene copy numbers, mutations and miRNAs exist but occur in low frequencies

RAB25 mRNA expression is significantly associated with the methylation status of the RAB25 promoter (Table 4.1 and Figure 4.2). Additionally, both RAB25 mRNA and RAB25 protein expression are associated with pN-status (Figure 4.3, Table 4.2). RAB25 methylation status (Supplemental Figure 4.2) is however not associated with pN+ status. Therefore, DNA methylation only partly explains the regulation of RAB25 protein expression. To assess the frequency of other (epi)genetic changes that might regulate RAB25 protein expression and a possible association with LN status, the frequency of RAB25 gene mutations and gene copy number alterations was assessed in 147 OSCC selected from the TCGA database. We found a single OSCC (1/147) with a RAB25 mutation (RAB25-Q98H). In 29 of 145 OSCC RAB25 copy number gain (1 case with high level amplification) and in 15 OSCC a hemizygous RAB25 deletion (not homozygous) were detected. RAB25 mRNA levels are significantly higher in OSCC with RAB25 gene copy number increase (P = 0.024), but RAB25 mRNA levels are not associated with hemizygous deletions of RAB25 (P = 0.330). Additionally, pN-status is not associated with RAB25 copy number gain (P = 0.540), RAB25 copy number loss (P = 0.785), or with RAB25 mRNA levels and RAB25 copy number gain (P = 0.143) or RAB25 copy number loss (P = 0.584).
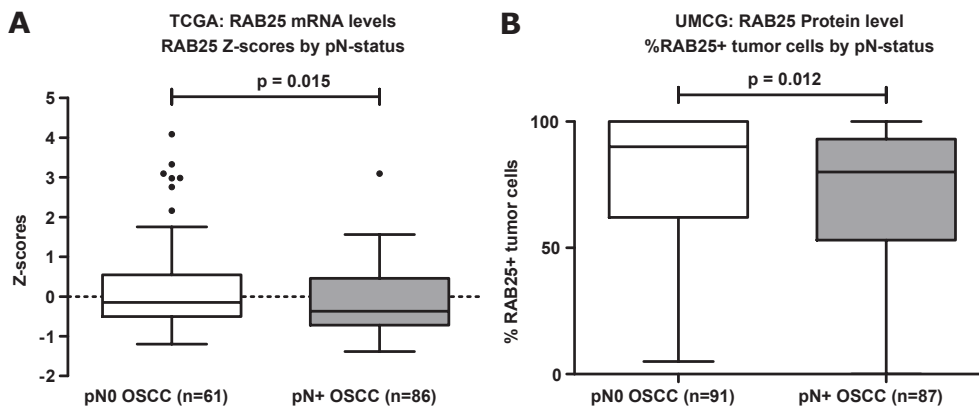
Figure 4.3. RAB25 expression levels between pN0 and pN+ OSCC in the UMCG and TCGA OSCC cohort. A) pN+ OSCC in the TCGA cohort (n=86) have significantly less RAB25 mRNA expression than pN0 OSCC (n=61) as revealed by Mann-Whitney-U test. B) pN+ OSCC in the UMCG cohort (n=87) have significantly less RAB25 positive tumor cells than pN0 OSCC (n=91) as revealed by Mann-Whitney-U test.
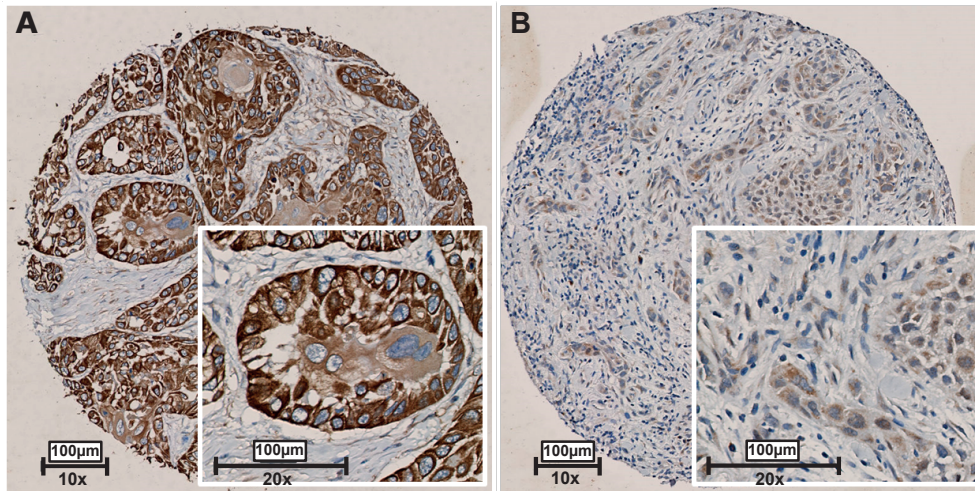
Figure 4.4. Representative examples of RAB25 expression in 2 OSCC using immunohistochemistry. Tissues were scored for the amount of RAB25 positive cells. (A) Example of a well differentiated OSCC with a high amount of RAB25 expressing cells; (B) example of a poorly OSCC with a very low amount of RAB25 positive cells.

The miRDB database contains 12 miRNAs putatively targeting RAB25 mRNA (hsa-miR-504-5p, hsa-miR-4725-5p, hsa-miR-608, hsa-miR-4651, hsa-miR-185-3p, hsa-miR-4520-3p, hsa-miR-4447, hsa-miR-8071, hsa-miR-4761-3p, hsa-miR-1296-3p, hsa-miR-6862-5p, hsa-miR-4253). For 6 miRNA expression, mutation and copy number data were available in the TCGA database. All 6 of these miRNAs did not display aberrant gene expression, mutations or copy numbers in the 530 HNSCC present in the TCGA database (data not shown).

DISCUSSION

We used a combination of genome-wide methylation analysis and a validated gene signature predictive for pN+ status in OOSCC to identify potential epigenetically regulated genes in the OOSCC metastatic phenotype. Of all analyzed genes RAB25 is the most likely epigenetically regulated and predictive gene for pN+ OOSCC. RAB25 is reported to a tumor suppressor gene lost in HNSCC subtypes [291], [292] as well as being hypermethylated in HNSCC cell lines compared to healthy tissue [291], [292], underlining the importance and epigenetic inhibition of RAB25 protein expression in carcinogenesis.

The RAB25 protein is a member of the RAB11 subfamily of small GTPases. These GTPases are emerging as novel and important regulators of cancer development and progression. Aberrant expression of small GTPases in general and RAB25 specifically [293], [294] has been detected in various cancers [295], [296] including HNSCC and OSCC [291], [292]. Interestingly, changes in RAB25 expression are correlated with tumor invasiveness in almost all cancer types [297]–[300], but only in triple-negative breast and HNSCC RAB25 functions as a tumor suppressor gene and loss of RAB25 leads to increased migration and invasion [291], [300]–[302].

Epigenetic down-regulation of RAB25 was reported in ovarian cancer compared to normal ovarian tissue [303], esophageal cancer and cell lines compared to paired normal esophageal tissue [291], and in HNSCC cell lines [291], [292]. This supports the hypothesis that loss of RAB25 expression in pN+ OSCC is caused by hypermethylation since both increased hypermethylation [97] and metastasis are associated with progressive cancer[225] and HNSCC specifically [252]. Additionally, epigenetic regulation of the expression of other small GTPases, to which RAB25 belong, has been shown in metastatic lung cancer [304] and in colon cancer [305].

We confirmed that loss of RAB25 protein expression correlated with the presence of LN metastasis in HNSCC and OOSCC specifically [291], [292], [301] and can be used to predict LN metastasis in OOSCC [78], [83], [281]. Additionally, MethylCap-Seq identified RAB25 as differentially methylated between pN0 OOSCC and pN+ OOSCC. These data suggest that RAB25 is epigenetically regulated and lost during cancer progression as a result of hypermethylation. However, we could not confirm differential methylation on a larger independent cohort using bisulfite pyrosequencing and Illumina Infinium 450k TCGA data, although we did find significant correlations between RAB25 mRNA levels and RAB25 DNA methylation levels. These data suggest that RAB25 is regulated by DNA methylation, but also potentially subjected to other forms of epigenetic regulation such as histone modification or miRNAs. However, previous reports show no relation between histone modifications and RAB25 expression in esophageal cancer [291] and alterations of 6 miRNA that regulate RAB25 was found to be almost non-existent in the TCGA OSCC database (this paper). Most RAB25 gene copy number alterations were amplifications and can thus not be responsible for down regulated RAB25 protein expression. The frequency of RAB25 loss was however too low in the TCGA OSCC database to draw firm conclusions.

Table 4.2. Correlations between RAB25 expression and tumor characteristics.

	A) RAB25 in TCGA cohort			B) RAB25 in UMCG cohort		
	Low	High		Low	High	
	N (%)	N (%)	P-value	N (%)	N (%)	P-value
Total tumours	58 (40)	89 (59)		18 (10)	160 (90)	
Total patients	58 (40)	89 (59)		18 (10)	160 (90)	
Gender						
Male	17 (36)	30 (64)	0.576	15 (83)	96 (60)	0.053
Female	41 (41)	59 (59)		3 (17)	64 (40)	
Age at diagnosis (yrs)						
Median	61	60	0.412	59	64	0.197
Range	26-85	19-87		38-80	25-94	
Site						
OSCC	58 (40)	89 (59)		14 (78)	142 (89)	0.180
Other	n.a.	n.a.		4 (22)	18 (11)	
pT status						
01-02	24 (41)	35 (59)	0.804	12 (67)	106 (66)	0.972
03-04	34 (39)	54 (61)		6 (33)	54 (34)	
pN status						
0	16 (26)	45 (74)	0.006	3 (17)	88 (55)	0.002
+	42 (49)	44 (51)		15 (83)	72 (45)	
Extranodal spread (only pN+)						
No	19 (53)	21 (70)	0.154	9 (6)	38 (53)	0.610
Yes	17 (47)	9 (30)		6 (40)	34 (47)	
Perineural invasion						
No	18 (35)	33 (65)	0.289	10 (67)	106 (73)	0.595
Yes	31 (45)	38 (55)		5 (33)	39 (27)	
Lymphovascular invasion						
No	27 (33)	56 (68)	0.029	12 (80)	112 (86)	0.573
Yes	17 (55)	14 (45)		3 (20)	19 (15)	
Histological differentiation						
Well	4 (22)	14 (78)	0.110	2 (11)	38 (25)	0.181
Moderate or Poor	54 (42)	75 (58)		16 (89)	112 (75)	
Infiltration depth (mm)						0.537
Median	n.a.	n.a.		9	15	
Range	n.a.	n.a.		3.1 - 22	0.07 - 40	
Infiltration depth (mm)						
<4 mm	n.a.	n.a.		3 (19)	24 (17)	0.823
>4 mm	n.a.	n.a.		13 (81)	121 (83)	

A) Associations between RAB25 mRNA expression and the clinical characteristics of the TCGA OSCC cohort. B) Associations between RAB25 protein expression and the clinical characteristics of the UMCG OSCC cohort.

In summary, our data suggest that epigenetic silencing of RAB25 contributes LN metastasis in OOSCC patients. Therefore, RAB25 protein expression assessment might contribute to better patient diagnosis and RAB25 epigenetic editing might open new therapeutic options for treatment of LN metastasis through epigenetic editing of demethylating agents to increase OOSCC patient prognosis and care. Genome wide methylation analysis using the MethylCap-Seq is a promising approach to identify important epigenetically regulated genes in carcinogenesis

Supplemental Table 4.1. Patient characteristics of the UMCG and TCGA database

N (%)	UMCG	TCGA
Total tumours	178 (100)	147 (100)
Total patients	178 (100)	147 (100)
Gender		
Male	111 (62)	100 (68)
Female	67 (38)	47 (32)
Age at diagnosis (yrs)		
Median	63	61
Range	25-94	19-87
Site		
Tongue	54 (30)	80 (54)
Floor of mouth	65 (37)	26 (18)
Cheek mucosa	7 (4)	0 (0)
Gum	17 (10)	41 (28)
Retromolar area	13 (7)	0 (0)
Oropharynx	18 (10)	0 (0)
Other	4 (2)	0 (0)
cN status		
cN0	108 (61)	73 (50)
cN+	70 (39)	73 (50)
Missing	0 (0)	1 (0)
pT status		
pT1	53 (30)	17 (12)
pT2	65 (37)	42 (29)
pT3	22 (12)	38 (26)
pT4	38 (21)	50 (34)
pN status		
pN0	91 (51)	61 (42)
pN1	34 (19)	26 (18)
pN2a	2 (1)	7 (5)
pN2b	43 (24)	38 (26)
pN2c	8 (5)	13 (9)
pN3	0 (0)	2 (1)
Extranodal spread (only pN+)		
No	47 (54)	40 (47)
Yes	40 (46)	26 (30)
Missing	0 (0)	20 (23)
Perineural invasion		
No	116 (65)	51 (35)
Yes	44 (25)	69 (47)
Missing	18 (10)	0 (0)
Lymphovascular invasion		
No	124 (70)	83 (57)
Yes	22 (12)	31 (21)
Missing	32 (18)	33 (22)
Histological differentiation		
Well	40 (23)	18 (12)
Moderate	103 (58)	102 (69)
Poor	25 (14)	27 (18)
Missing	10 (6)	0 (0)
HPV16 status		
Negative	178 (100)	28 (19)
Positive	0 (0)	0 (0)
Missing	0 (0)	119 (81)
Infiltration depth (mm) (n = 173)		
Median	8	n.a.
Range	0.07-40	n.a.

Supplemental Table 4.2. Overview of RAB25 primers and PCR conditions for MSP and pyrosequencing

Primer	Primer Sequence	Tannealing (°C)	[MgCl2]	Cycles	Amplicon size (bp)
PCR Forward 1	TTTTAAGTAGTTGGGTTTATAGTTATGTG	59	2	50	202
PCR Reverse 1	Biotin-CAACTAATAAACAAAAAATAACCCCTCAA				
PCR Forward 2	Biotin-TAGTTTTTAGTGGGTTGTTTTTGAAG	59	2.5	50	164
PCR Reverse 2	ATAACTAAAAACCTAAAACCCAAATAAATA				
Sequencing 1	TTAATTTTGTATTTTTTTAGTAGAA				
Sequencing 2	GTTTTTTAAAGTGTTGGGA				
Sequencing 3	AAACCCAAATAAATAAAAAAATAAT				

4

**Supplemental Table 4.3. Epigenetically regulated expression of genes associated with pN-status in OSCC.**

	Methylation Core data			MethylCap-Seq reads												DNA methylation data	
Gene	Chr	to TSS (bp)	size (bp)	pN0 (n=6)						pN+ (n=6)						P-Value	Hypermeth
ACTA1	1	-756	273	1	0	1	1	0	1	3	6	2	2	8	1	0.01	pN+
AGPAT2	9	-1318	88	0	2	0	0	0	2	4	4	1	1	4	2	0.04	pN+
BRUNOL4	18	-1543	24	2	0	0	0	0	0	3	2	3	3	1	0	0.03	pN+
COBLL1	2	-1247	191	2	2	1	2	0	0	5	2	4	3	2	5	0.02	pN+
FAM20C	7	247	210	0	0	0	0	0	0	0	1	1	0	2	1	0.03	pN+
FAM20C	7	-1727	146	4	3	2	3	3	5	1	1	0	0	3	2	0.02	pN0
GFRA1	10	-809	120	0	1	0	0	0	1	4	3	1	3	2	0	0.04	pN+
H2AFY	5	-1065	90	0	0	0	0	1	0	1	2	1	1	5	0	0.03	pN+
H2AFY	5	-1065	90	0	0	0	0	1	0	1	2	1	1	5	0	0.03	pN+
IGFBP7	4	118	228	0	0	1	0	0	0	1	1	1	1	1	2	0.01	pN+
IL22RA1	1	114	229	1	0	0	1	3	0	2	3	1	2	3	3	0.05	pN+
KCNQ5	6	0	101	3	0	4	1	0	1	6	4	4	4	2	3	0.04	pN+
KRT17	17	-296	1	0	0	1	0	0	0	2	2	1	1	3	0	0.02	pN+
LAMA3	18	173	220	0	0	0	1	0	0	1	2	1	2	1	0	0.03	pN+
LAMP3	3	0	284	0	1	1	1	0	3	3	6	4	1	3	1	0.05	pN+
MALL	2	413	152	0	0	0	0	1	0	1	1	0	1	5	2	0.03	pN+
MAST4	5	-271	57	0	0	1	0	0	0	1	0	1	1	1	1	0.03	pN+
NDUFA10	2	-1155	9	0	0	1	0	0	0	3	2	2	0	2	1	0.02	pN+
PAG1	8	-705	146	0	1	1	3	0	2	3	3	4	4	1	3	0.03	pN+
RAB25	1	-108	233	1	1	0	2	0	2	4	6	7	3	5	1	0.02	pN+
S100A9	1	490	125	0	0	2	1	1	0	1	2	2	1	5	4	0.04	pN+
THBS2	6	-1189	213	0	0	1	2	1	1	2	3	3	2	3	2	0.01	pN+
TPM2	9	0	132	1	0	1	1	1	2	0	0	0	0	0	0	0.01	pN0
TPM2	9	0	132	1	0	1	1	1	2	0	0	0	0	0	0	0.01	pN0
WDR13	X	0	54	0	1	1	0	0	0	0	4	1	2	2	1	0.05	pN+

The 24 Methylation Core (MC), representing 23 genes that were identified as potentially epigenetically regulated genes that predict N-status in OSCC, the location of the identified MC, the length of the MC, the average distance of the MC to the TSS of the associated gene, the raw read count measured by MethylCap-Seq, the p-value of read distribution between the pN0 OSCC and pN+ OSCC calculated by Mann-Whitney U, mean expression of the associated gene in pN+ OSCC, the group in which the MC is hypermethylated, the group in which the gene is down-regulated, whether the gene is indicated to be epigenetically silenced in pN+ OSCC. For FAM20C 2 significantly differentially methylated MC were found by MethylCap-Seq, for H2AFY 2 and TPM2 2 different probes were present in the expression microarray analysis. The FAM20C MC was both significant hypermethylation and hypomethylation and was therefore excluded from further analyzes.

Supplemental table 4.4. All additional probe annotation of the RAB25 probes.

Probe	Chr	Position	CpG Island	Closest TSS	to TSS (bp)	SNP
cg27550984	1	156027790	Open sea	RAB25	-3175	NA
cg15896939	1	156030809	Island: chr1: 156029612-156031006	RAB25	-156	NA
cg09243900	1	156030844	Island: chr1: 156029612-156031006	RAB25	-121	NA
cg19580810	1	156031182	Shore: chr1: 156031007 156033007	RAB25	217	NA
cg19406511	1	156036311	Open sea	RAB25	5346	rs74864564

All Infinium 450k probes associated with the RAB25 TSS their location; their CpG status, distance to the RAB25 TSS and known SNPs in the Infinium 450k probe target according to Gene Set Expression omnibus GSE42409 dataset [242].
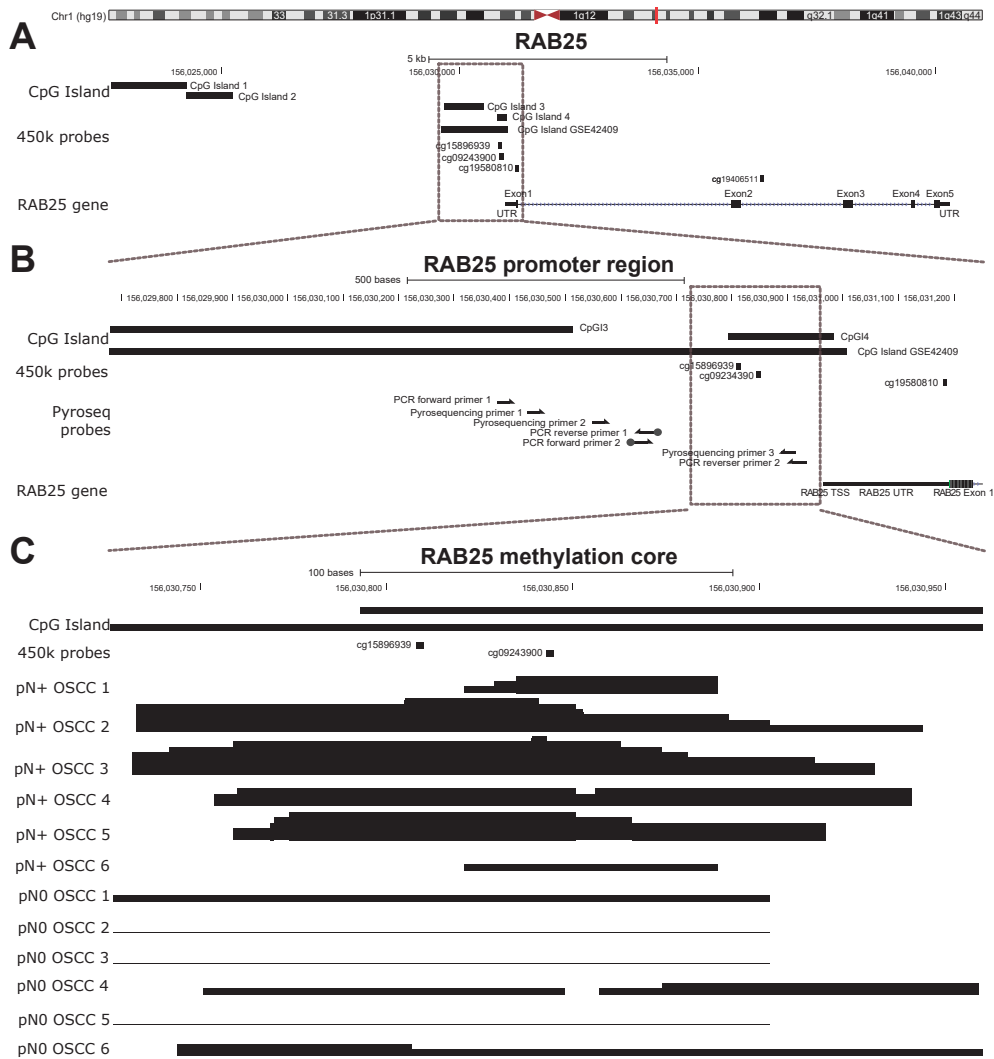
Supplemental figure 4.1. The RAB25 differentially methylated region revealed by MethylCap-Seq and the location of the bisulfite pyrosequenced CpGs. (A) Schematic representation of the genomic region around the RAB25 gene (chr1: 156,022,695 – 156,041,000, GRCh37/hg19) as extracted from the UCSC browser), the RAB25 associated Infinium 450k probes and the CpG islands as extracted from Tong et al. [291] and the GSE42409 database; (B) the RAB25 promoter region (chr1: 156,029,679-156,031,250, GRCh37/hg19) including the overlapping CpG islands as extracted from Tong et al. [291] and the GSE42409 database; the Infinium 450k probes; the RAB25 bisulfite pyrosequencing PCR and sequencing primers; The RAB25 TSS, untranslated region (UTR) and exon 1; (C) the RAB25 MC extracted from the Map of the Human Methylome is located chr1: 156,030,727-156,030,960, 6 to 239 bp from the RAB25 TSS and is located in a CpG island reported by Tong et al. [291] from 156,030,793 to 156,030,983 bp and in a CpG island reported by Price et al. [242] from 156,029,612 bp to 156,031,006 bp; the Infinium 450k probes and the reads measured by MethylCap-Seq in the methylation marker discovery cohort of 6 pN+ OSCC and 6 pN0 OSCC from the UMCG.
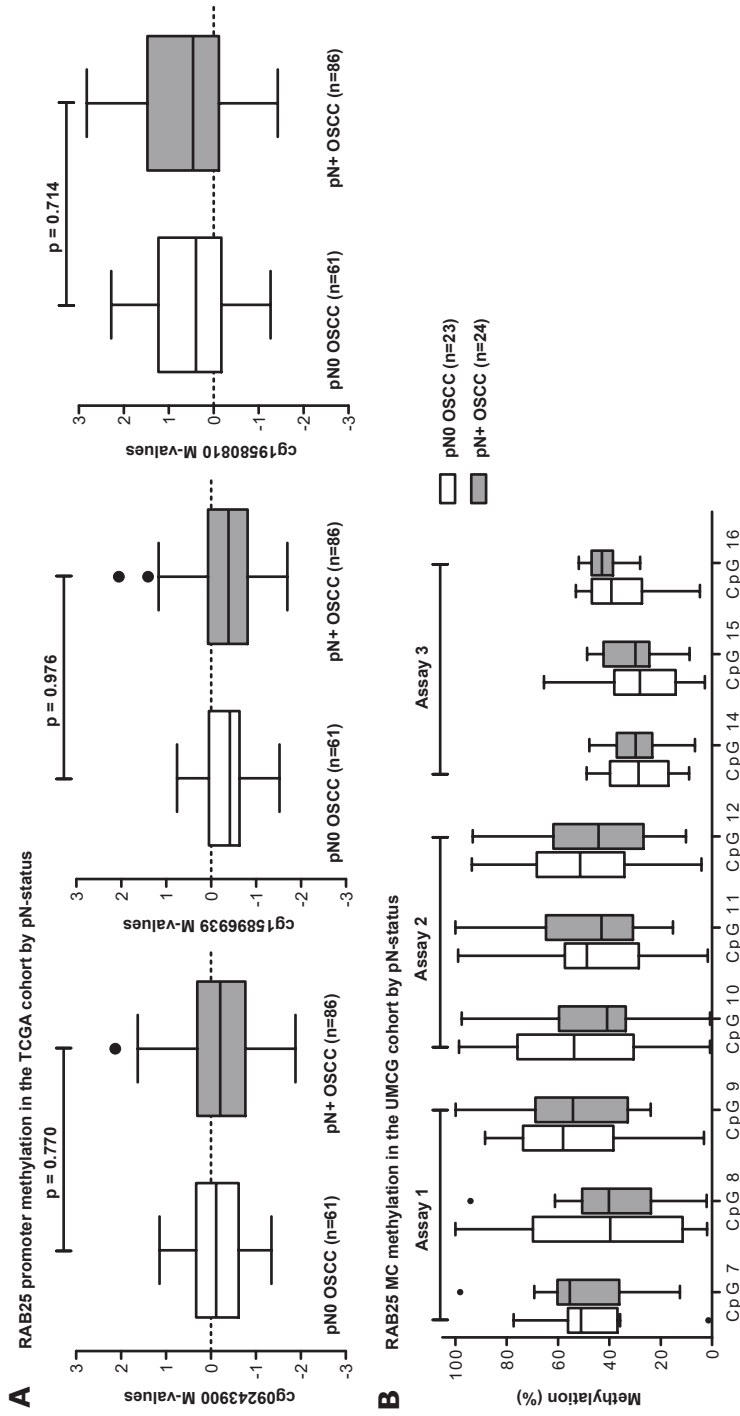
Supplemental figure 4.2. Validation of differential methylation of RAB25 between pN0 OSCC and pN+ OSCC in the TCGA and UMCG cohort. A) RAB25 promoter methylation between pN0 OSCC and pN+ OSC in the TCGA cohort (n=147). The M-values of the 3 RAB25 Infinium 450k promoter probes showed no significantly differential methylation between pN0 OSCC (n=61) compared to pN+ OSCC (n=86). B) RAB25 methylation measured by bisulfite pyrosequencing between pN0 (n=23) and pN+ OSCC (n=24). In total 9 CpG sites were analyzed using 3 bisulfite pyrosequencing assays. All CpG sites were located in a CpG island previously reported by Tong et al.[29]] and the same CpG site numbering was used. None of the 9 analyzed CpG sites were differentially methylated between pN0 OSCC and pN+ OSCC.

4

Chapter 5

Epigenetic regulation of S100A9 expression is related to lymph node metastasis and disease specific survival in patients with oral squamous cell carcinoma

M.J.A.M. Clausen[1,2], L.J. Melchers[1,2], L. Slagter-Menkema[1 3], M.F. Mastik[1], G.B.A. Wisman[4], B. van der Vegt[1], R. Grénman[5], T. de Meyer[6], W. van Criekinge[6], M.J.H. Witjes[2], J.L.N. Roodenburg[2*], E. Schuuring[1*]

* Both authors contributed equally to this work.
[1] Departments of Pathology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[2] Departments of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[3] Departments of Otorhinolaryngology/Head & Neck Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[4]Departments of Gynecologic Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[5]Department of Otorhinolaryngology – Head and Neck Surgery and Department of Medical Biochemistry, Turku University Hospital, University of Turku, Turku, Finland
[6]Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium

In preparation

ABSTRACT

A dilemma in the treatment of Oral Squamous Cell Carcinomas (OSCC) is the management of the "clinical negative neck" in which there is no evidence for lymph node (LN) metastases while a substantial risk for occult metastases is present. Accurate diagnosis is important because LN metastases severely impact patient survival. The purpose of this study was to identify methylated tumor biomarkers that predict LN metastases in OSCC and could serve as potential therapeutic targets.

Materials and Methods:

A multistep selection algorithm was performed using our OSCC-specific Methylome database, a gene expression signature and The Cancer Genome Atlas (TCGA) data to identify epigenetically down-regulated genes predictive for LN metastasis in OSCC. The gene with the most supportive evidence was characterized by immunohistochemistry and methylation-specific PCR using a cohort of OSSC and HNSCC cell lines.

Results:

From a list of 26 previously identified markers, S100A9 was identified as the most promising biomarker for LN metastases. TCGA data showed that S100A9 methylation was negatively correlated with S100A9 expression and significantly associated with the presence of LN metastasis. In an independent OSCC cohort reduced S100A9 expression was significantly correlated with LN metastasis and decreased patient survival. In HNSCC cell lines, treatment with demethylating drugs resulted in significant demethylation of the promoter and concomitant upregulation of S100A9 expression.

Conclusion:

This study shows that epigenetic down-regulation of S100A9 contributes to LN metastasis in OSCC providing a new tumor biomarker and a potential therapeutic target for the detection and treatment of OSCC patients with LN metastases.

INTRODUCTION

DNA methylation has been widely acknowledged as a very potent new biomarker. In oncology specifically, DNA methylation has been established as a regulator of cancer progression and patient survival (reviewed in [225]). DNA methylation impacts cell phenotypes and specific patterns in DNA methylation predict biological behavior and clinical characteristics such as treatment response and lymph node (LN) metastasis [228]. Additionally, in contrast to conventional genetic tumor biomarkers such as mutations, changes in DNA methylation are reversible. Therefore, DNA methylation is capable of increasing tumor diagnostics assessment as well as providing novel treatment strategies.

In oral squamous cell carcinoma (OSCC) a low survival rate is frequently seen which is often caused by lymph node (LN) metastases [224]. Additionally, due to the low sensitivity and specificity of lymph node detection by palpation and imaging techniques, under- and overtreatment of OSCC patients occur frequently [18], [170]. As a result of the complexity of the anatomy of the neck and the presence of micrometastases, only minor improvement is to be expected from developments in imaging techniques. Identifying DNA methylation patterns in the primary tumor predictive for the presence of metastases could be of great value for the accurate detection of metastases and consequently provide the most optimal treatment in this patient group. Indeed, DNA methylation status of several genes has been reported to be predictive for nodal metastasis in OSCC (reviewed in [278]) such as WISP1 [279], RAB25[306], TWIST1 [229], IGF2 [287], CDKN2A, MGMT, MLH1 and DAPK [231]. However, these tumor markers have not resulted in improved nodal metastasis detection in the clinic.

In order to successfully introduce a new DNA methylation marker into the clinic for lymph node metastasis detection, several characteristics have to be present. DNA methylation of the gene should be highly predictive for the presence of lymph node metastasis (N-status) in OSCC, preferably due to a known biological effect of the gene. The gene expression of the associated gene should change as a result of the altered DNA methylation status. Subsequently, the associated protein should be differentially expressed between OSCC with nodal metastasis (N+) and OSCC without metastasis (N0). And ultimately, this new tumor marker should also be a new potential drug target through demethylating treatment leading to its re-expression or by direct targeting the pathway that is affected by this gene.

In this study we report on the identification of S100A9 as a differentially methylated, expressed and epigenetically regulated gene in OSCC. A multistep selection algorithm was performed to select potential biomarkers for the prediction LN metastases in OSCC. DNA methylation markers were initially identified by the genome-wide methylation assessment of MethylCap-Seq, subsequently cross-validated with a gene signature predictive for N-status in OSCC and finally additional validation in the independent OSCC cohort from The Cancer Genome Atlas (TCGA). This resulted in the selection of S100A9 as the most supported DNA methylation marker. Additionally, clinical validation was performed using immunohistochemistry as well as functional validation of the epigenetic regulation of S100A9 in several HNSCC cell lines.

5

MATERIALS AND METHODS

Patient selection

All patients with OSCC that were included in this study were selected from a large cohort described previously [49], [307]. All patient characteristics as well as the precise selection for this study have been reported previously [279], [306]. Briefly, patients that were referred for the treatment of an OSCC to the Multidisciplinary Head and Neck Cancer Team of the UMCG, between 1997 and 2008 and no history of cancer treatment in the head and neck area were included in this study. For all eligible OSCC, the initial histopathological diagnoses were revised by an experienced head and neck pathologist using the original haematoxylin and eosin (HE). All selected OSCC received primary surgery as well as a neck dissection. Postoperative radiotherapy sometimes in combination with systemic therapy was given when indicated according to the national protocol. In cases with a low risk for a cervical lymph node metastasis, based on tumor characteristics, a watchful waiting policy was conducted. In these cases, the follow-up was at least two years to check for transformations from a negative to a positive clinical N-status. For the MethylCap-Seq, six pN+ OSCCs and six pN0 OSCC, matched for age and primary tumor site, were selected from the total cohort as reported previously [279], [306]. For technical validation in this study, a subgroup of 65 OSCC cases with matched age and primary tumor site, was selected from the 227 OSCC patient cohort. All clinical pathological characteristics of both included patients as well as tumors are presented in Table 5.1. To investigate the association between clinical outcome and expression of candidate markers using immunohistochemistry, 227 oral and oropharyngeal squamous cell carcinomas on tissue microarrays was used as described previously [278], [279], [306].

This study was performed according to the relevant institutional and national guidelines including the Code of Conduct for proper secondary use of human tissue in the Netherlands (www.federa.org). Because this study is performed retrospectively on data acquired from patients previously treated according to the Dutch national guidelines for oral cavity cancer, no approval from the hospital research ethics board was required according to the Dutch ethical regulations [308], [309].

DNA isolation

DNA isolation was performed as reported previously [279], [306]. Briefly, two 10 µm sections were cut from FFPE blocks for DNA extraction. Additionally, a 3 µm section was cut and HE-stained to check tumor load. Samples were deparaffinized for two hours using 750 µl xylene, and incubated overnight in 300 µl 1%SDS-proteinase K at 60ºC. Subsequently, DNA isolation was performed using phenol-chloroform extraction and ethanol precipitation. Subsequently, the isolated DNA was dissolved in 50ul TE-4 buffer for storage at 4˚C.

MethylCap-Seq analysis

MethylCap-Seq was performed as reported previously [279], [306]. Briefly, 500 ng DNA was fragmented using Covaris S2 (Covaris, Woburn, MA, USA) and methylated DNA fragments was enriched with the

MethylCap kit (Diagenode, Belgium) according to the manufacturer's protocol. Subsequently, the captures DNA was paired-end sequencing with the Illumina GA II as reported previously [151], [232]. The acquired reads were mapped back to the human reference genome (NCBI build 37.3) using the BOWTIE software [237], [283]. Reads were excluded when they were mapped back to multiple genomic loci as well as when the distance between paired ends after mapping exceeded 400 bp. In case of multiple copies of identical reads only a single read was included. Mapped reads were summarized using the "Map of the Human Methylome" [168], [382].

Table 5.1. Clinico-pathological characteristics of UMCG cohort and TCGA OSCC cohort.

N (%)	IHC	UMCG Validation	Pyroseq	TCGA 450K
Total tumors	227 (100)	27 (100)	61 (100)	147 (100)
Total patients	227 (100)	27 (100)	61 (100)	147 (100)
Gender				
Male	136 (60)	15 (56)	32 (53)	100 (68)
Female	91 (40)	12 (44)	29 (48)	47 (32)
Age at diagnosis (years)				
Median	63	66	64	61
Range	25-94	25-94	25-94	19-87
Site				
Tongue	66 (29)	9 (33)	23 (38)	80 (54)
Floor of mouth	28 (12)	6 (22)	19 (31)	26 (18)
Cheek mucosa	76 (34)	0 (0)	1 (2)	0 (0)
Gum	7 (3)	2 (7)	4 (6)	41 (28)
Retromolar area	17 (8)	4 (15)	5 (8)	0 (0)
Oropharynx	27 (12)	6 (22)	8 (13)	0 (0)
Other	6 (3)	0 (0)	1 (2)	0 (0)
cN status				
0	139 (61)	17 (63)	43 (71)	73 (50)
+	88 (39)	6 (22)	18 (29)	73 (50)
Missing	0 (0)	4 (15)	0 (0)	1 (0)
pT status				
1	61 (27)	5 (19)	21 (34)	17 (12)
2	81 (36)	11 (41)	23 (38)	42 (29)
3	28 (12)	4 (15)	5 (8)	38 (26)
4	57 (25)	7 (16)	12 (20)	50 (34)
pN status				
pN0	115 (51)	11 (41)	29 (48)	61 (42)
pN+	112 (49)	16 (59)	32 (53)	86 (58)
Extranodal spread (only pN+)				
No	64 (57)	9 (56)	18 (56)	40 (47)
Yes	48 (43)	7 (44)	14 (44)	26 (30)
Missing	0 (0)	0 (0)	0 (0)	20 (23)
Perineural invasion				
No	132 (68)	20 (74)	44 (72)	51 (35)
Yes	48 (25)	7 (26)	14 (23)	69 (47)
Missing	15 (8)	0 (0)	3 (5)	0 (0)
Lymphovascular invasion				
No	141 (72)	18 (67)	41 (68)	83 (57)
Yes	24 (12)	7 (26)	10 (16)	31 (21)
Missing	30 (15)	2 (7)	10 (16)	33 (22)
Histological differentiation				
Well	50 (23)	3 (11)	34 (56)	18 (12)
Moderate	130 (61)	15 (56)	15 (25)	102 (69)
Poor	34 (15)	5 (19)	4 (7)	27 (18)
Infiltration depth (mm) (n = 173)				
Median	8	8	7	Not available
Range	0-40.0	0-30.0	0-30.0	Not available

5

Gene Selection

Previously we combined a genome-wide methylation dataset with markers differentially methylated when comparing biopsies from pts with/without LNM with a gene expression signature predictive for N-status in OSCC [78] to select 26 potentially epigenetically deregulated genes indicative of pN-status in OSCC [279], [306]. Briefly, the "Map of the Human Methylome" build 2 [168], [382] is a database comprised of experimentally identified hotspots of differentially methylation named "Methylation Cores" (MC) across several different tissue types as well as healthy and pathological cells. MethylCap-Seq was performed on six pN0 and six pN+ OSCC to identify MC that are differentially methylated between pN0 and pN+ OSCC patients. Subsequently, the MC located between 2000 bp upstream and 500 bp downstream of a transcription start site (TSS) or in the first exon of gene annotated in the Ensemble database (v65) were ranked by False Discovery Rate (FDR). Then a Mann-Whitney-U test was applied to the 5000 MC with the lowest FDR to filter for genes with as little methylation as possible in either the pN0 or the pN+ OSCC. Finally, MC (n=887) associated with genes annotated in the UniProtKB/Swiss-Prot database were selected. The 26 genes of these differentially methylated genes that were also present in a previously reported expression signature predictive for pN-status in OSCC (n=696) [78] were selected for further validation.

In order to further narrow-down the previously selected 26 potential DNA methylation markers, data was acquired from the publicly database as a first step in silico validation. The Cancer Genome Atlas (TCGA) data was collected as reported previously [279], [306]. Briefly, all available clinical data (n=423) of all HNSCC patients was downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) on April 7th 2013. Subsequently, only those patients with a tumor reported to be located in the "Floor of Mouth", "Oral Cavity" or "Oral Tongue", an available pN-status were selected (n=189) (Table 5.1). Each pN-status was dichotomized for further analyses as reported previously [279], [306]. All available mRNA Expression z-scores (RNA Seq V2 RSEM) (n=147) as well as all available Methylation analysis level 3 methylation Infinium 450k data (n=147) for the previously selected 189 TCGA OSCC was acquired from the "provisional cancer study" cBioportal public portal (http://www.cbioportal.org/public-portal/) [286], [287]. Additional Infinium 450k probe information was acquired from the gene expression omnibus (GEO) accession number GSE42409 including: distance to TSS; associated CpG island and chromosomal localization. Of our selected 26 genes, the mRNA Z-scores from the TCGA database were statistically compared between pN0 and pN+ OSCC by Mann-Whitney test using the basic R function Wilcox.test. All probes located up to 2000 upstream and 500 bp downstream of a TSS were selected for further analyses. R (version 3.0.3), Rstudio (RStudio, Inc) and the Lumi package [284] were used to convert the 450k probe beta values to M-values using the beta2m function. Subsequently, all M-values were quantile-normalized by thenormalizeBetweenArrays function of R package Limma [285]. Using the eBayes function of the Lumi R package [284] all 450k probes located 2000 bp upstream to 500 bp downstream of the TSS of the selected (n=3) were statistically compared between pN0 OSCC (n=61) and pN+ OSCC (n=86). Correlation between the normalized Methylation M-values and the Expression Z-scores was calculated by the basic R function cor.test. Subsequently only genes with an overlap of the differentially methylated Methylation Core in the MethylCap-Seq data and the differentially methylated Infinium 450k probes in the TCGA data were selected (n=5).

CpG methylation analysis using pyrosequencing

For methylation analysis of the S100A9 MC, we performed pyrosequencing using FFPE biopsies of our own cohort of 30 pN0 OSCC and 35 pN+ OSCC. Genomic DNA (1 µg/sample) was bisulfite treated with the EZ DNA methylation kit (Zymo Research, Corp, Irvine, CA) according to the manufacturer's protocol. All primer sequences and PCR conditions are described in Table 5.2. For control of genomic DNA quality, sample DNA was amplified according to the BIOMED-2 protocol [196]. Only cases with products ≥200 bp were included for further analyses. Pyrosequencing primers were designed using PyroMark Assay design version 2.0.1.15 (Qiagen, Venlo, The Netherlands) (Table 5.2). Bisulfite treated DNA was amplified the PyroMark PCR kit according to the company protocol (Qiagen). The efficiency of the cytosine to uracil conversion by bisulfite treatment of each DNA sample was checked by Methylation Specific PCR (MSP) for beta-actin and DAPK (Table 5.2) as reported previously [139]. DKO (for double DNMT1–/– &DNMT3b–/– knock-out cells) and leukocytes DNA from healthy controls were included as negative controls (for non-methylated DNA) and DKO DNA that was in vitro methylated by SssI methyltransferase (New England BioLabs Inc., Bioké, Leiden, The Netherlands) as an optimal methylated control DNA (IV-DKO). Pyrosequencing was performed using the PyroMark Q24 (Qiagen) according to the manufacturer's protocol. Methylation percentages of all measured CpG sites were analyzed using the provided PyroMark Q24 software version 2.0.6 (Qiagen). Average methylation of all measured CpG's per pyrosequenced loci were compared as well as all individual CpG's were compared between pN0 and pN+ OSCC.

Immunohistochemistry of S100A9 expression

Immunohistochemistry was performed on tissue microarrays (TMAs) that were previously constructed [49]. TMA sections were deparaffinized in xylene and rehydrated in a graded alcohol series. Antigen retrieval was performed by heating in a microwave oven for 15 min in Tris-HCL pH=9.0. Subsequently endogenous peroxide was blocked by incubating in 0.3% peroxide solution. The slides were incubated in mouse anti-Human MRP14 (S100-A9) monoclonal antibody clone S36.48 (BMA Medicals) diluted 1:400 for one hour, followed by a 30 min incubation with HRP conjugated Rabbit anti Mouse Immunoglobulin (RaMpo, DAKO) 1:100. Finally, the slides were incubated for 30 min with HRP conjugated Goat anti Rabbit Immunoglobulin (GaRpo, DAKO) 1:100. All antibodies were diluted in 1% BSA-PBS. The slides were developed with 3,3'-di-aminobenzidine (DAB) chromogen solution (DAKO) and counterstained with Haematoxylin.Both nuclear as well as cytoplasmic staining were semi-quantitatively scored, assessing percentage of tumor cells with immunostaining. The immuno-staining was independently scored by two persons. Cases with discordant results were discussed until consensus was reached. Because cutoffs for S100A positivity have not been described in literature, we chose the median percentage of tumor cells with any staining as cutoff. A case was considered as high expression levels when the percentage of tumor cells stained was the same or higher than the median of all cases.

5

Cell lines and culture conditions

Eight HNSCC were used: UT-SCC-9 (glottis larynx), UT-SCC-23 (transglottic larynx), UT-SCC-24A (tongue), UT-SCC-32 (tongue), UT-SCC-76A (tongue) [310] (provided by Dr. Grenman from the University of Turku, Turku ,Finland), 92VU078 (oral cavity) [311] (obtained from Dr. Brakenhoff, VUmc, Amsterdam, The Netherlands), FaDu (pharynx) (ATCC® HTB-43™) and NKI-SC263 (unknown origin) (RRID:CVCL_LI51, obtained from Dr. Begg, NKI/AvL, Amsterdam, The Netherlands). All cell lines were cultured in Dulbecco's Modified Eagle Medium (Lonza, Basel, Switzerland) supplemented with 10% fetal bovine serum (Sigma-Aldrich, St. Louis, USA), 2 mM L-Glutamine (Lonza) and 100 units/ml penicillin/streptomycin (Lonza) at 37°C in 5% CO_2. Cell lines were transferred twice before treatment start. All eight HNSCC cell lines were treated with different demethylating agents and conditions. Cells were either treated for 72 hours with a low concentration (200 nM) of 5-aza-2'deoxycytidine (DAC); for 72 hours with high concentration (1 µM) DAC; for 24 hours with 300 nM trichostatin A (TSA) (Sigma), for 72 hours with a low concentration (200 nM) of DAC in combination with 300 nM TSA after 48 hours, or left untreated as described previously [142]. All cells were split to a low density 24 hours before start of demethylating treatment. DAC was refreshed every 24 hours. At the end of the DAC and/or TSA treatment, cells were collected for RNA and DNA isolation.

Cell line RNA and DNA isolation

Total RNA was isolated described previously [307]. Briefly, after washing cells with cold PBS, TRIzol® reagent according to manufacturer's protocol (Invitrogen, Carlsbad, USA) was added, the lysates collected and stored at -80°C. Subsequently, RNA was treated with DNase I (Ambion®-free Kit, Life Technologies, Carlsbad, CA USA) for 30 minutes at 37°C. Finally, the RNA was reverse transcribed using 500 ng total RNA, 300 ng random hexamer primers (Invitrogen, Carlsbad, USA) and Superscript II (Invitrogen, Carlsbad, USA) according to the manufacturer's protocol. DNA from these cell lines was isolated using a standard high salt extraction method as reported previously [312].

Real time Quantitative RT-PCR

Gene expression was analyzed by real-time PCR using the LightCylcler®480 system (Roche, Basel, Switzerland) and related software LightCycler® 480 Software release 1.5.0 version 1.5.0.39 (Roche) following LightCycler® 480 SYBR Green I Master (Roche) protocol. Reactions were carried out using intron spanning primers specific for subsequent S100A9 exons as reported in the mRNA sequence NM_002965.4 (designed by Clone Manager software (Sci-Ed Software, Denver, USA). Primers for the constitutively expressed RNA Polymerase II, RPII [313] were designed to functions as controls for normalizing mRNA expression levels. All rtQPCR Primer sequences are available in Table 5.2. PCR was performed with 2x SYBR Green I Master Mix (Roche, Basel, Switzerland) using 2,5 µl of diluted cDNA from an RT initiated with 5 ng of RNA and 900 nM primers. All samples were analyzed in triplicates and template-free blanks were also included. A series of dilutes of In vitro methylated leukocytes were used to establish a calibration curve. The relative mRNA expression was calculated using the $2-\Delta\Delta CT$ method [314].

Statistical analyses

S100A9 Methylation Core methylation percentages were compared by one-way Mann-Whitney-U test using GraphPad version 5.0. (GraphPad Software Inc., San Diego, CA, USA). S100A9 mRNA expression levels were analyzed using two-way ANOVA and Tukey's honest significance test using R version 3.5.2 and Rstudio 1.1.463 (RStudio, Inc., Boston, MA, USA).

Table 5.2. All sequence of primers used for Pyrosequencing, Q-rtPCR and Methylation Specific PCR in this study.

Primer	Sequence 5'-3'
S100A9 Pyrosequencing Forward	GTAGGAAGTGTTAAAGAAGTTTGATAGT
S100A9 Pyrosequencing Reverse	Biotin-TCAAAATATCTAAATACCCCAACTTCAC
S100A9 Pyrosequencing Sequencing	TTTTATTATATAGATAGAGTGTAAG
ACTA1 Pyrosequencing Forward	TGAGTTTTAGGAAGGGAAGGA
ACTA1 Pyrosequencing Reverse	Biotin-TCCCCCCCCCCAATTATCTATCCT
ACTA1 Pyrosequencing Sequencing	TTTGAATTTAAAAAGTTGAGTTA
IRS2 Pyrosequencing Forward	GGTTTATTTAGATGAAGAAGAAGTTGT
IRS2 Pyrosequencing Reverse	Biotin-ACAAATAAACCTAAAACCCAAAAATCT
IRS2 Pyrosequencing Sequencing	GGTTGTTAGTAGTTGAG
KCNAB1 Pyrosequencing Forward	GGTTGGATGATTTTGTAATTAGTAGTAT
KCNAB1 Pyrosequencing Reverse	Biotin-AACACCTAACAATAACCAAAACTCA
KCNAB1 Pyrosequencing Sequencing	GTAATTAGTAGTATTTGTATGTTAT
LAMP3 Pyrosequencing Forward	GGGTGTTGTGGTGTTGTT
LAMP3 Pyrosequencing Reverse	Biotin-CCCTAACATTCCTAACATTCATATTACAAA
LAMP3 Pyrosequencing Sequencing	GTGATGAAGTTTTTTGGTTAT
S100A9 exon 1-2 Q-rtPCR forward	GCTTTGACAGAGTGCAAGACGAT
S100A9 exon 1-2 Q-rtPCR reverse	GGAAGGTGTTGATGATGGTCTCTA
S100A9 exon 3-4 Q-rtPCR forward	CAGGGGGAATTCAAAGAGC
S100A9 exon 3-4 Q-rtPCR reverse	TGAACTCCTCGAAGCTCAG
RPII Q-rtPCR forward	CGTACGCACCACGTCCAAT
RPII Q-rtPCR reverse	CAAGAGAGCCAAGTGTCGGTAA
ACTB Q-rtPCR forward	TAGGGAGTATATAGGTTGGGGAAGTT
ACTB Q-rtPCR reverse	AACACACAATAACAAACACAAATTCAC
DAPK1 meth. MSP forward	GGATAGTCGGATCGAGTTAACGTC
DAPK1 meth. MSP reverse	CCCTCCCAAACGCCGA
DAPK1 unmeth. MSP forward	GGAGGATAGTTGGATTGAGTTAATGTT
DAPK1 unmeth. MSP reverse	CCCTCCCAAACACCAACC

S100A9 msRNA expression Z-scores from the TCGA databases were dichotomized based on the median S100A9 mRNA Z-score of -0.34685. Dichotomized S100A9 mRNA scores were compared to other categorial data by Chi-Square test using R version 3.5.2 and Rstudio 1.1.463 (RStudio, Inc., Boston, MA, USA). Dichotomized S100A9 mRNA scores were compared to Age at diagnosis (yrs) by Mann-Whitney-U test using R and Rstudio.

5

RESULTS

The identification of S100A9 gene

To identify genes epigenetically down-regulated in pN+ OSCC is regulated by methylation, 696 genes reported to be included in a gene signature predictive for pN-status in OSCC were combined with 887 differentially methylated genes annotated by MethylCap-Seq as reported previously [306]. 26 genes were found to be present in both the differentially methylated gene panel as well as in the gene signature. To confirm the down-regulation of expression, hypermethylation and finally the correlation between down-regulation and hypermethylation of these 26 selected genes, expression data (n = 147) for these 26 genes from the publicly available TCGA database OSCC cases was used. All 26 genes were found to be significantly differentially expressed between pN0 and pN+ OSCC (Figure 5.1). In the next step, the differential methylations status of the selected genes against the LN status was validated using methylation data selected from the same 147 OSCC cases in the TCGA dataset and 25 genes fulfilled this criterium (Figure 5.1). Only a single gene, RAB25, was found to be not significantly differentially methylated between pN0 and pN+ OSCC in the TGCA dataset.

To confirm down-regulation of gene expression by methylation of the 25 selected genes, the correlation between the mRNA Expression z-scores (RNA Seq V2 RSEM) and level 3 methylation Infinium 450k M-values were compared for each 450k probes overlapping with the annotated MC. For 21 genes (Figure 5.1), the M-values of the 450k probes overlapping with the annotated Methylation Core, were significantly negatively correlated with Z-scores of the associated gene by Spearman Correlation (p < 0.05). When the differentially methylated Illumina 450k probes of these 21 genes were aligned with the Methylation Core annotated by MethylCap-Seq, 5 genes overlapped (Figure 5.1).

To validate which of these five markers showed specific promoter methylation in clinical tumor samples, pyrosequencing assays were designed. For two genes (ACTA and IRS) no proper functioning pyrosequencing assays could not be designed due to pyrosequencing primer design limitation. The MC methylation of the three other selected genes (KCNAB1, LAMP3 and S100A9) was validated by pyrosequencing on 30 pN0 and 35 pN+ OSCC.

In total three CpG sites in the promoter of KCNAB1, five for LAMP3 and three for S100A9 were analyzed separately. LAMP3 showed the lowest average methylation levels, varying between 0 to 43 % methylation and no significant differential methylation was found between any of the five CpG sites in pN0 compared to pN+ (Supplemental Figure 5.1). The KCNAB1 promoter was 100% methylated in some OSCC but no significant differences were found between pN0 and pN+ OSCC (Supplemental figure 5.1). However, the promoter of S100A9 was not only 100% methylated in some OSCC, two of the three S100A9 MC CpG sites were significantly hypermethylated in pN+ OSCC compared to pN0 OSCC (Figure 5.2).

Analysis of mRNA expression z-scores acquired from the TCGA database, revealed that S100A9 had significantly lower mRNA levels in pN+ OSCC compared to pN0 OSCC (p < 0.001, Figure 5.3). Additionally, the two annotated S100A9 probes (cg23277715 and cg26937038) were significantly hypermethylated in

pN+ OSCC compared to pN0 OSCC (both with p < 0.001, Figure 5.4A). Finally, S100A9 methylation levels of the two annotated probes are both significantly negatively correlated with S100A9 mRNA levels in all OSCC (p <0.001, p = 0.0012, Figure 5.4B).

The stepwise analysis (Figure 5.1) resulted in the identification of S100A9 as the most predictive gene for pN+ as well as the most significantly epigenetically downregulated gene in pN+ OSCC.

S100A expression is associated with nodal status and survival

In order to validate whether any loss of expression of S100A9 was associated with pN-status, we performed immunohistochemistry (see Figure 5.5 for typical examples) on an independent group of 227 oral and oropharyngeal squamous cell carcinomas, for which the nodal status had been determined by performing a neck dissection (pN status) and described in detail previously [49] (Table 5.1). This analysis revealed that low S100A9 expression was associated with the presence of nodal metastases, both when considering nuclear and cytoplasmic expression (Table 5.3) which is in concordance with the TCGA mRNA expression z-score analyses (Figure 5.3).

Using Cox-regression analysis, low S100A expression was associated with a shorter disease-specific survival, both when analyzing nuclear (OR=0.496, 95% CI: 0.259-0.951, p = 0.006), as well as cytoplasmic expression (OR=0.495, 95% CI: 0.258-0.948, p = 0.037). Additionally, Kaplan-Meier and log-rank test confirmed a significantly decreased 5-year disease-free survival in OSCC patients with low S100A9 protein levels compared to OSCC patients with high S100A9 protein levels in both the cytoplasm (p = 0.03) as well as the nucleus (p = 0.025) (Figure 5.6). Because survival data of the TCGA OSCC cohort in general are incomplete, we could not perform survival analysis on TGCA data.

To evaluate whether expression of S100A9 is regulated by DNA methylation, eight established in vitro HNSCC cell lines (UT-SCC-9, UT-SCC-23, UT-SCC-24A, UT-SCC-32, UT-SCC-76A, VU-SCC-078, FaDu and NKI-SC263) were cultured under various conditions to induce demethylation including low or high concentration of the 5-aza-2'deoxycytidine (DAC). Because demethylation is dependent on the dividing status and/or toxicity of DAC for each separate cell lines, various combinations with DAC and trichostatin (TSA) were used as reported previously in cervical cancer cell lines [142]. UT-SCC-23 cells treated with both low concentration DAC and trichostatin were not viable and therefore this data is missing.

5
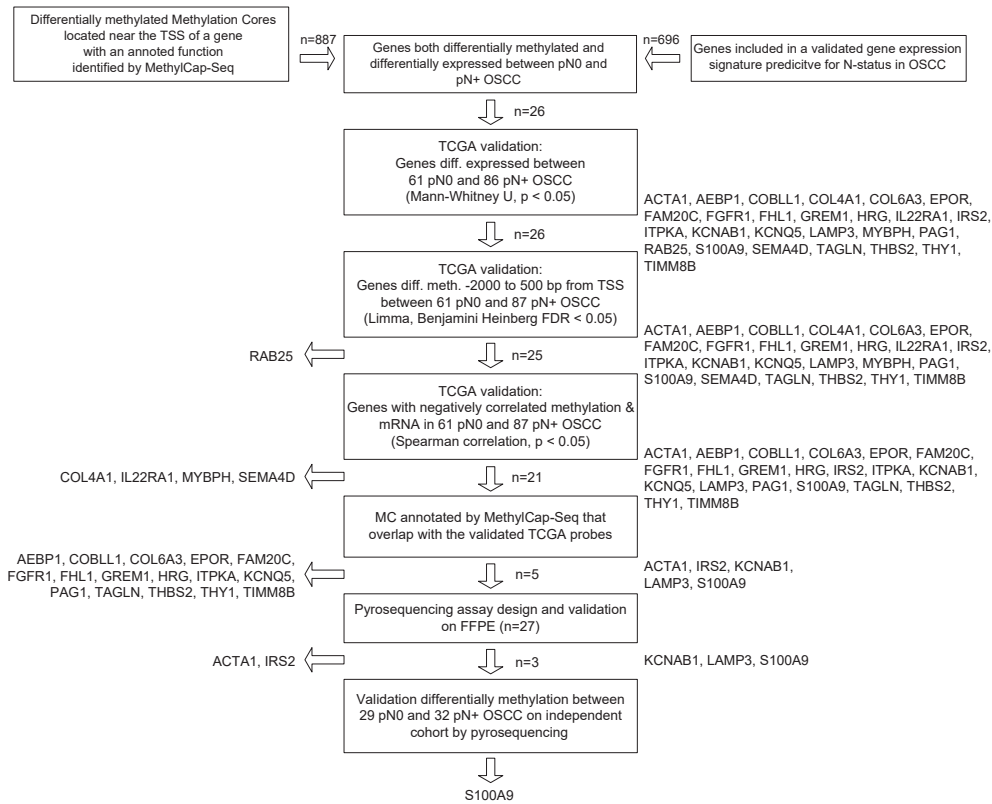
Figure 5. 1. Strategy to identify epigenetically down-regulated genes in pN+ OSCC. On the left side is the identification of differentially methylated biomarkers in OSCC as reported previously [279], [306]: MethylCap-Seq was performed on 6 pN0 OSCC and pN+ OSCC [279], [306]. All reads in MC associated with the TSS of a gene were ranked by False Discovery Rate. The 1709 genes that tested significantly differentially methylated between pN0 and pN+ OSCC were selected for further evaluation. Finally, only genes with an annotated function in the UniProtKB/Swiss-Prot database, were selected (n = 887) for cross-validation with expression data. On the right side is the identification of differentially expressed biomarkers in OSCC as reported previously [279], [306]: all 696 genes reported in a reported and validated gene signature predictive of pN-status in OSCC were used as differentially expressed genes between pN0 and pN+ OSCC [78], [83], [281]. In the middle: the gene signature and methylation data were compared to select epigenetically regulated genes in pN+ OSCC (n=26). These 26 genes were tested on data acquired from the TCGA database to selected genes significantly differentially methylated between pN0 and pN+ OSCC in the MC annotated by MethylCap-Seq, differentially expressed between pN0 and pN+ OSCC, showed significant negative correlation between methylation and mRNA levels and finally validated using Pyrosequencing. Finally, S100A9 was selected as the most significantly epigenetically down-regulated gene in pN+ OSCC compared to pN0 OSCC (see text for detailed description).

Expression of S100A9 is regulated by methylation in HNSCC cell lines using demethylation treatment

The S100A9 promoter methylation status was measured using pyrosequencing as the mean of CpG1/3 (see Figure 5.2) for all demethylation conditions. Interestingly, all HNSCC cell lines were highly methylated on CpG1/3 with an average S100A9 methylation percentage of 89.7, with the lowest (75%) in Fadu and a maximum of 99% in 92VU078. In all eight HNSCC cell lines, decreased S100A9 methylation levels were observed upon demethylating treatment (Figure 5.7). To evaluate whether decreased methylation levels
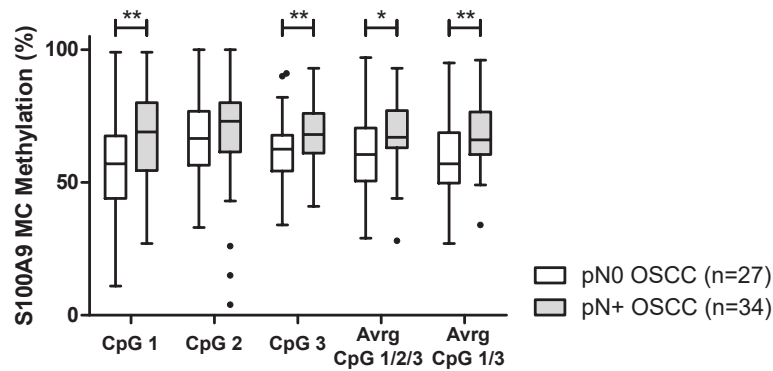
Figure 5.2. Pyrosequencing results of three CpG sites located in the Methylation Core of the S100A9 promoter annotated by the MethylCap-Seq data. The first and the third CpG site are significantly hypermethylated in pN+ OSCC compared to pN0 OSCC (p = 0.018, p = 0.020) while the second CpG site is not differentially methylated (p = 0.489). Significance of differential methylation between pN0 and pN+ by averaging the methylation of the first and third CpG site (p = 0.009) but the average of all three CpG sites located in the MethylCap-Seq annotated MC (p = 0.083).

in these same cell lines correlated with increased S100A9 mRNA expression, quantitative rtPCR analysis was performed on the same cell populations of the eight HNSCC cell lines treated with demethylating treatment. Two different intron spanning primer sets (exon 1-2 and exon 3-4) were used for measuring relative S100A9 mRNA levels. In five of the eight cell lines S100A9 expression was significantly increased after demethylating treatment (Figure 5.8). The lack of S100A9 upregulating in both NKI-SC263 and FaDu could be due to these cell lines either being sensitive to the cytotoxic effect of these compounds on cancer cells [315] or because the upregulation of S100A9 results in a reduced proliferation in these as S100A9 can have a tumor suppressing effect, resulting in lower levels of S100A9 mRNA. Overall, our analysis of these cell lines demonstrated that expression of S100A9 can be regulated epigenetically.
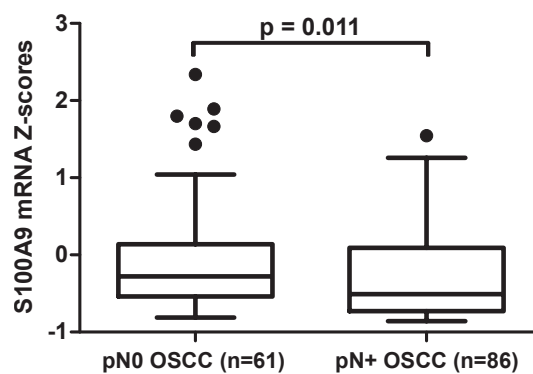
Figure 5.3. S100A9 mRNA expression levels of pN0 and pN+ OSCC in the TCGA OSCC cohort. pN+ OSCC in the TCGA cohort (n=86) have significant lower S100A9 mRNA Z-scores compared to pN0 OSCC (n=61) as revealed by two-sided Mann-Whitney-U test.
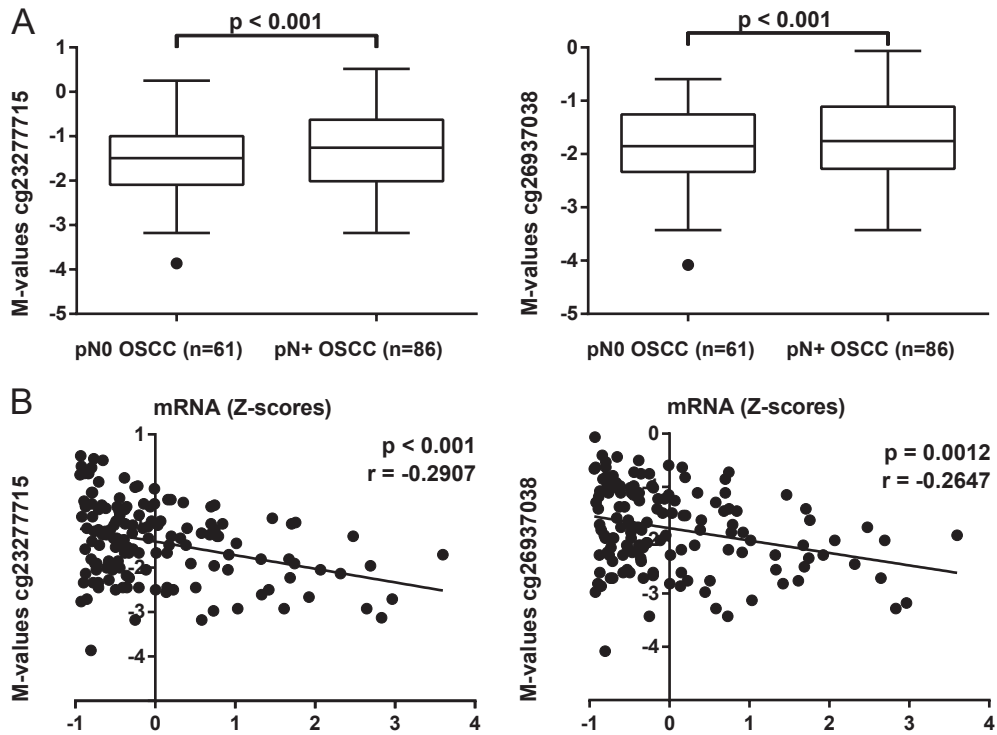
Figure 5.4. S100A9 methylation levels of pN0 and pN+ OSCC cases in the TCGA database in relation to S100A9 mRNA Z-scores.
A) S100A9 promoter Illumina 450k probe probes (cg23277715 and cg26937038) methylation M-values were significantly lower in pN0 OSCC compared to pN+ OSCC. B) Spearman correlation of S1009 promoter methylation M-values and S100A9 mRNA Z-scores levels. M-values of both S100A9 promoter probes (cg23277715 and cg26937038) were significantly negative correlated with S100A9 mRNA Z-scores.
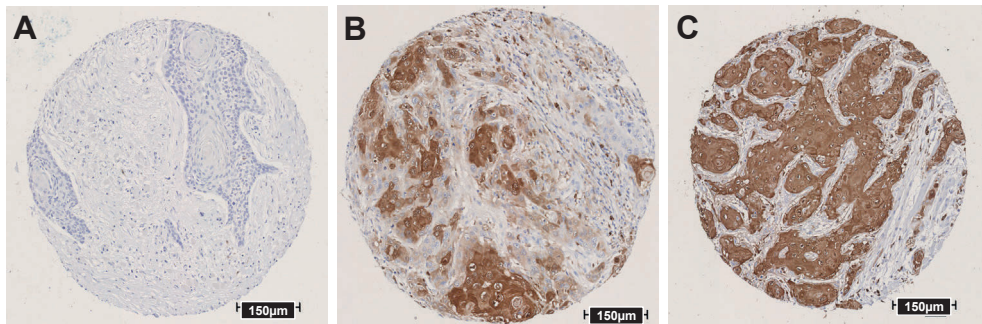
Figure 5.5. Representative examples of the different S100A9 intensity in three cores of OSCC using immunohistochemistry.
Tissues were scored for both immunoreactivity intensity in the cytoplasm as well as of the nucleus. A) A TMA core completely negative for S100A9 immunoreactivity intensity. B) A TMA core with different immunoreactivity intensity between the cytoplasm and the nucleus as well as between different cells within the same core. C) A TMA core with both cytoplasm positive for S100A9 immunoreactivity intensity as well as moderate S100A9 immunoreactivity intensity in the nuclei. TMA: Tissue Microarray

Table 5.3. Chi-square table of expression and pN status.

		S100A nuclear expression		S100A cytoplasmic expression	
		-	+	-	+
pN status	0	41	61	41	56
	+	58	39	48	35
		p = 0.006		p = 0.037	
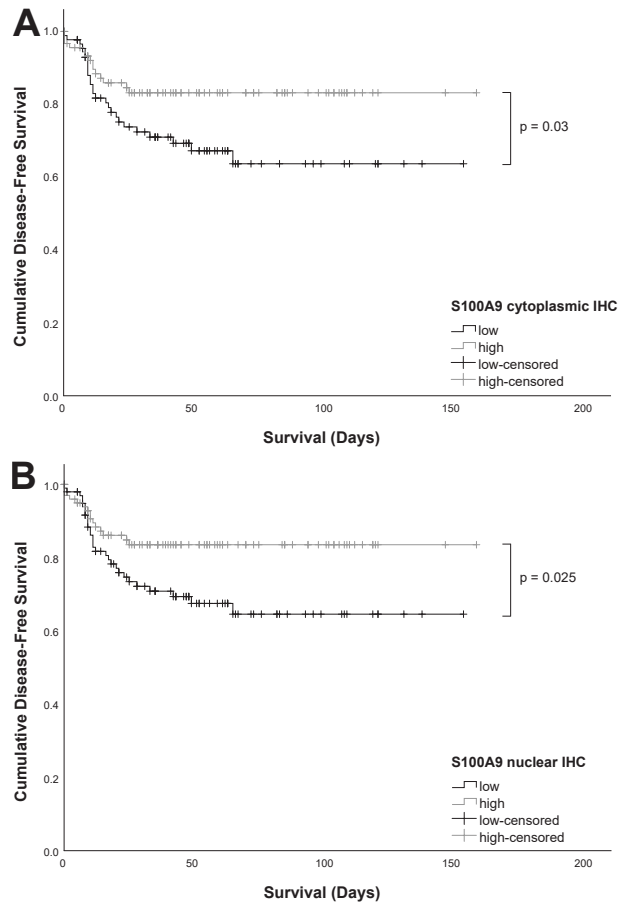
Figure 5.6. Kaplan–Meier curves. (A) Disease-specific survival of 227 OSCC stratified according to low or high S100A9 cytoplasmic IHC staining; and B) Disease-specific survival of 227 OSCC stratified according to low or high S100A9 nuclear IHC staining. IHC: Immunohistochemistry

DISCUSSION

Here we report on the identification of S100A9 whose expression is associated with OSCC without LN meta, but when down-regulated by DNA methylation becomes associated with LN metastasis.

S100A9 is part of a protein family consisting of over 25 different proteins with two highly homologue calcium-binding EF-hands and it is differentially expressed in a wide range of different cancers [316]. S100A9 can exist in three different dimensional structures, of which a heterodimer with S100A8 is the most common[317]. The activity of S100A9 is dependent on this heterodimer with S100A8 as well as on calcium. S100A9 was initially identified as a protein involved in multiple inflammatory processes[317]. More recently, S100A9 was found to influence differentiation, cell cycle, cell growth, apoptosis and the tumor microenvironment [316] through interactions with RAGE, activating downstream, proteins including upregulation of NF- B, and suppression of MAPK and AKT [318]. Additionally, S100A9 was found to induce p53-dependent apoptosis while S100A9 is upregulated by binding of p53 to the S100A9 promoter [316].

Interestingly, S100A9 is involved in a wide array of pathways which are associated with LN metastasis. In particular pathways in which a wide array of previously reported biomarkers for LN metastasis have been classified (reviewed by [75]).

S100A9 is also found to be expressed a variety of tumors including colon, colorectal, breast, cervical, gastric, hepatocellular, pulmonary and non-small cell lung [317]. On the other hand, in breast and leukemia S100A9 has been also found to be down-regulated [317], [319], [320]. More specifically, S100A9 has been found to be a negative regulator of lymph node metastasis in gastric adenocarcinoma [321]. In head and neck cancers most studies did not investigate the role of S100A9 as a marker associated with LN metastasis, but mainly reported on the expression levels of S100A9 in tumor tissue compared to (matched) normal tissue. In the majority of these studies a reduced S100A9 expression was observed in tumor [322]–[325] [326] Only one study in metastatic laryngeal SCC S100A9 was found to be down-regulated compared to both non-metastatic LSCC and normal laryngeal tissue [327] in line with our observation in OSCC.
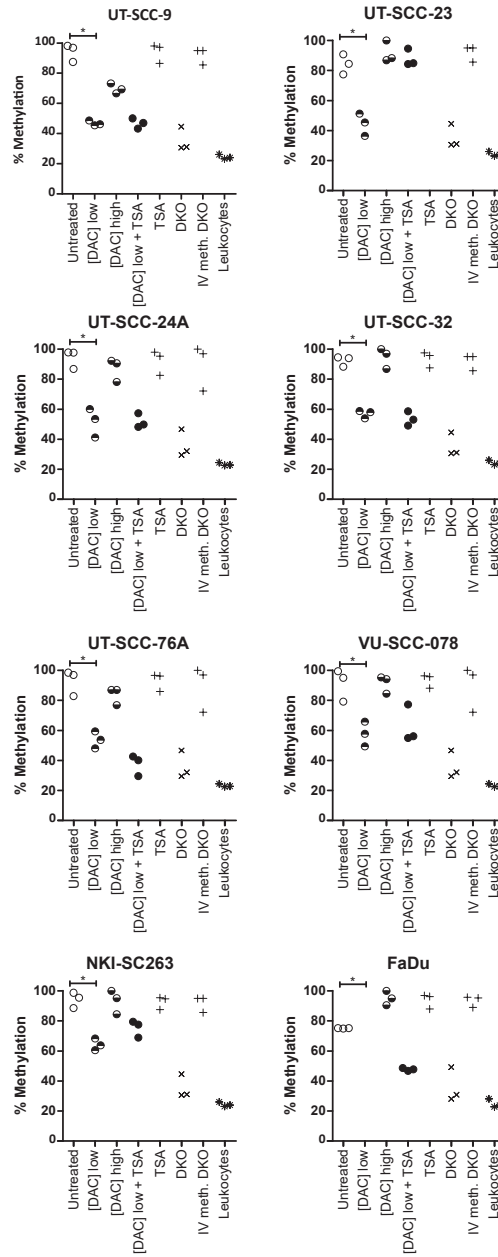
Figure 5.7. Methylations status of the S100A9 annotated MC in eight HNSCC cell lines measured by pyrosequencing. The eight HNSCC cell lines UT-SCC-9, UT-SCC-23, UT-SCC-24A, UT-SCC-32, UT-SCC-76A, 93VU078 (marked as VU-SCC-078), NKI-SC263 and FaDu were treated for 72 hours with 5-aza-2'deoxycytidine (DAC) in low concentration (200nM) or high concentration (5 μM), with 300 nM trichostatin (TSA), low concentration DAC and 300 nM TSA. Untreated cell lines are controls for baseline methylation levels of S100A9, DKO DNA and leukocytes as negative controls for non-methylated DNA and IV-DKO as an optimal methylated control. Methylation status was determined using pyrosequencing on the mean of CpG 1/3 (in triplicate) in the promoter of S100A9. UT-SCC-23 cells treated with both low concentration DAC and trichostatin were not viable and therefore this data is missing in this figure.
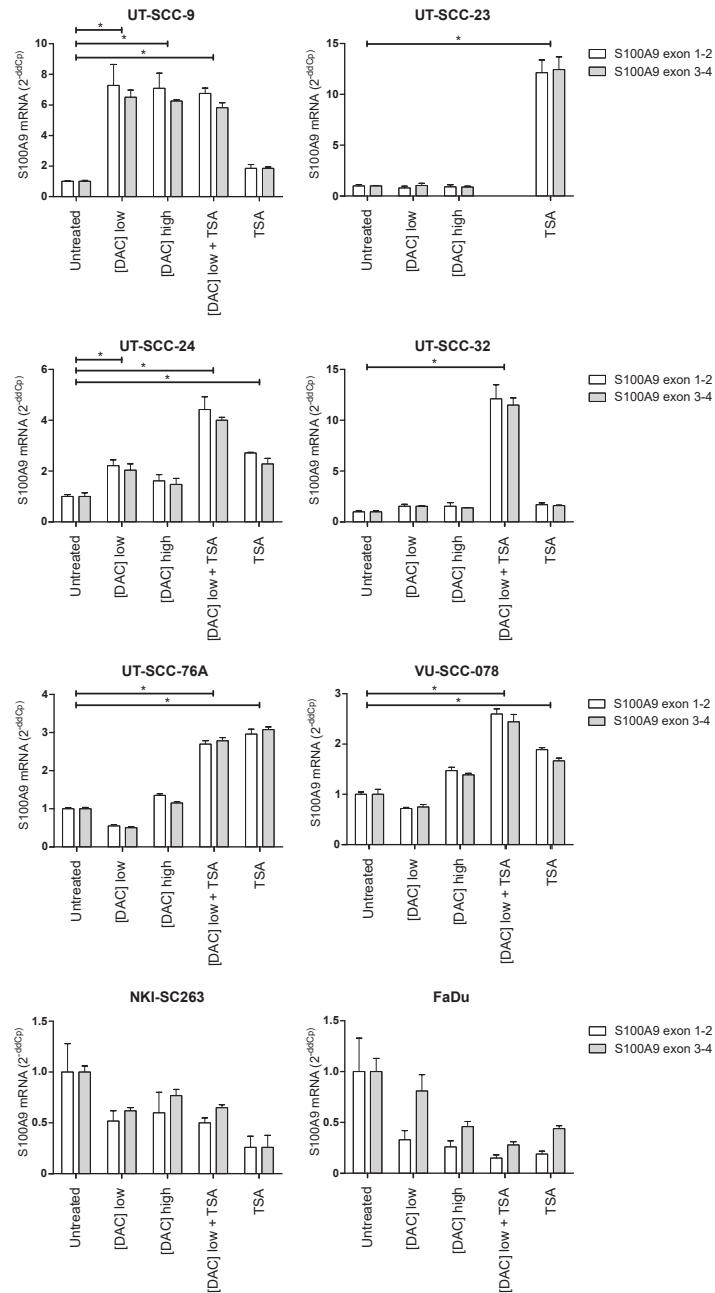
Figure 5.8. S100A9 mRNA levels in HNSCC cell lines treated under various demethylating conditions. From the same eight HNSCC cell lines treated with various demethylation conditions used to define the S100A9 methylation status (see Figure 5.7), also RNA was extracted. S100A9 mRNA levels were determined using two different primer sets for S100A9 exon 1-2 and S100A9 exon 3-4 by QRT-PCR. The level of RPII were used as controls for normalizing mRNA expression levels. The relative expression levels of S100A9 between the different demethylation conditions in each cell line were calculated compared to the levels in the untreated cell lines (set at ratio 1.0). UT-SCC-23 cells treated with both low concentration DAC and trichostatin were not viable and therefore this data is missing in this figure.

Epigenetic regulation of S100A9

Although the epigenetic regulation of the S100 protein family has been described and has been extensively investigated, DNA methylation of the S100A9 gene has not been investigated extensively in cancer and more notably HNSCC. Possible explanation is the lack of a CpG Island associated with the S100A9 gene. However, according to the ENCODE data in the UCSC genome browser there is a H3K27Ac regulatory region associated with the S100A9 TSS [328] which is associated with a lack of methylation [329]. Additionally, MeCp2, a protein that is a transcriptional repressor by DNA methylation and chromatin remodeling, was found to directly bind to the S100A9 promoter suggesting epigenetic regulation of the S100A9 gene [330]. Moreover, S100A9 expression was upregulated in breast cancer cell lines after treatment with the demethylating treatment [331]. S100A9 upregulation was also found in hematopoietic malignancy cell lines after 5-aza-2dC treatment [319].

Direct evidence for epigenetic regulation of S100A9 has been published in relation to ulcerative colitis [332]. The effects of DNA methylation on S100A9 expression was studied in bladder cancer but although S100A9 was upregulated in bladder tumors in comparison to paired normal bladder tissue and that DNA methylation of the S100A9 promoter was more abundant in cells with low S100A9 mRNA levels, S100A9 expression was not increased after demethylating treatment [333].

Although we found hypermethylation and S100A9 downregulation in metastasized OSCC, reports on whether the correlation between S100A9 levels and expression in cancer is positive or negative are inconsistent[317]. Whether S100A9 functions switches once the S100A9 expression levels cross a certain threshold is hypothesized [334]. At high levels of S100A9 seems to function as a tumor suppressor by promoting apoptosis [335] while in low levels S100A9 promotes tumor growth [336]–[342] and metastasis [339]–[344]. These findings imply that the epigenetic down-regulation of S100A9 in pN+ OSCC results in S100A9 promoting metastasis while in pN0 OSCC S100A9 functions a tumor suppressor gene.

In good agreement with all these observations, our analysis of eight HNSCC cell lines treated with different demethylation conditions demonstrated that the expression of S100A9 can be regulated by DNA methylation.

Using similar approaches, we reported previously the identification of WISP1 and RAB25 as biomarkers for LN metastasis in OSCC [279], [306]. In this study, we expanded on our previously reported in silico selection of DNA methylation markers predictive for LN metastases in OSCC by combining our analysis with an independent cohort retrieved from publicly available data (TCGA) and extensive validation on clinical samples resulting in the identification of S100A9. Remarkably, all three proteins have cellular functions that can be connected to calcium signaling. WISP1 is part of the Wnt signaling pathway in which Calcium functions as a second messenger [345], [346]. RAB25 is part of the Raf/MEK/ERK pathway that regulates cell differentiation and which is induced by calcium stimulation [347]. Epigenetic silencing of RAB25 could prevent calcium induced differentiation further promoting tumor growth. And finally, S100A9 is a well-known protein directly binding Calcium singling pathway components [316]. Moreover, S100A9 is also known to interact with the ERK pathway [348]. Interestingly, pathway analysis using The Database for Annotation, Visualization and Integrated Discovery (DAVID) [349], [350] revealed the calcium signaling

5

pathway as the only significantly enriched pathway amongst the 887 differentially methylated genes identified and select with MethylCap-Seq found earlier [279], [306] (Supplemental figure 5.2B). Analysis of the 696 genes reported to be included in a gene signature predictive for pN-status in OSCC [78] revealed that eight pathways were significantly enriched in this gene signature including the extracellular matrix, focal adhesion and calcium binding (Supplemental figure 5.2A). Interestingly, a majority of the currently known biomarkers for OSCC can be classified in nine major pathways: Cell cycle regulation, proliferation and apoptosis; Cell motility, cell adhesion and microenvironment; Transcription factors, immune system and angiogenesis [75]. The calcium pathway is related to all these pathways: cell cycle regulation [351], [352], cell proliferation [353], [354], apoptosis [351], [352], [354], cell motility [355], cell adhesion [352], [356], [357], microenvironnement [354], [358], transcription factors [354], [359], the immune system [352], [360], and angiogenesis [352], [361], [362].

In this study, immunohistochemistry was performed for the S100A9 protein in a well-established cohort of OSCC [49]. We found that higher S100A9 is associated with both pN0 status as well as a better survival rate in OSCC patients. This confirms that S100A9 acts as a potential tumor suppressor in OSCC that inhibits OSCC LN metastasis [363]. More research is needed to confirm through mechanisms S100A9 acts as a tumor suppressor. Our data provide additional information that S100A9 might act as a tumor suppressor or oncogene in OSCC due to the large patient cohorts that have previously not been used for S100A9 validation [364]. Analysis of S100A9 mRNA data from the TCGA database confirmed correlation between high S100A9 mRNA levels with pN0 status as well as other clinical factors associated with pN0s status including cN0 status, absence of extranodal spread, the absence of lymphovascular invasion as well as good histological differentiation. These data confirm that S100A9 acts as a tumor suppressor gene as well as inhibitor of LN metastasis in OSCC.

Because S100A9 is epigenetically down-regulated in pN+ OSCC, the hypermethylation of the S100A9 promoter could serve as a potential therapeutic target for treatment with demethylating agents. In fact, two specific genome-wide demethylating agents, Azacitidine and Decitabine, are being used in the clinic to reduce overall DNA methylation in myelodysplastic syndromes [365], [366]. Additionally, in 2014 a clinical trial (NCT02178072) started where HNSCC patients were treated with Azacitidine [367]. In fact, several preclinical studies have shown that Azacitidine treatment of HNSCC results in the reversal of chemoresistance and the induction of apoptosis [226]. However, genome-wide methylation could also cause harmful side effects such as the demethylation of epigenetically silenced oncogenes of metastasis promoting gene [174] such as we have shown is the case for WISP1 [279]. An additional treatment option for S100A9 promoter is a modification of the CRISPR-Cas9 complex [368]. By fusing the RNA guided enzymatic CRISPR with the catalytic domain of the demethylation enzyme TET1 instead of Cas9, this variant of the CRISPR-Cas9 system has been used to unmethylated the targeted DNA [369].
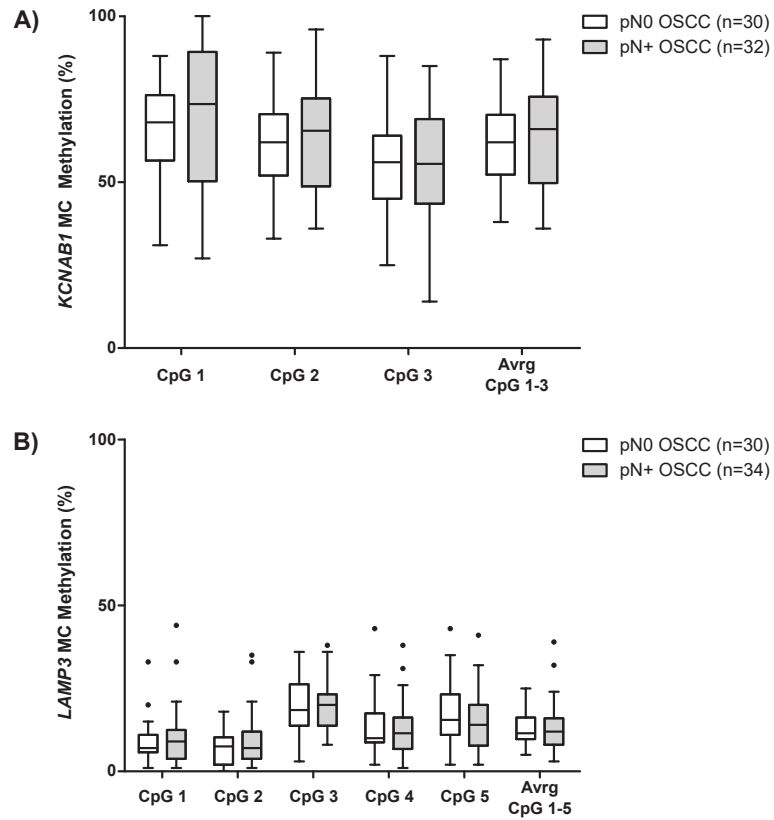
Supplemental figure 5.1. Pyrosequencing results of the KCNAB1 and LAMP3 Methylation Cores annotated by the MethylCap-Seq data. A) Three CpG sites were analyzed by pyrosequencing for the KCNAB1 MC annotated by MethylCap-Seq. None of the individual CpG sites nor the average methylation of any of the analyzed CpG sites were significantly differentially methylated between pN0 and pN+ OSCC. B) Five CpG sites were analyzed by pyrosequencing for the LAMP3 MC annotated by MethylCap-Seq. None of the individual CpG sites nor the average methylation of any of the analyzed CpG sites were significantly differentially methylated between pN0 and pN+ OSCC.

A

Top 10 enriched pathways in diff. expressed genes (n=696)

FDR = 0.05

Categories (top to bottom):
- SP PIR Keywords: extracellular matrix
- GO CC:extracellular matrix
- GO BP: ectoderm development
- GO BP: epidermis development
- KEGG pathway: ECM-receptor interaction
- KEGG pathway: Focal adhesion
- GO CC: collagen
- SP PIR Keywords: calcium binding
- Up Seq Feature: disulfide bond
- SP PIR Keywords: muscle protein

-Log (False Discovery Rate)
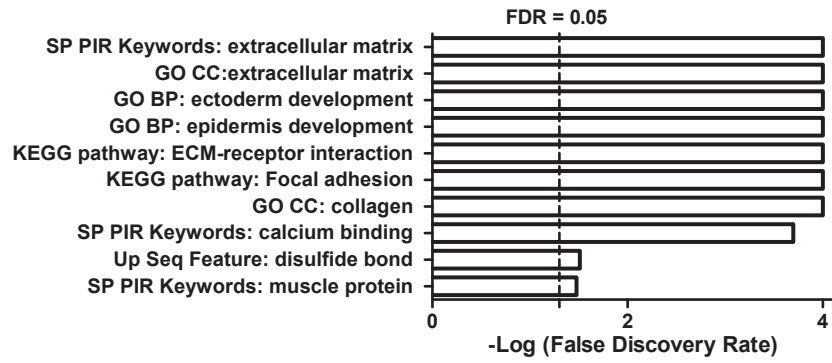
B

Top 10 enriched pathways in diff. methylated genes (n=887)

FDR = 0.05

Categories (top to bottom):
- KEGG Pathway: Calcium signaling pathway
- SP PIR Keywords: disulfide bond
- SP PIR Keywords: glycoprotein
- GO BP: muscle system process
- KEGG Pathway: Leukocyte transendothelial migration
- SP PIR Keywords: developmental protein
- SP PIR Keywords: Homeobox
- SP PIR Keywords: membrane
- SP PIR Keywords: muscle protein
- GO BP: regulation of system process

-Log (False Discovery Rate)

Supplemental figure 5.2. Pathway analysis of used gene panels A) Pathway analysis of all 696 gene included in the gene expression signature predictive for N-status in OSCC. B) Pathway analysis of all 887 differentially methylated genes identified by MethylCap-Seq.

5

Chapter 6

DNA hypermethylation of KCNA5 and TIMP3 is associated with tumor cells in saliva from patients with OSCC

M.J.A.M. Clausen[1,2]*, K. Boeve[1,2]*, L.J. Melchers[1,2]*, L. Slagter-Menkema[2,3],
M.F. Mastik[2], G.B.A. Wisman[4], B. van der Vegt[2], T. de Meyer[5,6] W. van Criekinge[5],
M.J.H. Witjes[1], J.L.N. Roodenburg[1]**, E. Schuuring[2]**

* Authors contributed equally to this work, ** Both authors contributed equally to this work.
[1]Departments of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[2] Departments of Pathology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[3] Departments of Otorhinolaryngology/Head & Neck Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[4]Departments of Gynecologic Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
[5]Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium
[6]Cancer Research Institute Ghent, Ghent University Hospital, Ghent University, Ghent, Belgium

ABSTRACT

Background:

The high local recurrence and/or second primary tumor rate of 20-30% in patients with oral squamous cell carcinoma (OSCC) is partly caused by residual tumor cells of the first primary tumor and the presence of precancerous epithelium that has not clinically manifested. Since OSCC cells are shed into the oral cavity, the detection of tumor-specific DNA methylation markers in saliva could be a tool for the early detection of local recurrences of OSCC. The aim of this study was to identify and validate new methylation markers to detect OSCC cells in saliva.

Materials and methods:

Molecular biomarkers methylated in OSCC and not in normal cells, were identified from a genome-wide methylation screening using MethylCap-Seq analysis of 12 OSCCs. Potential OSCC-specific hypermethylation markers were validated on saliva from ten OSCC patients and five younger and five age-matched healthy controls using quantitative methylation specific PCR (QMSP). These new methylation markers were compared to markers reported to be methylated in saliva by others (EDNRB, HOXA9, NID2 and TIMP3).

Results:

Using our OSCC methylome, seven genomic locations representing six genes (C11orf85, CMTM2, FERMT3, KCNA5, SIPA1 and TBX4) were identified that were significantly hypermethylated in tissues of OSCC compared to DNA from controls. QMSP analysis showed significant hypermethylation of KCNA5 in saliva of OSCC patients compared to saliva of age-matched controls (p < 0.003). Moreover, when combining QMSP results of KCNA5 with TIMP3, a 100% accuracy in detecting saliva from OSCC patients compared to non-cancer controls was observed.

Conclusions:

This study identified several new OSCC-specific methylation markers with a high sensitivity and high negative predictive value for the detection of OSCC. Two methylation (KCNA5 and TIMP3) markers might be useful for early detection of OSCC local regional recurrence in saliva cells. A larger prospective study should be done to confirm the clinical relevance of these two markers.

Keywords: DNA Methylation, Head and Neck Cancer, Saliva, Genome-Wide methylation detection, Oral Squamous Cell Carcinoma, Biomarkers for Early Detection, Quantitative Methylation Specific PCR, MethylCap-Seq.

Background

Oral Squamous Cell Carcinoma (OSCC) is the most common subtype of head and neck cancer. It is the sixth most common cancer worldwide, accounting for 650,000 new cases and 350,000 related deaths annually (www.WHO.org). Over the last 30 years, the incidence of OSCC has almost doubled, while the 5-year survival increased by 10% [370], reaching a 5-year survival of only 48% [1]. Risk factors for recurrence of OSCC are locally residual cancer after treatment or field cancerization of the oral mucosa. In addition, tumors that can metastasize show a higher chance of regional recurrence [3], [371].

Residual tumor cells are isolated cells of the first primary tumor which can remain after treatment and have the potential to develop into a local recurrence. Due to the small size of isolated cells and often submerged location, these residual tumor cells are often discovered late by regular clinical examination or imaging [372].

Due to the long-term exposition to tobacco and alcohol the epithelium of the upper aerodigestive tract might harbor areas with accumulation of pre-cancerous (epi)genetic changes [373], [374], with or without clinical manifestation which is known as field cancerization [372]. These (epi)genetic changes drive carcinogenesis [374] and therefore areas with field cancerization are at risk of developing a new malignant tumor [375].

Besides the difficulty in detecting residual tumor cells/precancerous epithelial cells and the challenge of detecting the conversion of clinical visible pre-cancerous fields (e.g. leukoplakia and erythroplakia) into new tumors as early as possible, the detection of local recurrences at an early stage is complicated by the consequences of earlier treatment. The resection area of the first primary tumor might be reconstructed with extra-oral tissue and fibrosis is induced by surgery and irradiation [376]. Although local recurrences and new primary tumors are clinically difficult to detect at an early stage, (epi)genetic alterations in DNA from residual primary tumor cells or field cancerization cells released into saliva might be detectable before clinical manifestation of recurrent disease [252]. Using (epi)genetic alterations to detect tumor DNA in saliva is therefore a promising new non-invasive strategy for the early detection of local recurrences.

Alteration in DNA methylation status is one of the epigenetic aberrations that drives tumor genesis in OSCC [123]. Changes in DNA methylation are associated with etiological factors such as cigarette smoking and alcohol consumption [373], [374] through inhibition of DNA methyltransferases (DNMT) [375], [377], [378]. Changes in DNMT expression might result in genome-wide hypermethylation associated with one of the hallmarks of cancer, chromosomal instability [97] as well as the downregulation of tumor suppressor genes [97]. Moreover, DNA methylation changes occurs early in tumorigenesis [97]. Therefore, DNA methylation markers might also be useful for the early detection of tumor cells or be detectable in shed DNA fragments in liquid biopsies such as plasma and sputum [252] and has been reported in lung [379], breast [380], colorectal [380] and hepatocellular cancer [381]. The detection of tumor cells in saliva of patients with head and neck SCC has been reported as well [128], [165], [167] and requires markers with high sensitivity and high specificity. In patients with OSCC, only few markers that are methylated in tumor tissue but not in normal epithelium have been reported [252]. To identify new methylation markers in

6

patients with OSCC that are associated with lymph node status, we recently used a novel genome-wide methylation screening method based on MethylCap-Seq analysis [151] and reported a methylome of several OSCC cases and numerous new differentially methylated tumor markers [306].

In the current study, we assessed the available methylome of a series of OSCC cases generated by MethylCap-Seq analysis [306] to identify new biomarkers associated with OSCC. We describe the identification of several new markers which are significantly hypermethylated in OSCC and not in healthy control samples. We validated the performance of these OSCC specific DNA hypermethylation markers using quantitative methylation specific PCR (QMSP). In addition, we included five DNA methylation markers previously reported to be associated with OSCC [165]–[167]. The aim of this study was to identify methylation markers with a high sensitivity and a high negative predictive value (NPV) for the detection of tumor cells in saliva from patients with OSCC.

MATERIALS AND METHODS

Identification of novel methylation markers using MethylCap-seq analysis

The strategy of methylation marker selection is summarized in Figure 6.1. To identify genomic loci hypermethylated in OSCC and not in normal tissue, in silico analysis was performed of MethylCap-Seq data [136] as reported previously [151], [232]. In summary, 12 OSCC samples and two pools of leukocytes of 500 ng DNA each were fragmented using Covaris S2 (Covaris, Woburn, MA, USA). Subsequently, methylated DNA fragments were separated from unmethylated fragments by enrichment with the MethylCap kit (Diagenode, Belgium), paired-end sequenced using the Illumina Genome Analyzer II and mapped to the human reference genome (NCBI build 37.3). For further analysis, only pair-end sequenced fragments (reads) were included that could be mapped to unique specific loci, and summarized using an in house generated "Map of the Human Methylome" for MethylCap-Seq data [382].

For further analyses only the Methylation Cores that are located either in a promotor region, between 2000 bp upstream to 500 bp downstream of the Transcription Start Site (TSS) or in the first exon of an Ensemble (v65), gene were selected and statistically compared using R with R-package Bayseq [239]. The most equally methylated MCs amongst all 12 OSCC were ranked according the likelihood of equal methylation. Additionally, an approximate false discovery rate (FDR) was calculated. The 5000 most equally methylated MCs with the lowest FDR were used for further analysis. These highest ranked 5000 MCs in OSCC were compared to the 2276 MCs available in the MethylCap-Seq data of the two leukocyte pools, by the Mann-Whitney U test (wilcox.test function in R). All MCs with a p-value < 0.05 were selected for further analyses (n = 334, Supplementary table 6.1). In the next step, all MCs were selected with a 100% positive and negative predictive value defined by ≤ 2 reads in both leukocytes pools as well as ≥ 3 reads in all 12 OSCC (Supplementary table 6.1). Finally, the MCs were compared to the semi-quantitative methylation data of the "Map of the Human Methylome" [382] to select MCs without methylation detected in the average methylome.

MethylCap-Seq analysis 12 OSCC and 2 leukocyte pools compared to 80 samples of tumors and healthy tissue, cell lines and stem cells	**Literature** Genes hypermethylated in saliva of HNSCC patients versus healthy controls
Selected Genes: *C11orf85, KCNA5, SIPA1*	**Selected Genes:** *EDNRB, HOXA9, NID2, TIMP3*

Technical validation pilot using QMSP
Comparison of saliva from 10 OSCC patients versus 10 healthy controls

Figure 6.1. Study design Methylation markers were selected using a MethylCap-Seq protocol. Selected genes were technically validated in a pilot study with saliva from 10 OSCC patients and 10 healthy controls (five younger and age-matched controls) and compared to methylation markers associated with OSCC and selected from literature. Abbreviations: OSCC, oral squamous cell carcinoma; HNSCC, head and neck squamous cell carcinoma; QMSP, quantitative methylation-specific PCR.

Technical validation of OSCC-methylation markers

For the validation, saliva from in total 10 OSCC patients were collected: seven males and three females with a median age of 63 years and with pT1-2 (n = 7) and pT3-4 (n = 3) tumors. For methylation status in the original tumor tissue, six fresh frozen (FF) tumor biopsies and nine formalin fixed paraffin embedded (FFPE) tumor resection tissues were available for DNA isolation. Saliva samples were collected from healthy controls. Five patients were planned to undergo benign corrective jaw surgery (median age 45 years, significant younger than the OSCC patients p = 0.050) and five patients were scheduled to receive dental implants (median age 67 years, age-matched with the OSCC patients). Characteristics of the patients and controls are summarized in Table 6.1. All patients and controls had no prior history of HNSCC or immunological diseases such as Sjögren's syndrome and no apparent infections in the oral cavity during saliva collection. Saliva was collected preoperatively on the day of surgery between 07:00 and 10:00 AM to exclude variation due to circadian rhythm. Patients and controls had at least 90 min without stimulation of the salivary glands by drinking, smoking or eating. Patients and controls deposited 2 ml whole saliva into a 15 ml falcon tube without a time limit. Samples were anonymized and coded for lab processing.

6

Table 6.1. Clinical characteristics of all included subjects

Patient characteristics	OSCC Patients (n)	Non age-matched Controls (n)	Age-matched Controls (n)
Total	10	5	5
Age (years) *			
Median (IQR)	63 (58 to 74)	45 (30 to 62)	67 (57 to 70)
Gender **			
Male	7	2	4
Female	3	3	1
Saliva DNA yield (µg)**			
Median (range)	64 (6 to 140)	32 (16 to 75)	32 (20 to 57)
FFPE tumor tissue	9	NA	NA
FF tumor tissue	6	NA	NA
Tumor localization		NA	NA
Tongue	4	NA	NA
Gum	2	NA	NA
Floor-of-mouth	3	NA	NA
Cheek	1	NA	NA
pT		NA	NA
01-02	7	NA	NA
03-04	3	NA	NA
pN		NA	NA
0	5	NA	NA
+	2	NA	NA
X	3	NA	NA
Infiltration depth (mm)		NA	NA
Median (range)	3 (1 to 23)	NA	NA
Tumor diameter (mm)		NA	NA
Median (range)	22 (7 to 52)	NA	NA

* OSCC versus orthognatic, p = 0.050; no significant differences between the other groups
** No significant differences between patient and control groups
*** No significant differences between patient and control groups
Abbreviations: OSCC, oral squamous cell carcinoma; IQR, interquartile range; ug, microgram; mm, millimeter; NA, not applicable; FFPE, formalin fixed, paraffin embedded; FF, fresh frozen.

Ethical considerations

Written approval and informed consent of all twenty patients and controls included in the validation study was obtained. Because of the non-invasive character of saliva sample collection, this research was not a clinical study with human subjects as meant in the Medical Research Involving Human Subjects Act as was concluded by the local Medical Ethics Review Board of the University Medical Center Groningen and no further approval was required.

DNA isolation

Saliva DNA integrity was preserved by adding 2.5 ml of 1 tablet Roche Complete mini Protease Inhibitor Cocktail (pro. #. 04693159001) dissolved in 10 ml filtered (4 °C) PBS. The saliva PBS mixture was equally divided in three 1.5 ml Eppendorf Tubes and centrifuged at 14000 rpm for 10 min at 4 °C. The pellets were

incubated in 600 μl 1% SDS-proteinase K. Both the pellet and the supernatant were separately stored at -80 °C.

Tumor DNA was isolated as follows. Approximately eight 10 μm thick sections were cut from the FFPE blocks. For quality control, the first and last section (3μm thick) were HE-stained to check for tumor load. A dedicated head and neck pathologist marked areas with >60% neoplastic cells. The 10μm FFPE sections were deparaffinized using xylene and neoplastic-enriched areas were macrodissected and used for DNA extraction. From the fresh frozen (FF) tissues, approximately four 10 μm thick sections were cut. Both, the FF and FFPE sections were incubated overnight at 60 °C in 300 μl 1% SDS-proteinase K solution.

DNA was extracted from sections and saliva cell pellets by phenol-chloroform extraction and ethanol precipitation as described previously [278]. Samples were dissolved in TE-4 buffer (50 μl for FFPE and FF, 300 μl for saliva) and stored at 4 °C. DNA quality and quantity was assessed using the Nanodrop and Biomed II PCR protocol (PCR products ≥ 200 bp) [196].

Bisulfite treatment and Quantitative Methylation Specific PCR (qMSP)

Isolated DNA was treated with bisulfite for methylation-specific-PCR (MSP) as previously described [278], [306]. Briefly, bisulfite treated DNA (bisDNA) was acquired using the EZ DNA methylation kit (Zymogen, BaseClear, Leiden, The Netherlands), according to the manufacturer's protocol. Methylation-specific-PCR (MSP) was performed on 20 ng bisDNA as follows: 10 min 95 °C, 40 cycli (1 min 95 °C, 1 min Tannealing, 1 min 72 °C), followed 10 min 72 °C and ∞ 4° C. Primer sequences and Tannealing are summarized in Table 6.2. As controls in each qMSP, leukocyte DNA from healthy individuals (as a control for endogenous methylation), leukocyte DNA that was in vitro methylated (I.V.) by SssI methyltransferase (New England BioLabs Inc., Bioké, Leiden, The Netherlands) (as a control for methylated DNA) and leukocyte DNA that was amplified according to manufacturer's protocol using whole genome amplification with the Illustra Ready-To-Go GenomiPhi HY DNA Amplification Kit (GE Healthcare, Little Chalfont, UK) (as a control for hypomethylation). Cytosine conversion by bisulfite treatment was checked with primers specific for bisulfite treated Beta-Actin (ACTB) and DAPK as described earlier [139], [278]. After MSP, PCR products were separated and visualized by custom Ethidium Bromide staining.

QMSP was performed as previously described with an internal dual-labeled hybridization probe (IDT, Coralville, IA) [139], [278]. For CMTM2 and FERMT3 no specific primers and probes with a minimum length of 250 bp within the methyl core region could be designed. For four genes (C11orf85, KCNA5, SIPA1 and TBX4), QMSP primers and probes were designed by Methyl Primer Express TM Software v1.0 (Thermo Fisher Scientific, Applied Biosystems, Leiden, The Netherlands) and checked using Clone Manager software (Sci-Ed software, Denver, USA) (Table 6.2). Serial dilutions of I.V. DNA were used to calculate standard curves for each primer-probe set, resulting in suitable conditions for the detection of methylation of C11orf85, KCNA5 and SIPA1. For TBX4 no optimal condition was found and therefore TBX4 was excluded for further analysis. The amount of bisulfite treated DNA input of each sample was determined by qMSP for ACTB (Table 6.2) as reported previously [139]. Fluorescence was measured in triplicates for 50 cycles using the following mixture: 7.5 μl of 2* LightCycler 480 Probes Master mix (Roche

6

Diagnostics GmbH, Mannheim), 300 nM of forward and reverse primers (IDT, Coralville, IA), 200 nM of probe (IDT) and 25 ng bisulfite-modified DNA. Each sample was analyzed by (LightCycler 480, Roche Diagnostics GmbH, Mannheim). Relative methylation levels for each sample were calculated as ratios using absolute measurements: the average DNA quantity of the gene of interest divided by the average DNA quantity of ACTB and then multiplied by 10,000.

The PubMed electronical database was searched for DNA methylation biomarkers that were reported to be hypermethylated in saliva of head and neck SCC patients compared to saliva of healthy controls. This search revealed four genes, EDNRB [165], HOXA9 [166], NID2 [166] and TIMP3 [167] which were used as a reference. QMSP primers and probes were selected from literature for EDNRB, HOXA9, NID2 and TIMP3 [165]–[167] (Table 6.2).

Statistical analysis

The Mann-Whitney U test was used for comparing MethylCap-Seq read counts of OSCC and leukocytes and was also used for comparing methylation levels between saliva of patients and controls. Optimal cut-offs and biomarker predictive values were determined by ROC-curves and crosstabs respectively. The accuracy of the biomarkers in detecting OSCC in saliva was determined by the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). Methylation levels in the tumors and saliva were compared using the related Wilcoxon signed-rank test. All tests were performed two-tailed. Results were considered significant when p < 0.05 or FDR < 0.05. Statistical analysis was performed with IBM SPSS Statistics 23 (Statistical Package for the Social Sciences, Inc., Chicago, IL, USA).

RESULTS

Selection of OSCC specific methylation markers

With the MethylCap-Seq analysis, a total of 11.6 to 22.3 x 10^6 reads were sequenced per sample [279], [306]. Approximately 6.91 to 14.6 x 10^6 unique reads could be mapped back to the genome per sample (Supplementary figure 6.1). Statistical analysis of reads around the transcription start site resulted in a ranking list of the 5000 most significant equally methylated regions among the 12 OSCCs. In total 334 MCs representing 319 genes were significantly differentially methylated between the 12 OSCC samples and the two leukocyte pools (Supplementary table 6.1). Of these 334 MCs, 53 MCs were hypermethylated in all OSCC and not in the leukocytes (≤2 reads). Seven MCs had a 100% positive and negative predictive value for the presence of OSCC tumor cells. These seven MCs were associated with six genes: C11orf85, CMTM2, FERMT3, KCNA5, SIPA1 and TBX4. Semi-quantitative comparison with the methylation data in the Map of the Human Methylome showed no methylation in a panel of 80 reference samples that were considered as samples not associated with OSCC (thus considered negative controls). For TBX4, CMTM2 and FERMT3 no suitable QMSP primers/probes could be designed or (Q)MSP did not pass the technical validation. The design and technical validation of C11orf85, KCNA5 and SIPA1 QMSP was optimal for further analysis. In

addition, based on literature search we included EDNRB [165], HOXA9 [166], NID2 [166] and TIMP3 [167], as these genes were reported to be associated with OSCC in saliva previously.

Technical validation of OSCC-specific methylation markers to detect tumor cells in saliva from patients with OSCC

For the validation of methylation markers, we collected a total of 2 ml of saliva from 10 patients with OSCC and from 10 non-cancer controls (five orthognathic and five dental implant patients) considered as healthy donors (Table 6.1). DNA was isolated from cell pellets collected after centrifugation of 667 µl saliva. Median amount of isolated DNA from saliva was 64 µg (range: 6 to 140 µg) among the OSCC patients, 32 µg (range: 16 to 75 µg) among the orthognathic patients and 32 µg (range: 20 to 57 µg) among the dental implant patients (Table 6.1). There were no significant differences in DNA yield from the pellets between the OSCC, orthognathic and dental implant patients.

QMSP analysis of the seven selected methylation markers on bisulfite-treated DNA from saliva cells from ten OSCC and ten control patients, revealed significant differences in methylation levels of EDNRB (p = 0.016) and KCNA5 (p < 0.001) (Figure 6.2). In fact, methylation of C11orf85, HOXA9, NID2 and SIPA1 was detected in all controls and methylation of EDNRB in 50% of the controls (not associated with age) (Figure 6.2). Five control patients were significantly younger than the OSCC patients (Table 6.1). Comparing QMSP data from saliva from OSCC with either control patients of similar or younger age, revealed a difference for only KCNA5 methylation (both OSCC-patients vs younger or older controls p = 0.001) and EDNRB methylation (only OSCC vs younger controls p = 0.003) (data not shown). Age-matched analysis did not affect the results of the other methylation markers.

One explanation for the fact that not all markers were methylated in saliva cells could be that the original tumor is not methylated for each of these methylation markers. To evaluate the effect on the sensitivity and NPV of detecting tumor cells in saliva in patients with methylated tumor tissues, the methylation status of these seven markers was tested in available tumor tissues of these same 10 OSCC patients. Methylation of four markers (EDNRB, C11orf85, KCNA5 and SIPA1) was detected in all 10 tumor tissues (Supplementary figure 6.2). Methylation was detected in nine (HOXA9) and seven (NID2 and TIMP3) of the 10 tumor tissues (Supplementary figure 6.2). Analysis of the QMSP data of saliva restricted to the three methylation markers (HOXA9, NID2 and TIMP3) with a concomitant methylated tumor tissue compared to either all 10 or five age-matched control saliva, revealed no differences in methylation levels for these three markers (data not shown).

To evaluate the possible clinical relevance for the detection of tumor cells in saliva independent on methylation status in the original OSCC tissue, we determined the optimal cut-off to discriminate between OSCC and non-cancer control DNA in saliva cells for each marker. A ROC analysis among all 20 patients (10 OSCC versus 10 control patients), revealed a high area under the curve (AUC) with a 100% sensitivity and 100% NPV of EDNRB (0.82) and KCNA5 (0.99) for detecting saliva cells of patients with OSCC (Table 6.3A). The other five markers showed a lower sensitivity and NPV when using the most optimal cut-off. The analysis on the age-matched patients (10 OSCC versus five aged-matched control patients) also showed a 100% sensitivity and 100% NPV for EDNRB (AUC 0.68) (Table 6.3B). For KCNA5 (AUC 0.98) both

6

the sensitivity (90%) and NPV (83%) decreased slightly, but interestingly with the highest specificity (100%) and positive predictive value (PPV 100%) (Table 6.3B). As none of the markers had a 100% accuracy, we combined one or more methylation markers in age matched analyses. This analysis revealed that KCNA5 combined with TIMP3 had the highest accuracy (100% for this limited dataset) in detecting saliva cells in patients with OSCC (data not shown).

Table 6.2. **The sequences of the primers and probes for all genes used for methylation detection by QMSP and MSP**

Gene name	Primer	Sequence 5'-3'	Amplicon	Tannealing (°C)	Reference
ACTB	QMSP forward	TGGTGATGGAGGAGGTTTAGTAAGT	133	60	NA
	QMSP reverse	AACCAATAAAACCTACTCCTCCCTTAA			
	QMSP probe	ACCACCACCCAACACACAATAACAAACACA			
C11orf85	QMSP forward	GAAATGCGTACGCGTAGATC	118	60	NA, MethylCap-Seq
	QMSP reverse	CAACTTCGAAACTCGTACCG			
	QMSP probe	TGGGAAGCGTATTTGCGCGTGC			
EDNRB	QMSP forward	GGGAGTTGTAGTTTAGTTAGTTAGGGAGTAG	75	60	[165]
	QMSP reverse	CCCGCGATTAAACTCGAAAA			
	QMSP probe	TTTTTATTCGTCGGGAGGAG			
HOXA9	QMSP forward	AATAAATTTTATCGTAGAGCGGTAC	226	60	[166]
	QMSP reverse	CATATAACAACTTAATAACACCGAA			
	QMSP probe	GCGCCCCATTAACCGTACGCGT			
NID2	QMSP forward	GCGGTTTTTAAGGAGTTTTATTTTC	99	62	[166]
	QMSP reverse	CTACGAAATTCCCTTTACGCT			
	QMSP probe	ACGCCGCTACCCCAAACCTTACGA			
KCNA5	QMSP forward	TTTTTTGACGTTAGGGTTAAGC	103	60	NA, MethylCap-Seq
	QMSP reverse	GAACGCCTAACGTCAAACTC			
	QMSP probe	AGAGGGGTCGGTCGATCGTTGG			
SIPA1	QMSP forward	TTCGAGTCGAGGTTAGTTC	124	60	NA, MethylCap-Seq
	QMSP reverse	CAAATCGACTAACCTCTTCG			
	QMSP probe	CGTAGCGGTAGCGATGTAGGC			
TBX4	QMSP forward	TTCGTTTTTAGTTCGAGTTGC	99	60	NA, MethylCap-Seq
	QMSP reverse	CTACGCTCTCAATCCTACGC			
	QMSP probe	CGGCGTTAGTGGACGCGG			
TIMP3	QMSP forward	GCGTCGGAGGTTAAGGTTGTT	95	62	[167]
	QMSP reverse	CTCTCCAAAATTACCGTACGCG			
	QMSP probe	AACTCGCTCGCCCGCCGAA			
ACTB	MSP Forward	TAGGGAGTATATAGGTTGGGGAAGTT	103	57	[278]
	MSP Revers	AACACACAATAACAAACACAAATTCAC			
DAPK meth	MSP Forward	GGATAGTCGGATCGAGTTAACGTC	98	60	[278]
	MSP Revers	CCCTCCCAAACGCCGA			
DAPK unmeth	MSP Forward	GGAGGATAGTTGGATTGAGTTAATGTT	101	60	[278]
	MSP Revers	CCCTCCCAAACACCAACC			

Abbreviations: NA, not applicable; (Q)MSP, (quantitative) methylation specific pcr; (un)meth, (un)methylated.
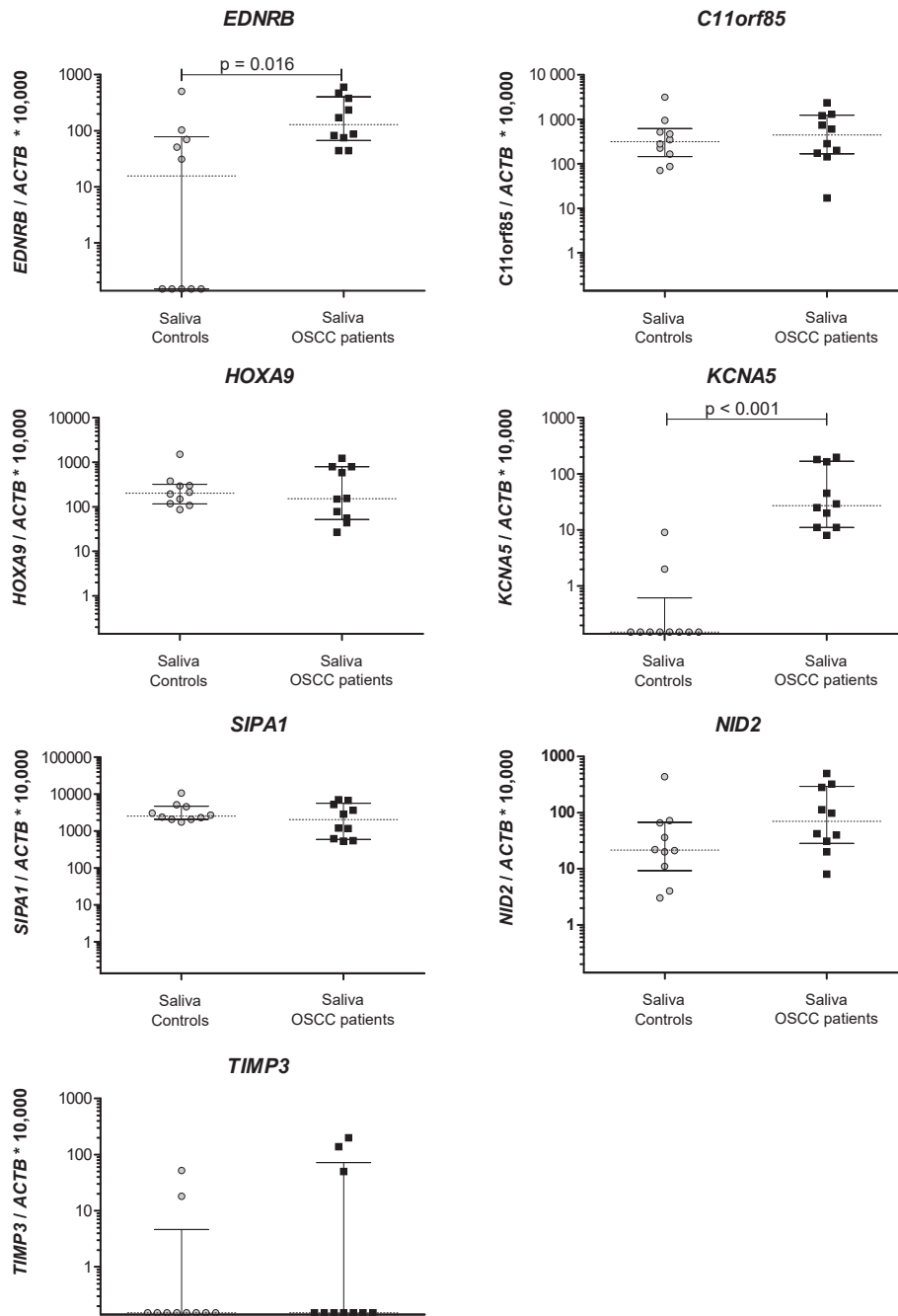
Figure 6.2. DNA methylation levels of seven OSCC-specific markers in saliva cells of OSCC patients and healthy controls. QMSP analysis of seven methylation markers using DNA extracted from cells in saliva collected from 10 OSCC patients (saliva OSCC patients) and as healthy control saliva from five younger and five age-matched controls (saliva controls). Methylation levels on the x-axis are defined as the average DNA quantity of the gene of interest divided by the average DNA quantity of ACTB and then multiplied by 10,000. Dotted and continuous line represents median with interquartile range. Only statistically significant differences ($p < 0.050$, using the Mann-Whitney-U test) between saliva of 10 controls and 10 OSCC samples are shown.

6

Table 6.3. OSCC detection accuracy in saliva of the selected methylation markers

A) OSCC patients (n = 10) vs all controls (n = 10)

Gene name	AUC	Optimal cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
C11orf85	0.56	559	50	80	71	62
EDNRB	0.82	38	100	60	71	100
HOXA9	0.44	479	40	90	80	60
KCNA5	0.99	5	100	80	83	100
NID2	0.72	27	80	60	67	75
SIPA1	0.4	5239	20	90	67	53
TIMP3	0.57	34	30	90	75	56

B) OSCC patients (n = 10) vs age-matched controls (n = 5)

Gene name	AUC	Optimal cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
C11orf85	0.44	533	50	60	71	38
EDNRB	0.68	22	100	40	77	100
HOXA9	0.39	479	40	80	80	40
KCNA5	0.98	10	90	100	100	83
NID2	0.66	38	70	80	88	57
SIPA1	0.32	5239	30	80	75	36
TIMP3	0.65	25	30	100	100	42

A ROC analysis of methylation in saliva between 10 OSCC patients and 10 control patients (A) and an age-matched analysis of saliva between 10 OSCC patients and five dental implant control patients (B) for the optimal cut-off points to detect OSCC in saliva. KCNA5 combined with TIMP3 could detect OSCC with a 100% sensitivity, specificity, PPV and NPV in an age-matched analysis.
Abbreviations: AUC, area under the curve; PPV, positive predictive value; negative predictive value; OSCC, oral squamous cell carcinoma.

DISCUSSION

DNA methylation of OSCC specific tumor markers might be useful as biomarkers for early detection of new primaries or local recurrences in OSCC patients, preferably prior to clinical manifestation. In this study we used the methylome of tissue biopsies of 12 patients with OSCC generated using genome-wide methylation screening by MethylCap-Seq analysis [279] to identify DNA methylation biomarkers with a high accuracy for the detection of OSCC. Seven new OSCC-specific biomarkers representing six genes were identified by selection of equally methylated markers between all 12 OSCC and not methylated in two pools with leukocytes from four different individuals. Moreover, the acquired highest ranking methylated candidate markers were compared to a vast methylome database of over 80 different samples considered as negative control samples. For the validation of these markers using QMSP, we could design optimal primers/probes assays for three markers (C11orf85, KCNA5 and SIPA1). To evaluate the performance of these biomarkers, DNA was isolated from saliva cells acquired from 10 OSCC patients and their corresponding tumor tissues. Saliva cells from five younger controls and five age-matched controls planned to undergo benign surgery served as healthy controls. KCNA5 was the best marker (independent of age) as it was significantly hypermethylated in OSCC saliva cells in comparison to control saliva. The possible clinical relevance of KCNA5 is further illustrated by the very high sensitivity (90%),

NPV (83%), specificity (100%) and PPV (100%), the highest of all markers tested in this study (Table 6.3B). Moreover, a panel of KCNA5 and TIMP3 could further improve the accuracy of detecting OSCC in saliva cells (100%) in an age matched analysis. Due to the limited size of our pilot group, the diagnostic potential of these biomarkers must be validated on a larger independent and prospective cohort. Similarly, a saliva database containing samples of 5-year-follow-up, pre- and post-operative as well as pre-malignant cases should be constructed for prospective studies and to assess the background methylation caused by non-tumor cells.

The use of molecular markers for the early detection and monitoring treatment response and disease progression using body fluids like saliva, sputum, plasma, cerebrospinal fluid and urine [252], [383] does not have clinical utility today [384] but has great promise to contribute to improved clinical care by early detection of OSCC or monitoring the treatment response. Since DNA methylation is important in carcinogenesis, occurs early in tumorigenesis and is detectable in patient saliva [385], DNA methylation markers could contribute to the early detection of local recurrences of OSCC. Additionally, aberrations in DNA methylation arise early in tumorigenesis [97]. Therefore, methylation of these reported genes is not suitable as methylation markers in the "older" age-matched OSCC cohort.

Several methylation markers for the detection of cells in saliva of patients with OSCC were reported previous (EDNRB, HOXA9, NID2 and TIMP3) [165]–[167]. As a comparison to our new markers, we analyzed these markers in parallel on the same samples using QMSP. In our cohort, methylation of HOXA9 and NID2 was detected in all saliva cells of health individuals. Methylation of EDNRB was observed in 50% of this saliva, but the difference between saliva of OSCC patients and of age-matched healthy controls was not significant. An explanation for the frequent methylation in normal control, especially in the saliva of the "older" age-matched "healthy" (non-cancer) saliva cells is that methylation of many genomic sequences has been reported to increase with age [386], that might explain the methylation of EDNRB in saliva in the older group of controls. This implies that methylation of these reported genes in saliva of "aged" OSCC cannot reliable discriminate between saliva with and without OSCC.

Note that the four markers selected from literature showed significant methylation in our age-matched samples from normal saliva is also an explanation why these markers were not present in our selected list of 2276 highest ranking methylation cores from the MethylCap-seq analysis of OSCC tissue samples.

With the genome-wide methylation analysis, within the methylome of millions of methylated DNA fragments in 12 OSCC samples, we eventually identified and validated three of the six new candidate markers for OSCC (C11orf85, KCNA5 and SIPA1). The pathophysiology of the novel genes related to OSCC or other types of cancer is not yet fully clarified. KCNA5 is a member of the voltage-gated potassium channel subfamily A [387]. In Ewing sarcoma cells methylation of the KCNA5 promoter region is correlated with cell survival and proliferation [388]. Signal-induced proliferation associated protein 1 (SIPA1) is located at the 11q13 chromosome close to CCND1 (Cyclin D1) and is known for influencing growth factors and cytokines by regulating RAP1 in hematopoietic cells [389], [390]. Loss of SIPA1 resulted in myeloproliferative disorders in mice [390]. The interaction between SIPA1 and RAP1 is also associated with metastasis in breast and prostate cancer by mediating cell adhesion signaling and metastasis suppressor gene signaling [390].

6

Recently, SIPA1 was found to be overexpressed in OSCC and correlated to lymph node metastasis [391]. C11orf85 also called MAJIN (membrane anchored junction protein) plays a role in telomere attachment to the inner nuclear membrane during meiosis [392], [393]. C11orf85/MAJIN is related to cancer as one of the genes in a 92-gene signature that is prognostic for overall survival in multiple myeloma patients [394]. Currently, no studies are available that report the exact role of C11orf85/MAJIN in oncogenesis. The biological significance of these three methylated genes in OSCC has not been elucidated in great detail and needs further investigation in future.

In conclusion, using the methylome of 12 OSCC tissue samples based on a genome-wide methylation screening approach, we have identified several novel biomarkers commonly methylated in OSCC. With one of these methylation markers (KCNA5) cells in saliva that are associated with OSCC patients could be detected with a high accuracy. Moreover, it is of interest to perform a larger scale evaluation for KCNA5 combined with TIMP3, given the 100% accuracy found for detecting OSCC cells in saliva. Irrespective of the small study size, our findings demonstrate the high sensitivity of Quantitative Methylation Specific PCR for detecting methylation on saliva cell DNA. DNA methylation detection using saliva has potential as an easy, low-cost, non-invasive and accurate diagnostic tool to improve the early detection of local recurrences or second primary tumors in OSCC. Our findings warrant evaluation of the clinical relevance of these methylation markers in larger cohorts.

Supplementary table 6.1 OSCC hypermethylation markers selected with MethylCap-Seq data

All 334 Methylation Cores (MCs) with a p-value < 0.05 between the 5000 highest ranked MCs in OSCC compared to the 2276 MCs available in the MethylCap-Seq data of the two leukocyte pools by Mann-Whitney-U using R and the wilcox.test function. The final seven MCs representing six genes with a 100% positive and negative predictive value defined by ≤ 2 reads in both leukocytes pools as well as ≥ 3 reads in all 12 OSCC are highlighted in bold and underlined. Location of the methylation as extracted from the "Map of the Human Methylome" [168], [382]. Abbreviations: Chr, chromosome; TSS, transcription start side; FDR, false discovery rate; bp, base pair).

Gene name	Chromosome locus	Gene regio start (GSE42409)	Gene regio end (GSE42409)	p-value (Mann-Whitney U test)	OSCC patient 1	OSCC patient 2	OSCC patient 3	OSCC patient 4	OSCC patient 5	OSCC patient 6	OSCC patient 7	OSCC patient 8	OSCC patient 9	OSCC patient 10	OSCC patient 11	OSCC patient 12	Leukocyte pool 1	Leukocyte pool 2
AC012074.3	2	25595094	25595647	0.000	9	9	9	5	9	10	6	13	9	8	10	10	14	14
PIGQ	16	614601	615239	0.000	5	9	9	11	2	4	8	5	9	9	14	7	15	15
AC021021.1	2	6635378	6635815	0.000	8	6	7	8	6	6	14	9	10	9	12	8	14	14
KCNA5	**12**	**5153088**	**5153505**	**0.000**	**6**	**4**	**11**	**6**	**5**	**10**	**12**	**11**	**7**	**6**	**10**	**3**	**1**	**1**
RAPSN	11	47470539	47471210	0.000	7	7	8	12	7	10	8	14	7	10	11	8	14	14
RP11-56M3.1	10	92913015	92913355	0.000	9	10	7	10	5	7	7	13	9	8	10	9	13	13
KIF22	16	29800874	29801501	0.000	3	8	5	8	6	9	11	5	6	10	9	7	12	12
AC007272.7	2	201963822	201964264	0.000	4	7	7	6	6	4	7	9	6	6	10	6	10	10
AL359844.1	10	70782804	70783197	0.000	5	12	10	6	4	6	9	13	9	6	11	8	14	14
NAT12	14	57855731	57856072	0.000	8	9	9	6	5	6	7	12	8	10	8	6	4	4
EIF2S2	20	32702015	32702335	0.000	6	10	8	6	9	9	14	15	8	9	11	4	3	3
ACTBP11	1	224052444	224052660	0.000	6	10	9	4	5	11	9	13	12	4	1	17	17	16
PTPRS	19	5341427	5341949	0.000	7	9	5	7	5	5	14	5	7	7	8	8	12	12
AC017104.4	2	232254143	232254467	0.000	5	12	9	10	3	11	12	11	10	5	13	7	15	15
TBX4	**17**	**59531961**	**59532563**	**0.000**	**5**	**13**	**7**	**5**	**4**	**9**	**13**	**9**	**11**	**4**	**12**	**4**	**1**	**0**
PABPCP2	2	147344801	147345449	0.000	8	13	9	10	9	4	13	15	8	9	10	7	15	15
AC113607.1	2	905373	905826	0.000	3	10	17	10	10	5	11	18	16	3	10	10	20	19
TH	11	2192576	2193202	0.000	5	11	6	12	8	2	7	15	4	5	13	5	15	15
SIPA1	**11**	**65408027**	**65408751**	**0.000**	**6**	**9**	**13**	**7**	**4**	**10**	**16**	**9**	**14**	**3**	**9**	**7**	**2**	**1**
GPR39	2	133174634	133175088	0.000	9	14	8	5	4	5	17	5	9	4	17	5	17	17
ING5	2	242665670	242666172	0.000	3	10	2	9	6	8	13	9	7	5	13	5	14	15
C11orf85	**11**	**64739412**	**64739716**	**0.000**	**9**	**8**	**9**	**6**	**4**	**7**	**15**	**11**	**8**	**4**	**11**	**3**	**2**	**2**
AC011530.1	19	46318069	46318747	0.000	6	10	12	9	9	3	13	11	10	3	11	5	3	3
AL139130.1	1	156357691	156358065	0.000	11	8	8	3	1	8	16	10	13	4	8	4	1	0
MUC2	11	1075204	1076015	0.000	4	13	7	7	7	10	8	17	7	8	17	10	16	16
SLC22A20	11	64981227	64981938	0.000	4	8	10	9	8	9	17	12	7	5	13	4	3	3
AP001476.3	21	47457079	47457365	0.000	3	10	6	9	5	7	13	11	7	6	16	4	14	15
C3orf24	3	10149532	10150224	0.000	5	10	7	8	6	11	10	10	15	6	10	7	13	13
AL031296.2	1	12587915	12588225	0.000	7	11	7	5	5	5	12	11	4	3	11	6	12	12
CENPB	20	3765976	3766968	0.000	5	13	7	12	5	12	6	14	9	7	15	4	15	15
CMTM2	**16**	**66613072**	**66613394**	**0.000**	**6**	**12**	**5**	**3**	**4**	**10**	**13**	**8**	**10**	**4**	**12**	**4**	**1**	**2**
ODF3	11	195013	195431	0.000	5	12	7	7	7	10	9	9	7	10	6	11	11	11
TBX4	**17**	**59532564**	**59532803**	**0.000**	**6**	**12**	**3**	**7**	**4**	**8**	**15**	**8**	**7**	**5**	**8**	**5**	**1**	**2**
AL512362.1	14	104917026	104917422	0.000	7	9	7	10	7	7	11	10	7	7	11	5	15	16
GSC	14	95237544	95237981	0.000	5	5	9	4	7	12	6	8	14	12	11	6	2	3
PATZ1	22	31740363	31741170	0.000	7	7	9	5	7	9	16	10	9	3	9	6	2	3
TBC1D3	17	36282803	36283142	0.000	5	11	4	10	4	13	16	8	5	8	11	8	15	14
AL355075.1	14	20903481	20903844	0.000	7	6	8	7	7	6	11	11	6	6	12	6	13	14
SNED1	2	241936268	241936700	0.001	6	8	3	9	9	6	6	13	5	5	9	7	13	12
AP001476.3	21	47456602	47457078	0.001	4	10	6	9	5	7	7	10	6	10	9	6	14	13
AC068993.1	12	79187669	79188167	0.001	4	7	4	8	5	7	9	10	6	5	11	5	10	10
PCK2	14	24562608	24563016	0.001	4	11	2	10	4	6	9	13	8	6	7	6	13	12
AC007189.1	2	49142926	49143500	0.001	8	12	13	11	7	9	21	15	14	8	18	10	19	18
DIP2C	10	737199	737975	0.001	11	17	10	7	6	4	16	11	8	10	5	16	15	15
C20orf197	20	58629936	58630496	0.001	1	10	4	9	7	13	8	10	10	3	9	13	3	2
5_8S_rRNA	16	33964201	33964553	0.001	8	12	8	9	3	5	6	16	11	5	9	7	13	13
AC021016.2	2	219218778	219219347	0.001	8	9	7	6	5	8	10	11	9	6	13	7	13	14
AC010928.2	18	58329757	58330190	0.001	9	13	6	7	3	4	12	6	9	3	7	7	12	12
FERMT3	**11**	**63974229**	**63974772**	**0.001**	**5**	**5**	**5**	**7**	**5**	**9**	**11**	**6**	**8**	**5**	**10**	**7**	**1**	**2**
CHST6	16	75529780	75530072	0.001	2	15	8	4	5	6	6	13	4	5	15	5	14	13
ATL3	11	63439439	63440134	0.001	5	11	7	8	5	5	9	12	7	7	11	6	11	11

RP11-165M6.1	13	107078138	107078692	0.001	12	6	5	6	7	5	12	11	7	7	13	5	12	12
RPL13AP3	14	56233075	56233429	0.001	7	13	10	7	9	8	9	13	12	6	12	9	15	14
OR1M1	19	9204180	9204744	0.001	6	11	7	7	6	4	11	11	10	7	8	4	12	13
AC131097.1	2	242845825	242846198	0.001	7	12	8	3	6	4	7	12	8	5	11	9	12	13
U3	17	42380910	42381319	0.001	6	10	4	6	7	4	8	6	12	6	7	7	4	4
WFIKKN1	16	678644	679024	0.001	7	12	3	9	8	10	12	13	10	7	11	14	15	14
RPS10L	20	820108	820639	0.001	7	9	5	8	4	9	5	9	10	6	10	6	10	10
AL135798.1	1	117284700	117285259	0.001	7	17	11	16	12	9	14	14	15	9	17	8	17	18
AL356957.13	1	149287471	149287898	0.001	10	5	6	6	5	9	9	11	6	5	10	6	3	2
AC008069.2	2	17036367	17036607	0.001	8	11	7	6	7	9	13	12	6	10	13	4	13	14
AL008723.1	22	32665325	32665766	0.001	6	14	9	8	4	11	7	14	7	8	11	6	13	14
ZNF547	19	57873156	57873573	0.001	4	11	11	4	5	5	8	11	10	8	13	3	12	12
IGHA1	14	106174533	106175030	0.001	8	14	6	13	13	7	10	10	16	11	15	7	15	15
BX322557.4	21	46772281	46773134	0.001	6	12	3	11	6	4	9	18	8	7	8	6	13	14
HES5	1	2463378	2464184	0.001	8	13	5	10	9	10	6	14	6	7	9	11	14	13
AP001466.1	21	15308923	15309360	0.001	5	10	4	5	9	9	10	12	7	6	10	3	11	11
C2orf85	2	242812123	242812826	0.001	7	14	7	6	7	9	10	10	14	10	13	7	13	13
AL451069.4	10	134243532	134244554	0.002	9	11	6	7	8	11	7	12	12	5	14	6	13	14
FCN3	1	27702197	27702709	0.002	6	10	6	5	6	6	9	8	6	4	9	7	9	9
AL139188.2	13	30438274	30438770	0.002	6	17	8	7	3	6	8	18	10	9	15	9	15	16
TMEM132C	12	128899599	128900228	0.002	5	13	5	8	4	12	10	13	12	4	10	9	14	13
AC074212.1	19	46236309	46236955	0.002	7	11	10	12	6	6	16	17	11	11	15	7	20	22
C21orf77	21	33948372	33948737	0.002	7	11	6	8	5	9	13	9	4	6	10	10	12	13
ACAP3	1	1246319	1246922	0.002	6	18	10	8	15	9	16	17	14	11	15	8	17	17
ADAD2	16	84224448	84224795	0.002	6	11	10	11	5	5	12	16	10	5	15	9	14	14
KIAA0323	14	24897965	24898530	0.002	4	8	5	9	4	6	6	13	7	7	9	7	10	10
C2CD2	21	43374507	43375097	0.002	4	14	10	9	11	7	10	16	16	8	14	3	15	16
TIMM13	19	2428888	2429827	0.002	10	10	12	10	6	4	16	15	9	9	12	8	14	14
AC110299.1	2	242456556	242457232	0.002	4	8	8	8	7	8	8	9	11	5	9	6	12	13
U1	1	146550207	146550837	0.002	7	18	7	4	4	9	22	15	8	5	8	5	2	3
AP005380.2	18	5132955	5133396	0.002	3	11	6	9	7	3	11	12	10	6	13	4	12	13
SETDB1	1	150896594	150896961	0.002	8	6	3	9	6	4	12	10	8	4	8	4	10	10
AC103563.10	2	95638175	95638602	0.002	3	9	8	6	6	4	6	10	6	9	12	6	10	10
CALM2	2	47405104	47405186	0.002	0	1	1	0	1	0	0	1	0	1	1	1	0	0
MSMB	10	51548149	51548150	0.002	1	0	1	0	1	0	1	0	1	1	1	0	0	0
GLS2	12	56881692	56881936	0.002	0	0	1	0	1	1	1	1	1	0	1	0	0	0
HSD11B2	16	67463366	67463367	0.002	0	0	1	1	1	0	1	0	1	0	1	1	0	0
ETV2	19	36132506	36132707	0.002	0	1	0	1	1	0	1	1	0	1	1	0	0	0
FCRLB	1	161689631	161689644	0.002	0	1	1	1	0	0	1	1	0	1	1	0	0	0
LIPA	10	91175701	91176201	0.002	3	10	7	5	4	8	10	10	7	7	10	5	10	10
SYT14	1	210111479	210111996	0.003	9	4	3	3	4	12	16	12	8	5	4	2	2	1
SKI	1	2157600	2158529	0.003	8	8	4	7	7	8	6	13	6	5	9	7	12	11
ELF5	11	34534937	34535481	0.003	7	13	7	6	4	7	7	15	4	7	16	9	13	13
AC133919.5	16	90160195	90160692	0.003	6	14	8	6	8	6	10	11	11	4	7	7	5	5
AC012075.1	2	81694490	81694825	0.003	5	13	6	9	4	8	8	17	8	6	14	3	14	13
SFRS6	20	42084031	42084724	0.003	7	8	3	5	8	7	5	10	6	9	8	6	9	9
AC006269.1	17	53638377	53639063	0.003	6	16	10	6	7	7	17	15	14	8	13	7	15	16
AC018804.7	2	130986044	130986416	0.003	10	9	7	7	6	4	13	16	10	10	12	5	13	14
AC025279.1	16	29300194	29301153	0.003	5	9	6	8	9	9	4	9	12	12	12	7	13	12
hsa-mir-410	14	101532662	101533132	0.003	3	11	5	12	7	7	5	13	13	6	10	5	12	13
WDR24	16	738987	739624	0.003	19	15	12	14	8	8	15	15	18	8	19	10	18	18
LRRC30	18	7231085	7231660	0.003	6	19	5	10	4	6	13	13	12	8	11	4	15	14
C16orf81	16	89225822	89226056	0.003	5	13	5	10	8	9	7	12	9	3	12	8	12	13
C17orf62	17	80409641	80409982	0.003	7	9	7	4	7	6	9	9	6	7	10	4	5	5
AC080112.1	17	38523323	38523956	0.003	6	9	6	8	7	4	5	13	7	7	10	7	10	10
F2	11	46741219	46741868	0.003	9	15	7	8	8	12	7	15	8	7	11	7	14	13
AL359457.2	13	20134785	20135212	0.003	5	12	8	6	5	6	6	16	10	8	11	7	13	12
AP001266.1	11	65546062	65546392	0.003	9	13	10	9	6	8	13	12	16	12	14	9	14	14
AL356961.2	13	112760878	112761112	0.003	8	4	8	4	1	9	17	11	7	1	9	2	2	1
SYT8	11	1846938	1847982	0.004	8	18	12	17	10	14	10	28	6	9	21	11	20	21
GPR25	1	200842396	200842812	0.004	7	7	8	7	9	14	19	10	7	2	11	10	5	4
C2orf65	2	74875016	74875561	0.004	7	9	6	6	8	8	7	12	7	4	12	5	5	5
CASP7	10	115478883	115479141	0.004	6	9	6	7	7	6	15	8	12	8	9	6	4	5
AL591848.2	1	246954217	246954457	0.004	8	13	6	4	6	7	11	9	8	6	11	4	12	11
hsa-mir-381	14	101512070	101512894	0.004	6	14	5	6	6	7	9	14	13	7	15	9	14	13
PYY2	17	26553901	26554658	0.005	7	10	9	5	9	4	13	14	10	5	11	5	12	13
AP000345.1	22	23909004	23909277	0.005	5	11	8	5	4	7	14	11	8	5	10	6	11	11
CTA-299D3.1	22	48943316	48944061	0.005	3	13	6	10	6	6	12	9	10	9	6	5	11	11
AL122127.9	14	106351596	106351950	0.005	2	8	11	9	5	7	6	10	9	10	8	6	5	5
CHRM4	11	46407278	46407882	0.005	12	16	9	10	11	9	14	11	9	10	13	7	15	14
AP001623.1	21	43720930	43721824	0.005	6	8	8	8	6	7	13	9	7	4	10	5	10	10

Gene	Chr	Start	End	P														
BSND	1	55464529	55465160	0.005	6	9	7	5	6	6	7	13	4	7	8	5	11	10
ZNF570	19	37959287	37959750	0.005	7	15	8	6	4	8	11	13	10	4	7	10	12	13
MAFB	20	39319652	39320415	0.005	8	12	11	6	6	7	17	26	5	7	9	4	4	3
C13orf36	13	37247969	37248446	0.005	11	2	6	5	3	10	2	16	3	4	8	5	2	2
PPPDE1	1	244814626	244815003	0.006	10	7	3	6	8	4	10	9	3	7	11	9	10	10
ZNF583	19	56916205	56916724	0.006	7	10	5	9	12	6	11	8	10	7	12	7	11	11
AP001931.1	11	57520088	57520392	0.006	9	10	11	8	7	10	14	15	12	11	7	5	7	7
TUBGCP2	10	135125479	135125859	0.006	4	9	6	9	6	5	6	8	11	5	11	4	11	10
AQP5	12	50355718	50356251	0.006	3	16	6	9	6	6	9	10	11	9	12	6	13	12
CCDC79	16	66835674	66836220	0.006	7	11	10	8	8	6	8	13	9	6	14	8	13	12
DLGAP1	18	3879613	3879881	0.006	6	26	8	11	7	4	14	21	14	8	12	9	18	18
MASP2	1	11108391	11108763	0.006	7	7	3	9	6	6	8	12	5	8	8	7	5	5
SSU_rRNA_5	21	9826641	9826839	0.006	28	10	16	76	35	5	12	37	12	30	49	25	8	9
FAM38A	16	88804843	88805299	0.006	8	13	7	9	6	12	7	18	9	4	9	11	13	13
CCDC79	16	66835180	66835673	0.007	8	18	13	8	8	7	13	14	11	8	13	7	14	15
AL691429.2	10	134778953	134779462	0.007	6	10	7	9	3	7	7	11	13	6	12	4	12	11
C15orf60	15	73735188	73735531	0.007	6	13	5	10	8	4	12	15	4	10	10	6	12	13
CCL15	17	34330963	34330964	0.007	1	1	0	0	0	1	1	0	1	1	0	0	0	0
RAB22A	20	56884284	56884425	0.007	1	0	1	1	0	0	0	0	1	0	1	1	0	0
RP11-529I10.1	10	103329231	103329589	0.007	0	0	1	1	1	0	0	1	1	1	0	0	0	0
HSD17B12	11	43702084	43702130	0.007	1	0	1	0	0	1	1	0	0	1	1	0	0	0
GYLTL1B	11	45944422	45944514	0.007	1	1	1	0	0	0	1	0	1	1	0	0	0	0
AP003108.2	11	61276076	61276077	0.007	0	0	1	1	0	0	0	1	1	1	1	0	0	0
BCAT1	12	25101998	25102064	0.007	0	0	1	1	0	1	0	0	1	1	1	0	0	0
AC084398.1	12	102323488	102323489	0.007	1	1	0	0	1	0	1	0	0	1	0	1	0	0
AB019437.11	14	107150907	107150953	0.007	0	1	0	1	0	0	0	1	1	1	1	0	0	0
ARNT2	15	80696362	80696410	0.007	1	0	0	1	1	0	0	0	1	1	1	0	0	0
PLD6	17	17109465	17109466	0.007	1	0	0	1	1	0	1	0	0	1	0	1	0	0
MFAP4	17	19290689	19290762	0.007	1	0	0	1	0	1	0	1	0	1	1	0	0	0
SECTM1	17	80291235	80291459	0.007	1	0	1	1	0	0	0	1	1	0	1	0	0	0
MYO1F	19	8644031	8644155	0.007	1	1	0	1	0	0	1	0	1	1	0	0	0	0
C1orf113	1	36772133	36772221	0.007	0	0	1	0	0	1	1	1	0	1	1	0	0	0
CACNA1S	1	201082700	201082731	0.007	0	0	1	1	1	0	1	0	0	0	1	1	0	0
TMEM18	2	678865	678866	0.007	1	0	0	1	1	0	0	1	0	1	1	0	0	0
ZFAND6	15	80350338	80350339	0.007	0	1	1	0	0	1	0	1	0	0	1	1	1	1
U6	1	22314468	22314516	0.007	0	1	0	1	0	0	1	0	1	1	1	0	1	1
SPI1	11	47400930	47401026	0.007	1	0	0	1	0	1	0	0	1	0	1	1	1	1
SLC15A3	11	60720092	60720229	0.007	0	0	0	1	1	0	1	1	0	0	1	1	1	1
AP000770.1	11	116510340	116510341	0.007	1	0	0	1	1	0	0	1	0	0	1	1	1	1
SCARNA11	12	8748456	8748579	0.007	1	1	0	1	0	0	0	0	1	0	1	1	1	1
COMP	19	18903207	18903230	0.007	0	0	1	1	0	0	1	1	0	0	1	1	1	1
INSM1	20	20348794	20348956	0.007	0	0	1	1	1	1	0	1	1	0	0	1	1	1
C16orf81	16	89226057	89226378	0.007	7	12	6	12	8	15	9	11	10	5	15	9	13	14
AC012652.1	15	41521764	41522036	0.007	8	8	2	9	11	5	7	10	5	7	11	7	10	10
CNIH	14	54910018	54910240	0.007	7	13	7	10	3	6	11	13	12	7	15	4	15	17
SPO11	20	55904448	55905206	0.008	6	16	8	9	8	18	7	15	6	11	14	6	18	16
AC018755.9	19	52101660	52102180	0.008	6	15	8	11	11	6	11	22	8	12	14	8	16	15
AC008271.1	2	15830701	15831610	0.008	10	8	7	9	7	13	17	13	11	14	14	9	8	8
TACR2	10	71175640	71176127	0.008	8	8	7	5	6	12	17	12	9	8	8	7	12	13
5S_rRNA	1	228770930	228771604	0.008	14	29	7	6	2	8	32	30	7	16	23	17	31	27
CDKN3	14	54861108	54861777	0.008	8	9	7	5	8	9	10	12	14	6	9	5	11	11
AL356957.13	1	149287899	149288411	0.008	8	9	6	7	4	5	10	9	7	3	15	3	4	4
AC118470.1	1	247802955	247803176	0.009	12	7	5	4	4	7	20	7	6	6	5	2	3	2
AL139161.2	1	236136540	236137128	0.009	6	6	7	5	6	6	9	7	8	7	9	5	8	8
TMEM85	15	34515638	34515959	0.009	3	10	6	7	6	7	7	7	9	11	7	10	11	
MRPL28	16	422537	423002	0.009	9	10	5	5	9	6	8	11	19	6	10	5	13	12
Y_RNA	14	100048449	100049145	0.009	6	14	6	13	8	5	14	13	15	8	7	8	13	14
AL928742.3	14	106004966	106005349	0.010	5	10	7	7	6	13	4	10	16	9	21	5	14	14
CALML5	10	5540692	5541293	0.010	4	15	7	8	8	4	9	15	4	8	16	11	13	13
INHA	2	220431911	220432323	0.010	8	8	4	12	6	10	11	9	11	5	17	4	13	12
GPS1	17	80008024	80008393	0.010	5	10	7	13	9	7	15	13	14	8	13	6	13	14
AC011491.1	19	6378815	6379216	0.011	9	14	7	6	6	5	13	11	7	6	10	4	12	11
AC104841.1	2	242165415	242165901	0.011	7	12	8	8	8	6	14	13	12	8	8	8	7	7
AL391244.1	1	1354451	1355097	0.011	10	14	11	13	8	12	8	17	14	11	18	6	15	15
ESPNP	1	17046419	17046814	0.011	8	15	10	7	6	7	11	10	8	6	19	10	17	15
RAB1B	11	66033857	66034662	0.011	11	9	4	10	8	8	11	19	13	7	9	7	17	15
TMEM79	1	156251158	156251235	0.012	0	1	1	0	1	0	1	1	0	0	2	0	0	0
GGT5	22	24642529	24642530	0.012	0	1	0	1	1	0	0	2	0	1	1	0	0	0
ALKBH7	19	6369985	6370742	0.012	6	7	7	7	6	6	11	8	7	7	11	4	9	9
BAIAP2	17	79006444	79007013	0.012	11	19	12	6	7	10	17	18	17	8	14	7	16	17
AL136038.2	14	64061697	64062069	0.012	11	22	8	9	10	5	13	17	19	13	16	5	17	17

6

131

540123-L-bw-Clausen

Processed on: 16-1-2020

AL358176.2	1	240799808	240800561	0.012	3	12	5	8	5	9	14	15	7	10	11	5	12	12
TBX4	17	59533693	59534236	0.012	6	11	9	6	5	7	15	11	10	6	13	5	3	1
GZMM	19	543219	543909	0.012	5	10	4	5	6	6	13	7	7	6	8	5	9	9
CATSPER1	11	65793385	65793796	0.013	8	11	8	10	9	17	14	9	11	10	12	6	14	13
BACH1	21	30670042	30670853	0.013	8	8	4	9	8	5	5	14	10	7	7	8	10	10
P4HA3	11	74022781	74023219	0.013	3	5	6	7	9	8	10	9	9	5	11	4	5	5
MSGN1	2	17997725	17998458	0.014	6	8	5	5	11	9	8	10	10	6	7	6	10	11
BTBD6	14	105713084	105713865	0.014	8	17	10	9	7	7	7	14	16	8	11	6	14	13
AL359737.3	13	19173660	19174402	0.014	5	8	5	10	5	6	10	11	10	7	6	3	5	5
PLEKHN1	1	900206	900744	0.015	8	12	7	6	7	8	11	11	6	10	11	7	12	11
GDF2	10	48416585	48416977	0.015	2	11	10	12	5	5	10	10	7	7	11	3	13	15
RRP15	1	218457065	218457480	0.015	9	6	7	6	7	6	15	10	7	3	9	4	11	10
RASGRF1	15	79382595	79382766	0.015	8	1	0	2	2	9	11	11	1	1	3	2	1	0
AC104024.2	17	16884183	16885006	0.015	3	7	6	7	5	8	9	9	5	8	9	4	10	9
FAM108A6	22	22471677	22472380	0.015	6	10	3	9	5	11	17	6	9	7	13	4	5	5
NKX2-2	20	21496275	21496756	0.016	0	1	2	6	4	10	1	12	0	6	10	2	1	1
GBP5	1	89739002	89739591	0.016	2	13	10	9	6	4	13	15	11	4	8	9	12	12
RASGRP1	15	38857242	38857791	0.016	7	6	3	11	8	8	10	14	13	5	9	7	12	11
ATHL1	11	289630	290144	0.016	6	14	7	15	7	12	5	14	15	9	18	6	7	7
AL451043.2	1	147716091	147716663	0.016	13	21	13	11	13	10	29	16	18	16	25	15	21	22
RP4-697K14.1	20	62199420	62200049	0.016	10	13	11	6	5	12	14	14	9	4	13	7	3	5
RUSC1	1	155290535	155291063	0.016	5	1	6	6	3	12	1	7	12	5	13	3	3	2
AC124861.4	2	241196681	241197295	0.017	7	16	9	7	7	14	12	17	13	8	13	10	14	14
USP18	22	18631319	18631711	0.017	9	14	6	12	7	9	12	10	7	8	9	8	7	6
AC009237.7	2	96190971	96191608	0.017	8	18	11	11	5	9	15	13	13	9	19	7	15	15
FAM3B	21	42675620	42676210	0.017	9	13	6	12	3	4	21	16	6	10	9	3	14	16
C11orf85	11	64739717	64740014	0.017	5	9	9	6	5	8	16	10	6	5	8	4	4	5
MAD2L2	1	11751298	11751523	0.017	0	0	1	1	0	1	1	0	0	1	1	1	1	1
WFIKKN1	16	680974	681202	0.017	1	0	1	0	0	1	1	0	0	1	1	1	1	1
IRX6	16	55357787	55358034	0.017	1	0	1	1	0	0	1	0	1	1	0	1	1	1
FAM71E1	19	50978981	50980006	0.017	6	7	10	9	9	5	13	8	7	9	15	4	12	11
FKBP4	12	2901927	2902637	0.017	8	11	7	8	5	8	7	11	7	7	14	6	6	5
C16orf81	16	89225495	89225821	0.017	2	13	6	9	8	12	5	18	15	2	16	7	16	14
FAM92A2	15	41455529	41456044	0.018	5	10	6	10	3	6	9	10	8	5	8	4	9	9
RPL12L3	20	19804150	19804670	0.018	3	11	8	7	3	14	4	17	4	7	8	7	12	11
AC138969.3	16	16459071	16459683	0.018	0	10	1	15	5	14	5	9	17	6	19	12	14	15
FSCN2	17	79492961	79493634	0.018	5	6	2	9	9	7	7	8	10	8	14	7	10	10
C1orf159	1	1053304	1053617	0.018	4	14	8	7	5	5	6	17	11	8	8	9	13	15
RBP3	10	48389910	48390847	0.018	6	15	7	11	9	12	17	13	11	6	10	6	14	13
SLA2	20	35274515	35274715	0.019	3	3	2	2	0	10	3	14	0	1	8	4	0	1
COX6A2	16	31439306	31439752	0.019	8	5	10	7	8	4	6	10	11	8	11	6	10	11
NACA2	17	59668192	59668741	0.019	6	10	6	10	8	5	14	10	12	8	13	9	7	7
C13orf35	13	113299262	113300283	0.019	15	8	10	8	6	9	15	11	6	6	12	7	12	12
AC015651.1	17	61926521	61927086	0.019	5	12	2	7	6	9	9	14	9	7	14	5	11	12
JMJD4	1	227921399	227921906	0.019	3	6	6	7	7	14	13	9	12	5	12	6	11	12
GALNT13	2	154728002	154728308	0.020	9	1	3	1	3	4	14	7	0	4	2	2	1	1
RNF17	13	25337805	25338421	0.020	5	9	10	7	4	6	8	16	8	5	14	8	12	11
AC112777.1	12	20704358	20704532	0.020	9	14	9	21	5	10	8	13	13	6	33	5	19	18
C21orf33	21	45551474	45551798	0.020	3	10	4	14	6	5	10	14	8	4	13	6	12	11
GP1BA	17	4836017	4836468	0.021	8	9	3	9	7	5	16	11	6	8	13	5	12	11
AL592464.2	1	2729504	2730299	0.021	8	17	8	11	8	8	20	18	13	9	14	4	16	15
SNX32	11	65601265	65601550	0.021	0	2	0	3	2	14	0	13	3	2	7	0	0	0
KAT2A	17	40274881	40275820	0.021	9	10	11	12	7	9	11	14	7	6	14	6	15	17
NNAT	20	36149825	36150208	0.021	7	12	7	7	3	11	11	20	10	8	8	5	5	6
KIAA0562	1	3774998	3775624	0.021	6	15	3	8	12	8	5	12	5	9	16	10	13	12
AL117692.1	14	50519321	50519571	0.021	6	9	5	5	9	3	7	12	7	5	10	4	10	9
AL109945.1	1	32815223	32815649	0.022	8	7	6	5	11	10	10	10	9	8	8	5	6	5
hsa-mir-380	14	101491469	101492406	0.022	5	15	7	8	9	12	10	15	13	7	16	6	13	14
GRK1	13	114321368	114322096	0.022	6	15	9	8	5	8	10	19	12	8	12	7	13	13
TCL1A	14	96179944	96180575	0.022	10	6	8	5	5	11	9	17	7	8	8	7	13	15
snoU13	17	77685085	77685964	0.022	3	11	4	8	4	10	7	10	9	5	15	6	12	14
AMN	14	103388267	103388745	0.023	5	12	5	5	6	9	3	14	8	7	11	6	10	11
DSCR4	21	39493391	39493628	0.023	6	12	4	8	9	7	12	15	8	9	6	8	11	11
ASPA	17	3375365	3375814	0.023	8	8	6	6	6	8	15	11	7	8	12	4	13	15
NTN5	19	49176094	49176747	0.023	10	8	12	7	12	8	16	17	8	8	16	8	17	15
NAV1	1	201591881	201592262	0.024	10	20	5	11	7	14	18	10	8	4	16	5	6	7
AC004448.7	17	19396521	19397115	0.024	4	7	8	8	6	5	8	5	6	7	10	7	8	8
NEFH	22	29876170	29876886	0.024	9	12	11	8	7	12	21	16	11	5	10	9	8	7
AL158216.1	1	42506718	42507155	0.025	9	10	4	5	6	10	15	8	6	8	4	10	11
PAOX	10	135193152	135194112	0.025	3	8	15	18	10	9	10	19	11	6	10	4	14	14
AC068134.5	2	233252919	233253570	0.025	5	13	8	6	9	8	14	14	8	8	13	2	13	15

MBD3	19	1593742	1594594	0.025	7	7	7	10	7	3	16	12	6	5	14	6	11	12
AP002347.1	11	59665177	59665838	0.025	10	14	10	8	5	7	16	10	8	11	9	10	13	12
P4HA3	11	74021586	74022694	0.026	4	5	9	7	7	9	11	8	10	6	15	9	13	15
ACTRT2	1	2936088	2936661	0.026	5	12	5	5	5	11	12	10	8	3	10	6	12	14
EP400NL	12	132567723	132568038	0.026	5	7	7	10	4	3	11	8	5	8	11	5	5	5
SFT2D3	2	128456522	128456875	0.026	7	20	10	6	11	8	8	15	10	7	19	9	14	15
PROKR2	20	5294594	5294876	0.027	5	19	6	9	9	14	13	15	14	7	12	7	14	14
AC018731.1	2	152042525	152042811	0.028	5	11	10	7	4	7	14	6	8	9	8	7	6	6
AC105272.1	1	104112490	104113282	0.028	8	9	5	5	11	10	9	16	8	5	9	10	11	11
AL034420.1	20	50481186	50481798	0.028	6	12	9	6	8	4	12	10	9	9	8	4	13	15
WDR90	16	697261	697875	0.028	9	11	6	8	5	15	4	11	14	8	12	6	15	13
FAM83E	19	49116077	49116544	0.029	4	17	8	7	6	3	7	16	4	8	7	7	11	11
AP001187.6	11	64658499	64658834	0.029	5	16	9	8	7	9	5	15	10	12	8	10	12	12
PROX1	1	214156052	214156507	0.029	5	2	5	6	4	13	20	10	1	3	11	4	3	1
hsa-mir-663	20	26188963	26189097	0.029	75	40	79	140	38	41	26	34	43	68	78	47	21	2
AL049812.1	20	40626799	40628118	0.030	14	25	14	14	14	9	23	19	13	8	19	9	19	19
MBD1	18	47808654	47809093	0.030	5	16	7	11	5	8	10	9	8	4	18	5	12	12
NPAS4	11	66188392	66189250	0.030	4	5	8	8	3	15	14	11	16	5	15	2	5	3
NTSR2	2	11809606	11810729	0.031	11	13	13	10	8	5	14	13	6	8	16	6	16	14
AL035669.3	20	61406620	61407125	0.032	7	9	4	9	3	8	9	11	5	3	9	8	10	9
AP001476.3	21	47455564	47456395	0.032	2	12	10	17	8	5	9	11	10	8	11	9	15	13
KRT85	12	52760680	52761247	0.032	6	12	5	11	7	9	20	14	10	12	17	4	14	14
5S_rRNA	12	34358079	34358737	0.033	5	12	9	5	5	11	14	17	11	7	12	5	13	15
UTS2R	17	80332010	80332560	0.033	7	15	5	15	10	11	16	15	10	9	17	7	15	17
MYEOV	11	69061709	69062020	0.033	1	10	12	4	5	4	5	6	5	6	27	7	13	12
AC010528.1	16	76268977	76269409	0.033	8	11	7	9	7	8	12	13	9	7	13	6	11	12
PSMA8	18	23713594	23714084	0.034	8	15	7	8	6	6	16	20	6	11	7	13	13	
TUBB6	18	12306268	12306837	0.034	8	11	7	6	6	2	8	11	3	7	14	4	5	4
CEACAM16	19	45199937	45200643	0.035	9	6	10	8	8	10	6	9	8	8	16	8	6	7
AL357712.1	10	8203710	8204202	0.035	6	10	6	7	8	4	10	9	5	9	9	4	10	9
MRPL20	1	1343891	1344780	0.035	4	15	7	6	8	4	15	10	7	11	10	7	11	12
AC093393.1	2	33952332	33952821	0.035	5	11	10	7	6	10	11	12	12	5	10	6	15	13
AL122018.1	1	236273020	236273412	0.036	6	12	5	9	8	7	11	9	10	5	12	7	15	13
RP11-56M3.1	10	92913356	92913775	0.036	9	13	11	9	6	9	15	7	14	8	12	8	13	12
AC008993.3	19	93193	93664	0.036	8	16	6	7	8	8	13	16	11	12	10	1	6	7
hsa-mir-663	20	26188638	26188962	0.036	76	42	78	139	39	40	24	33	41	66	79	44	21	1
AL391244.1	1	1353425	1353858	0.036	2	8	8	7	5	8	9	10	5	7	11	8	9	9
TM7SF2	11	64878569	64879196	0.037	5	19	5	10	7	11	15	16	13	9	13	9	14	14
AL358237.2	20	58662433	58662514	0.037	0	1	0	6	0	0	6	0	1	0	3	0	3	3
AC215219.3	12	94127	94541	0.038	6	15	5	10	11	4	13	9	15	5	18	4	18	15
C10orf139	10	1205273	1205468	0.038	5	12	8	9	6	3	13	7	6	8	17	6	11	11
MSLN	16	834001	834714	0.038	3	12	9	11	7	5	5	9	10	9	12	4	11	10
AL355376.2	10	29084569	29085115	0.038	7	10	7	6	6	3	10	8	7	5	12	9	14	12
HTR6	1	19991919	19993142	0.038	2	8	4	12	6	11	12	15	7	5	12	5	11	11
RPS6KB2	11	67194641	67195339	0.038	4	12	9	8	5	2	8	12	4	7	11	7	13	11
KRT71	12	52946425	52946854	0.038	5	16	3	8	10	6	15	6	10	7	16	5	12	12
GPHA2	11	64702199	64702982	0.039	6	13	10	10	7	15	13	21	17	10	15	11	7	9
AC012075.1	2	81694040	81694489	0.040	12	17	12	7	6	8	14	15	13	7	20	9	15	17
CDK2AP1	12	123758033	123758565	0.040	6	9	5	7	5	9	8	11	16	5	9	6	10	11
AC007248.2	2	102866968	102867275	0.040	8	12	5	5	6	10	10	16	8	12	17	5	6	7
AGRN	1	953352	954148	0.041	3	9	8	7	6	9	10	9	12	11	8	8	10	11
ZNRF4	19	5455175	5455830	0.041	7	11	10	13	8	4	19	17	11	8	11	5	14	13
PHACTR4	1	28695091	28695513	0.042	4	9	5	7	9	3	10	12	4	5	14	4	5	4
AC022748.1	15	79042124	79042505	0.042	12	17	5	10	5	9	9	15	15	9	11	12	13	14
CIRBP	19	1268296	1268904	0.042	7	11	6	11	6	6	6	9	5	9	19	10	12	14
MLNR	13	49794460	49795110	0.042	11	3	12	5	6	6	8	19	7	11	6	5	5	3
FSIP2	2	186603232	186604189	0.043	3	9	7	7	6	5	11	11	9	4	13	4	4	2
FGF3	11	69633336	69634068	0.044	2	2	12	5	1	7	9	6	10	7	38	3	2	2
CLSTN3	12	7280735	7280996	0.044	6	10	9	9	11	7	11	11	11	8	17	5	15	13
DHODH	16	72041197	72041688	0.046	5	8	5	9	6	7	10	6	10	8	11	4	10	9
AC022400.1	10	75491360	75491675	0.046	10	15	10	12	7	5	19	12	10	16	17	6	9	8
KRT33A	17	39506596	39507113	0.046	8	13	6	4	10	8	15	9	8	3	14	8	11	12
AL512638.1	1	115826147	115826583	0.047	6	11	7	6	6	6	8	11	7	5	11	7	9	9
AC092810.2	1	209405064	209405472	0.047	5	16	5	7	7	7	8	15	5	4	13	5	3	5
TNNT3	11	1940716	1941338	0.048	9	10	10	6	5	9	6	12	4	10	12	7	14	12
HSF5	17	56565440	56565821	0.048	11	11	9	9	9	4	10	11	13	6	14	4	6	4
AP002748.2	11	66304685	66305454	0.048	6	11	5	7	4	10	10	5	8	9	16	5	6	5
FAM100A	16	4665443	4666047	0.048	3	11	9	15	5	7	6	9	11	12	13	6	12	11
EDARADD	1	236511487	236512106	0.049	4	13	12	7	6	8	18	16	11	7	14	5	13	13
FAM21C	10	46220649	46221042	0.050	3	5	6	9	5	11	8	9	13	8	10	5	11	13

6
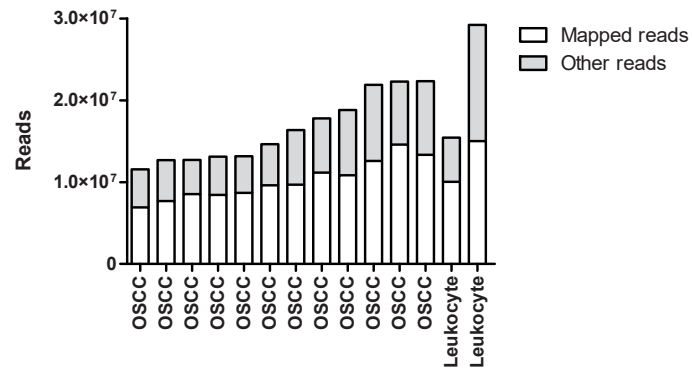
Supplementary figure 6.1. The number of total sequenced reads by MethylCap-Seq and the mapped reads for all included OSCCs In total 11.6 to 22.3 x 10^6 reads were sequenced for each OSCC. Between 6.91 to 14.6x10^6 were mapped back to the genome (60-67 %). For the Leukocytes 10.0 to 15.0 x 10^6 reads were mapped to the genome (51 to 65 %). Data from the MethylCap-Seq analysis were reported previously [279], [306].
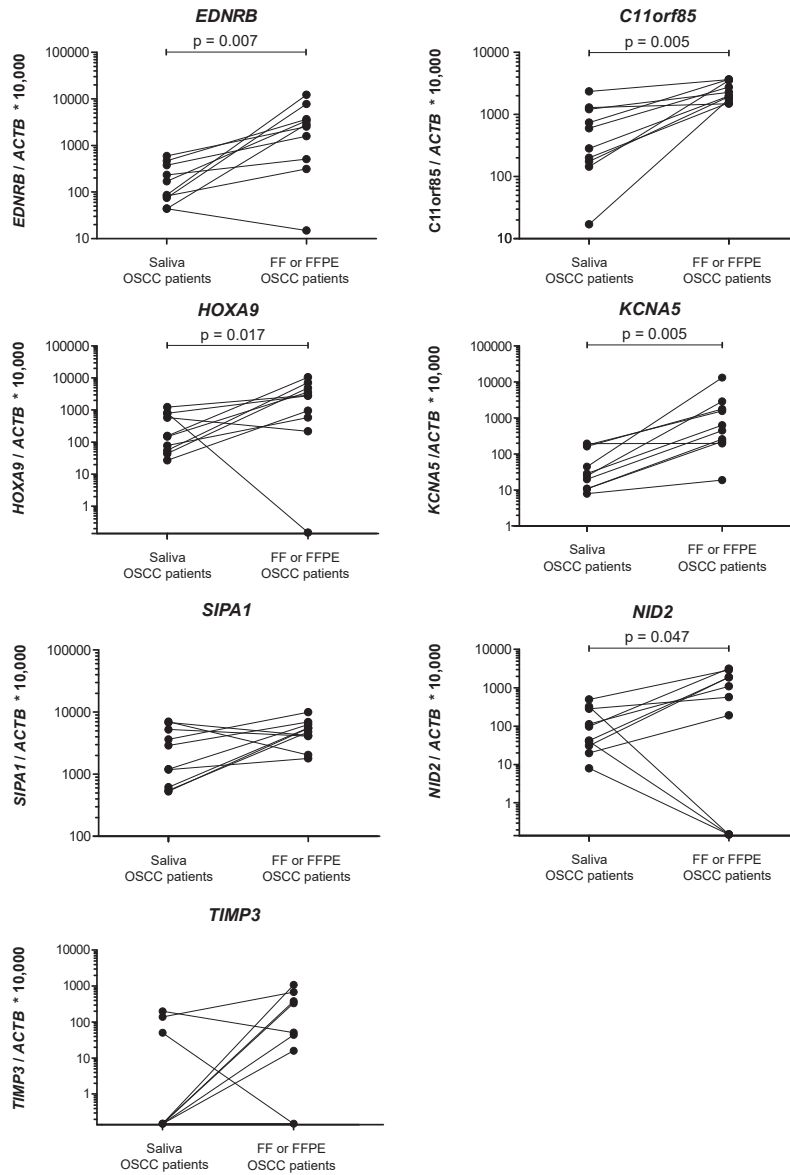
Supplementary figure 6.2. DNA methylation levels of seven OSCC specific markers in saliva and tumor tissues of OSCC patients
Differences in methylation level of the markers between DNA isolated of saliva (saliva patients), fresh frozen (FF tissue) and formalin fixed paraffin embedded tissue (FFPE tissue) of oral squamous cell carcinoma (OSCC) patients. Methylation levels on the x-axis are defined as the average DNA quantity of the gene of interest divided by the average DNA quantity of ACTB and then multiplied by 10,000. Saliva and tumor samples from the same patient are connected by a continuous line in the figure. Tumors were defined as methylated if methylation was present in FFPE or FF tumor tissue. Differences between saliva, FF or FFPE were compared using the Wilcoxon rank test, only significant differences (p < 0.05) are shown.

6

Chapter 7

General Discussion

In recent years great progress has been made in improving detection of LN metastasis in OSCC using clinical and histological parameters in the primary tumor. However, the detection of the presence of LN using these parameters, cannot be applied on a specific subset of OSCC that appear to behave differently than the majority of the OSCC. OSCC of this particular subset initially appear as low risk and as clinically negative for nodal spread, but eventually develop LN metastases. Since these OSCC cannot be distinguished from the other cN0 OSCC using the current clinical and histological parameters, other biomarkers might prove to be helpful in identifying this subset of tumors that potentially develop subclinical metastases in the neck ensuring the proper elective treatment for these LN metastases.

In this thesis the identification and application of markers regulated by DNA methylation for use as predictor for biological behavior of Oral Squamous Cell Carcinomas is reported.

In chapter 2, we first report on the evolution of DNA methylation markers from the literature associated with LN metastasis in our cohort of OSCC [278]. From a panel of 28 genes only the methylation status of DAPK and MGMT were predictive for pN-status of the neck in OSCC. Although the negative predictive value of the methylation status of DAPK1 and MGMT combined was 76%, it did not outperform current clinical nodal staging techniques such as sentinel lymph node biopsy with a reported negative predictive value 88 to 95% [25], [26], [193], [219]. DAPK1 and MGMT are two of the most widely studied methylated genes and have been reported the most for clinical application. However, these two genes aren't sufficient to solve the clinical negative neck dilemma [209], [395], [396]. Thus, to improve upon the predictive value using a DNA methylation gene panel predictive for N-status in OSCC, new DNA methylation markers are necessary and genome-wide discovery approaches are needed to identify these novel DNA methylation markers.

Using genome-wide discovery MethylCap-Seq analysis followed by statistical analysis as well as in silico validation, we identified and characterized three different methylation markers predictive for the presence of lymph node metastasis in OSCC: WISP1, RAB25 and S100A9. WISP1 was found to be hypomethylated and overexpressed in pN+ OSCC (**Chapter 3**), while RAB25 (**Chapter 4**) and S100A9 (**Chapter 5**) are hypermethylated and down regulated in pN+ OSCC. WISP1 and S100A9 aberrant protein levels and DNA methylation levels were also found to be related to patients' survival. Our analyses revealed that all three DNA methylation biomarkers can be used to improve current diagnostic modalities and potentially provide additional treatment options. Finally, WISP1 is involved in the Wnt-pathway, RAB25 in the Raf/MEK/ERK pathways and S100A9 is a well-known calcium-binding protein, and calcium levels are both involved in the Wnt-pathway and the Raf/MEK/ERK pathways.

Unfortunately, the studies reported in this thesis revealed that the predictive value for the individual patient of these newly discovered biomarkers is not outperforming current diagnostic modalities [77], [397], [398]. Especially, when comparing the negative predictive values of these DNA methylation biomarkers to recent advancements as the Sentinel Node Biopsies [25], [26]. Summarizing, the methylation status of WISP1, RAB25 and S100A9 were found to be not clinically applicable biomarkers as compared to for example MGMT is in glioblastomas [114]. There are several potential reasons for why the impact of methylation of these three genes is not as predictive for the phenotype of OSCC as that of MGMT in glioblastomas.

First of all, the OSCC methylome is more variable between OSCC cases compared to other cancers. This broader variation in methylomes could be due to OSCC life-style related risk factors which drive carcinogenesis also happen to influence genome-wide methylation levels. The main tumorigenic factors in OSCC are cigarette smoking and alcohol consumption [2]. The combination of both drinking and smoking is especially hazardous and synergistic [2]. Interestingly, both alcohol consumption and cigarette smoking have been shown to be related to global DNA hypomethylation in head and neck cancer which is one of the new generation of hallmark of cancer as defined by Hanahan and Weinberg in 2011 [97], [375], [399]. More specifically, both alcohol and cigarette smoke are known to cause inhibition of the DNMT proteins responsible for DNA methylation maintenance, causing passive global demethylation (reviewed by [400]. For example, alcohol consumption is known to down regulate the mRNA levels of DNMT3a and DNMT3b [401]. In addition, DNMT1 activity can be down-regulated as a result of interaction with ethanol resulting passive genome-wide demethylation (reviewed by [401]). Specifically, alcohol consumption was found to be related to hypermethylation of several genes involved in HNSCC such as CDKN2A and E-cadherin [402]. Besides alcohol, DNMT1 activity is also downregulated by tobacco smoke contributing to the same passive genome-wide hypomethylation as influenced by ethanol (reviewed by [403],[404]). These aberrations of a normal methylation machinery is exemplary of a carcinogenic methylome [403]. Global hypomethylation might be further stimulated by cigarette smoking through DNMT3b downregulation [405]. Finally, several studies report cigarette smoking-induced hypermethylation of several specific genes in different types of cancers [406], [407]. The direct effects of the etiologic effects of OSCC, and especially the effect on the DNMT family, could cause a wider variability of methylation changes in OSCC that drive tumor growth compared to other cancers in tissue less frequently exposed to cigarette smoke and alcohol.

The diversity of carcinogenic aberrations in oral squamous cell carcinoma is further supported by the presence of field cancerization in the oral cavity [3]. Field cancerization is the process where several (pre-)neoplastic lesions develop simultaneously at multiple locations, under the assumption that these premalignant sites develop independently [3]. Field cancerization explains the occurrence of second primary tumors in this area. This field-effect illustrates the diversity and scale of the collection of genetic aberrations in the oral cavity in relation to the etiology of the variability in OSCC methylomes. The high impact on DNA methylation of alcohol consumption and cigarette smoking also explains the absence of a universal marker like MGMT in glioblastoma because risk factors for glioblastomas are radiation exposure and certain genetic syndromes which could cause more homogenous drivers of carcinogenesis [408].

A second reason for not finding a universal methylation marker associated with pN-status in OSCC is the focus on more conventional genomic loci as well as large changes in epigenetic changes in our experimental design. Traditionally, methylation studies focused on changes in methylation levels in CpG rich regions, specifically CpG islands and especially those CpG islands associated with Transcription Start Sites. To prevent being limited to CpG islands we chose MethylCap-Seq analysis, a genome-wide platform as a discovery tool for DNA methylation biomarkers [136]. However, the protein that was used for methylated DNA fragment enrichment (MBD2) only enriches DNA fragments that have a certain minimum amount of methylated CpG within a certain minimum distance of each other [136]. This means that certain

7

DNA fragments with either a single or just a few methylated CpG or with too much space between these methylated CpGs, could be missed by this MBD2 enrichment. The potential impact of this MBD2 bias for CpG rich DNA fragments has become more apparent over the last few years due to the discovery of the biological relevance of DNA methylation of CpG-poor regions. Originally, DNA methylation research was mainly directed at CpG Islands. However, meta-analysis of Illumina 450k data revealed high transcriptional regulation by DNA methylation in CpG island neighboring regions. Specifically, DNA methylation of regions up to 2000 bp adjacent on either sides of CpG islands, referred to as CpG Island Shores, were identified as having a strong regulatory function on gene expression [193]. Additionally, the next 2000 bp further upstream and downstream of these CpG Island Shores were also found to be involved in epigenetic regulation of gene transcription and were dubbed CpG shelves (Figure 7.1) [246].

Figure 7.1. The relative locations of CpG shelves, CpG Shores and CpG islands. CpG islands are genomic regions that confirm to certain statistics that define these loci as especially CpG site rich. The adjacent 2 kb upstream and downstream of these CpG islands are referred to as CpG Island Shores. The next 2kb further downstream or upstream of CpG islands and CpG islands shores are called CpG shelves. Any regions further located from CpG islands are referred to as the Open Sea.

Besides regions of differential methylation, some gene regulatory regions were identified that across different normal tissues have very stable methylation levels. These regions have been called CpG Canyons and CpG Ravines [409], [410]. The consistency of these methylation patterns across tissue suggests an important biological function of these regions and hypo- or hypermethylation of these regions could be especially pathogenic.

Besides the before mentioned technical bias of MDB2 to more highly methylated fragments, we have introduced biases towards conventional regions of DNA methylation in our algorithms and in silico analyses. First of all, we selected regions located 2000 bp upstream to 500 bp downstream of TSS. Additionally, we focused on differential methylation of Methylation Cores identified in the map of the human Methylome [168], [382] The recent addition of these Shores, Shelves, Ravines and Canyons to the targets of DNA methylation research expand and complicate the selection of DNA methylation techniques. Moreover, enhancer sequences, which lie distal to TSS but still impact gene expressions can also be epigenetically regulated [411]. Their distance to CpG islands associated with genes complicates the epigenetic regulation of a gene and could be easily overlooked in experiment design and data analysis [412]. While the ENCODE [413], [414] and FANTOM5 [415] projects have made leaps forward in identifying these distal regulatory regions, these enhancers are often not included in methylation studies. For example, the widely used Illumina 450k lacks these probes covering these long-distance enhancers [412]. In our experimental design we focused only on the transcriptional impact of methylation of regions in the proximity of gene transcription start sites. So, most studies might only cover a very limited part of any the epigenetic iceberg.

Thirdly, our studies focused on methylated cytosines as a very stable compound. However, the methylation of cytosines is a much more complicated process. The epigenetic machinery consistent besides DNA methyltransferases also of the TET protein family. The three ten-eleven translocation (TET) proteins, TET1, TET2 and TET3, facilitate the oxidation of methylated cytosines to hydroxymethyl-cytosines, formyl-cytosines and carbo-cytosines (reviewed by [416], [417]), which are intermediates between methylated and unmethylated cytosines (Figure 7.2).
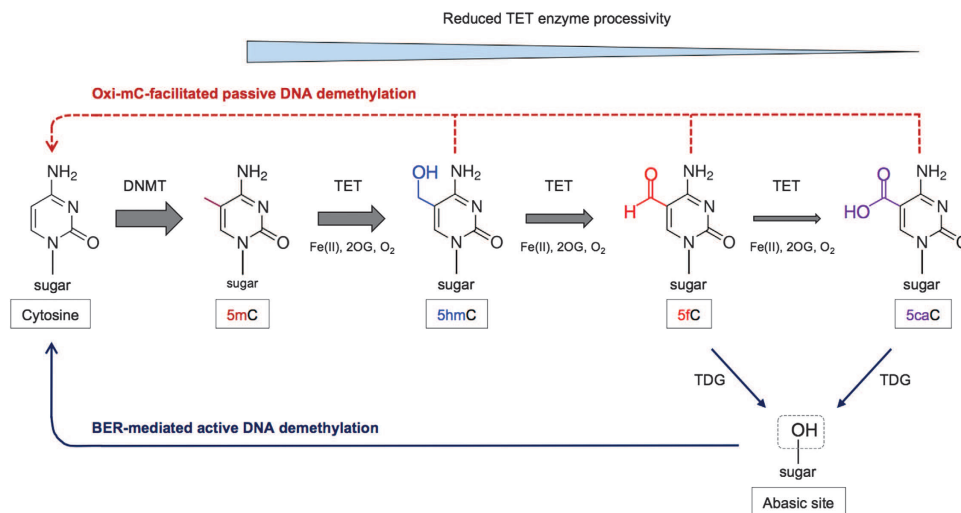
Figure 7.2. TET protein mediated methylated cytosine (5mC) modifications. The family of TET proteins further modify methylated cytosines (5mC) to hydroxymethyl-cytosines (5hmC), formyl-cytosines (5fC) and carbo-cytosines (5caC). Both 5fC and 5caC are intermediates for active demethylation of cytosines by base excision repair (BER). Adapted from [418].

Interestingly, there is much support of a role of TET proteins in carcinogenesis. TET proteins are frequently mutated in tumors and TET proteins has been shown to function as tumor suppressor genes [94], [416], [417], [419]. Regardless of their role in tumorigenesis, TET proteins are often neglected in DNA methylation studies. While occurrences of these TET mediated cytosine modifications is much more rare than DNA methylation, about 0.03 to 0.7% of all cytosines, compared to a 5% of methylated cytosines [420], a major concern is that the widely used bisulfite-based methods of methylation detection such as MSP and bisulfite Sequencing do not differentiate between methylated and the three TET catalyzed intermediates as all different modified cytosines are converted to uracil by bisulfite treatment [421]. Moreover, in general the effects of DNA methylation and the TET DNA methylation intermediates on gene expression are opposite, causing inaccurate methylation detection as well as wrong conclusions of the effects of the detected methylation of biological behavior [417]. In contrast, MethylCap-Seq analysis does distinguish between these different cytosine modifications because the MBD2 enrichment is only specific for methylated cytosines. However, because we performed further gene validation by MSP and since bisulfite Sequencing can't distinguish between the different cytosine modifications [422], the presence of these non-methylation modified cytosines might have distorted the relation between methylation and protein expression found in our studies resulting in an overestimate of the actually DNA methylation levels and

7

thus underestimate of the epigenetic regulation of the studies genes.

In general DNA methylation of gene regulatory sequences results in the down-regulation of the associated genes' expression. However, the relation between methylation and gene expression is further complicated by the observation that some genes appear to be upregulated by hypermethylation. For example, the hTERT gene, a gene involved in regulating telomerase activity, has been found to be upregulated by an increase in DNA methylation [423], [424]. Furthermore, traditionally, gene expression regulation by CpG site methylation had been thought to be confined to the CpG islands. However, as discussed above studies have shown that methylation of CpG island shores might have an even greater influence on gene expression regulation [193]. The influence of DNA methylation on gene expression has been shown to also be capable of long-range interaction between CpGs and target genes [425], [426]. These findings contest the traditionally model that CpG methylation influences gene expression solely by methylation of CpG islands that overlap with regulatory regions. More importantly, a recent genomic study reported that the influence of long-range DNA methylation on its long-range target genes is far greater than the influence of CpG island methylation [427]. Additionally, the same study that found tumor and normal tissue display far greater differences in methylation levels of the long-range regulatory regions [427]. Therefore, it is not prudent to assume that CpG Islands always influence the gene that is closest. To properly identify epigenetic regulation of genes by DNA methylation, direct correlation of mRNA levels and DNA methylation is insufficient. This might explain why we did not find a significant correlation between RAB25 expression and RAB25 TSS methylation. Incorporation of long-range regulatory regions data should be part of algorithms for the identification and characterization of DNA methylation biomarkers.

Besides huge differences between different primary tumors, within a single tumor also lots of variability are present due to tumor heterogeneity. A single tumor can originate from various sub clones that individually collected different genetic aberrations. Differences in these carcinogenic drives can cause variability in molecular and biological differences within a single tumor [428]. A recent study found that the amount of tumor heterogeneity itself is correlated with the presence of LN metastases [429]. Even though, our studies took tumor heterogeneity into account during biomarkers validation by taking at least three different biopsies from a single tumor sample based on HE-coupes for our tissue microarray studies, it's not possible to completely compensate for tumor heterogeneity. The true variability within a tumor is unclear and difficult to assess, and probably would require sequencing of many different tumor subpopulations acquired by macro-dissection to identify clonal subpopulations [430]. Possible heterogeneity of the tumor DNA used for our MethylCap-Seq could have obscured consistent methylation data due to the presence of several tumor subclones. The chance of heterogeneity in OSCC is especially high due to field cancerization as discussed earlier [3]. Because of the nature of the etiology and risk-factors of OSCC such as tobacco smoking and alcohol consumption, the whole oral epithelial is simultaneously subjected to the same risk factors causing a lot of different pre-cancerous genetic aberrations can arise within the oral epithelial layers increasing the risk of tumor subclones.

Finally, the microenvironment of a cancer is an important contributor to the tumorigenesis and might provide another variable to connect our OSCC methylome with the metastatic phenotype [431]. The

influence of macrophages, fibroblasts and endothelial cells are known to partially drive the nodal spread of OSCC, through direct cell-to-cell interactions and paracrine signaling (reviewed by [431]). Additionally, it is possible that these cells were present in the OSCC tissue used for MethylCap-Seq analysis we cannot exclude any influence of these cells in the assessment of our methylome.

While OSCC methylation markers are regularly reported in the literature, a universal OSCC N-status methylation marker has currently not been found yet. Major bottleneck in the process of marker identification is the intra-center validation. Often major variations, improper or incompletely reported in patient population and selection, technique selection and execution, sample size or experimental design prevent successful marker validation. Recent endeavors such as The Cancer Genome Atlas, The International Cancer Genome Consortium, and the Gene Expression Omnibus have provided platforms to make intra-cohort validation easier [409], [432], [433]. Both databases provide huge amounts of publicly available datasets as well as analysis tools for easy validation of markers [434], [435].

A possible biological model for the OSCC metastatic methylome

We report on three different approaches for the identification and selection of DNA methylation markers that are predictive for the presence of nodal metastases in OSCC. All three approaches have led to the identification of a different biomarker: WISP1, RAB25 and S100A9. Initially, these three genes have different functions and are involved in different pathways, showing no direct connection or overlapping pathway. However, all three proteins and their cellular functions can eventually be traced to calcium signaling and calcium levels. For example, WISP1 is part of the Wnt-pathway in which calcium has been shown to act as a second messenger [345], [346]. Additionally, RAB25 is part of the Raf/MEK/ERK pathway which is involved in cellular differentiation and this process is known to be induced by calcium stimulation [347]. And finally, S100A9 is a well-known protein directly binding calcium signaling pathway components [316].

Additionally, we performed a pathway analysis using The Database for Annotation, Visualization and Integrated Discovery (DAVID) [349], [350] on 887 differentially methylated genes identified and selected with MethylCap-Seq. This analysis revealed that the calcium signaling pathway is the only significantly enriched pathway amongst these genes. This confirms the calcium signaling pathway as the main contributor to the metastatic phenotype of pN+ OSCC.

In the introduction we discussed nine major pathways in which the biomarkers predictive for LN metastasis in OSCC can be classified: Cell cycle regulation, proliferation and apoptosis; Cell motility, cell adhesion and microenvironment; Transcription factors, immune system and angiogenesis [75]. Interestingly, all of these nine pathways have been shown to be related to the calcium signaling: cell cycle regulation [351], [352], cell proliferation [353], [354], apoptosis [351], [352], [354], cell motility [355], cell adhesion [352], [356], [357], microenvironment [354], [358] transcription factors [354], [359], the immune system [352], [360], and angiogenesis [352], [361], [362].

7

Summarizing, the calcium pathway appears to be the interconnecting factor between all three identified DNA methylation markers. Interestingly, the calcium pathway can be linked through several mechanisms to metastasis, specifically to mesenchymal-to-epithelial-transition as well as epithelial -to- mesenchymal -transition [436].
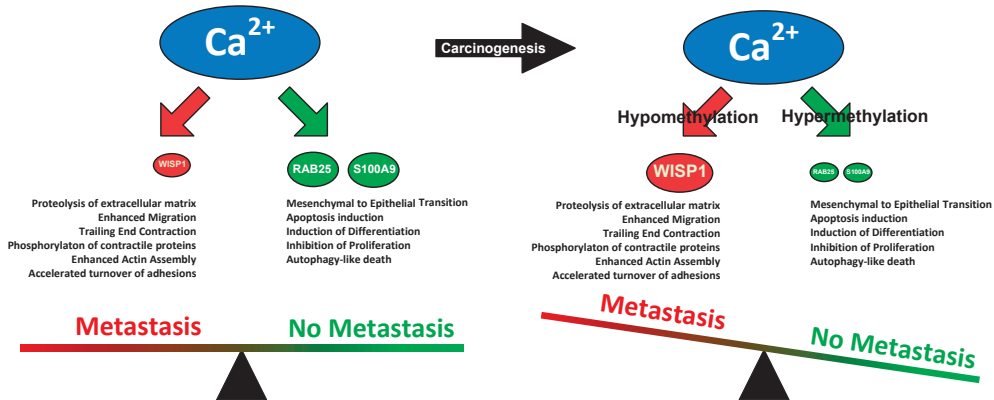
Figure 7.3. Proposed model for how DNA methylation can disrupt the balance between metastasis promoting and metastasis inhibiting pathways through calcium levels.

In addition, calcium promotes proteolysis of the extracellular matrix detaching cells from the matrix, enhanced migration, trailing end contraction of cells allowing cell motility, phosphorylation of contractile proteins activating them, enhancing actin assembly promoting a dynamic cytoskeleton and thus motility of cells, as well as enhancing turnover of adhesion molecules [437]. On the other side, calcium is also involved in other process associated with carcinogenesis such as mesenchymal-to-epithelial transition, apoptosis, inducing cell differentiation, inhibiting cell proliferation and inducing autophagy-like cell death[437].

Furthermore, changes in the DNA methylation patterns occur in close association with calcium signaling during carcinogenesis [438]. And finally, all three DNMTs have been found to be Ca2+ ion- and redox state-dependent, providing a direct link between the aberrant methylation in OSCC as well as the Calcium related genes found to be epigenetically deregulated in OSCC [439]. It must be noted though that very high Ca2+ concentrations are required before this effect is seen and therefore the clinical relevance of this finding might prove limited.

We hypothesize, that the changes of DNA methylation during carcinogenesis causes a shift in calcium-induced metastasis by inhibiting tumor suppressor genes through promoter hypermethylation of RAB25 and S100A9 as well as hypomethylation and upregulation of WISP1 (Figure 7.3) based upon the identification and characterization of these three new DNA methylation biomarkers and their role in the calcium pathway analysis as well as the pathway analysis of our shortlists of genes annotated by MethylCap-Seq analysis.

OSCC DNA methylation potential therapeutic targets

The three novel methylation markers might provide viable therapeutic targets for pN+ OSCC patients. Therapeutic upregulation of WISP1 might already be soon clinically possible due to the nature of the epigenetic changes that lead to overexpression of the WISP1 protein. Although no direct medication for WISP1 is available, drugs targeting the Wnt-pathway, of which WISP1 is a part, are being developed such as Compound LGK974 (WNT974) which specifically is being developed for treatment of several cancers including HNSCC [440], [441]. The epigenetic repression of RAB25 and S100A9 might be less easily directly targeted due to the lowered levels of these proteins as a result of hypermethylation. The epigenetic down-regulation of RAB25 and S100A9 could be treated with general demethylating medication. Two inhibitors of DNMT activity have already been approved by the FDA. Both Azacitidine and Decitabine are being used in the clinic to reduce overall DNA methylation in myelodysplastic syndromes [365], [366]. In 2014 a clinical trial (NCT02178072) started where HNSCC patients were treated with Azacitidine [367]. In fact, several preclinical studies have been performed where great promise of Azacitidine treatment of HNSCC was show through the reversal of chemoresistance and the induction of apoptosis [226]. However, genome-wide methylation might also induce harmful side effects such as the demethylation of epigenetically silenced oncogenes of metastasis promoting gene [174] such as we have shown is the case for WISP1 (chapter 3).

Additional treatment options for epigenetically silenced tumor suppressor genes such as S100A9 and RAB25 might be provided through advancements in the clinical application of CRISPR-Cas9. The Crispr-Cas9 complex is a RNA-guided DNA endonuclease that originated from bacterial immune systems (reviewed by [368]). The guide RNA is 20 nucleotides long and can be replaced with any other 20 nt long sequence to target any desired DNA sequence. This has led to a great interest for the CRISPR-Cas9 complex as a genomic editing tool. A recent study has shown that the CRISP-Cas9 inactive Cas9 endonuclease domain can be fused with a DNMT3A functional domain to specifically methylate a the guide RNA's target sequence[442]. Additionally, by fusing the RNA guided enzymatic system with the catalytic domain of the demethylation enzyme TET1, the CRISPR-Cas9 system has been used to unmethylated the targeted DNA [369]. These two targeted epigenetic editing tools might provide future therapeutic options to reverse pathogenic methylation levels of the hypomethylated WISP1 as well as the hypermethylated RAB25 and S100A9.

Finally, since all three biomarkers share a relation with calcium induced metastatic potential, the aberrant calcium-dependent pathways in tumor cells could also provide a therapeutic target for treatment of pN+ OSCC. For example, carboxyamido-triazole (CAI), a drug that alters Ca2+ concentration through inhibition of receptor-mediated Ca2+ influx and was shown to affect malignant proliferation and metastasis [362]. The relevance of these new treatment strategies however needs additional clinical validation as well as biological characterization.

7

Reflection on choosing MethylCap-Seq combined with Methylation Specific PCR

When constructing our experimental designs, we choose to use MethylCap-Seq to establish the OSCC-specific methylome. Our reasoning for choosing MethylCap-Seq was its preference of CG-rich hypermethylated sequences which would focus on promoter regions which would lend itself especially well for identifying epigenetically silenced genes. However, this traditional view on DNA methylation induced gene silencing has been contested after those initial experiments due to the identification of regions that are less CpG-rich than CpG islands but heavily impact gene expression such as CpG island shores and CpG shelves [193], [246]. In hind-sight, an algorithm that would better include these regulatory regions would give a more complete view of the OSCC methylome. Unfortunately, we introduced a bias towards CpG Islands by only selection MC associated with transcription start sites. Additionally, the bias of MBD2, the enrichment protein used in MethylCap-Seq, only binds fragments with a high amount of methylated CpG site. Possibly, the bias eliminated the less CpG-rich CpG shores and CpG shelves. However, these insights were not available at the time of the MethylCap-Seq selection. Moreover, MethylCap-Seq was the best choice at the time that was less constrained to certain regions such as the Illumina 27k platform that was available at the time. Other available options of profiling techniques were also not more suited at the time. While frozen sample DNA used for MethylCap-Seq, the source of the DNA of the same patients used for the validation originated from FFPE. DNA extracted from FFPE is associated with a relatively lower DNA quality, the required high amount and high-quality DNA that is required for e.g. Whole-Genome Shotgun Bisulfite Sequencing, WGSBS was not suited for our validation OSCC cohort.

The other DNA methylation discovery tools available at the time of study design like MeDIP and Reduced representation bisulfite sequencing were comparable to the MethylCap-Seq. The deciding factors for MethylCap-Seq included the bias of MethylCap-Seq for DNA fragments with a certain threshold of number of CpG sites and distance between CpG sites compared to MeDip and RRBS. Additionally, MethylCap-Seq was chosen for the compatibility of this data with the data available in the Map of the Human Methylome. This, in hindsight, was probably one of the essential pros of MethylCap-Seq for the outcome of this thesis. Inclusion of validation of biomarker selection with the data of the Map of the Human Methylome phase proven to be essential of both the LN metastasis study as well as the saliva study as described in Chapter 6.

Future prospective

The extensive use of available public databases of genome-wide OSCC methylation data provided in both the TCGA and GEO databases to validate and further refine our long list of potential DNA methylation biomarkers predictive for the presence of LN metastasis in OSCC, was an essential part to select a restricted number of candidate methylation markers. Validation of most of the very first selected methylation markers from our own MethylCap-Seq data with other methylation detection techniques failed on our larger OSCC validation cohort. The sample size of six pN0 and six pN+ OSCC used for our discovery experiment proved too small for error-free biomarker identification often leading to false positives and false negatives. By using publicly available 450k data of pN0 and pN+ OSCC provided by the TCGA studies as a first validation set, we could save a lot of time and experimental testing by

removing any underperforming methylation markers from our prior selections. Additionally, the TCGA 450k data originates from an array of different centers and patient populations, assisting in decreasing any bias introduced by using a single center patient cohort. Overall, the public access to such vast DNA methylation data greatly assisted us in identifying more promising OSCC N-status methylation markers. Not only did the TCGA portal provide methylation data but also mRNA and mutation data to further elucidate the biology behind metastatic OSCC. We also found that methylation is not the only mechanism underlying the deregulation of the gene of interest. Therefore, endeavors like the TCGA and The International Cancer Genome Consortium (ICGC) are the way forward in biomarker studies. The TCGA and ICGC projects give smaller labs access to greater patient populations and datasets then they could muster by themselves. Additionally, the amount of data and research questions that could be answered with these data is to vast to be done by a single institute. For example, the different labs working on the first TCGA papers concerning OSCC did not investigate a potential pN predictive marker at all while all the data for identifying these pN status predictive marker were present [443].

The GEO datasets provided invaluable updated annotation of the 450k probe locations. Not only do these open platforms provide lots of data for other labs to use, it also supports open and verifiable research by making all data publicly available as well as providing a platform for comparable analyses. However, there is still some progress to be made in the application and accessibility of these platforms. Reanalysis of these data still requires above-novice bioinformatic skills to which most labs do not have access. GEO does a great job in providing some R analysis on their GEO cloud with GEO2R but this only allows analysis of a single GEO dataset and not on multiple similar GEO datasets (https://www.ncbi.nlm.nih.gov/geo/geo2r/). Comparably, TCGA does provide some simple analysis tools on their cBioportal website but for larger more complex datasets such as the 450k data is not possible with the provided tools. Essential steps such as data normalization, especially when dealing with multiple 450k experiments from different centers, are not provided. Also, while a great amount of clinical data is available for the TCGA patients some clinical parameters such as patient survival is incomplete or inconsistently reported between different participating centers and treatment-response data are generally lacking or incomplete. For GEO datasets the extensive individual clinical data is generally lacking where patient group designation is the only parameter provided.

Nevertheless, the possibilities with these open datasets are endless. I might even foresee some labs completely dedicating themselves to answering research questions using only publicly available data. Some major insights have already been achieved with this approach. For example, the before-mentioned CpG Island canyons have been identified by analyzing publicly available 450k datasets. I also believe that extensive mining of these datasets will provide many more insights for which these datasets have not been used yet. The list of all TCGA project publications can be found online (https://tcga-data. nci.nih.gov/docs/publications/) and only a single paper is present for the characterizations of HNSCC but that paper does not compare pN0 and pN+ OSCC but compares all HNSCC samples together as on group [443]. Additional examples of reusage of the same TCGA data, in our lab we have performed a pilot study for the detection of local-recurrences by analyzing DNA methylation patterns in saliva of patient undergoing OSCC follow-up (Chapter 6). For the identification of OSCC recurrence markers

7

we have used the same MethylCap-Seq and 450k but instead of looking for differentially methylated regions between the pN0 and pN+ we have compared all OSCC with normal oral epithelial cells to identify universal hypermethylation markers in OSCC. Using the vast majority of the same datasets as used for the methylation N-status markers, we have solved another research question.

This thesis has created blueprints for DNA methylation biomarker identification experimental designs. We have laid the foundations for other studies to use when trying to identify biologically relevant biomarkers. The steps taken and described in this thesis have set an example for other groups how to incorporate the data available in both the Gene Expression Omnibus and The Cancer Genome Atlas. We have shown how to use these databases to both validate the predictive value of the selected DNA methylation markers and the biological validity of these markers. These steps can now be applied to different cancers and research questions. These steps will both reduce the costs for biomarker discovery studies as well as increase the odds of finding a biomarker with a high predictive value.

In our lab we have applied this same approach to identifying DNA methylation biomarkers to predict radiotherapy response in Oropharyngeal and laryngeal tumors (Clausen et al. in preparation). For this study another MethylCap-Seq was performed on three radioresistant and three radiosensitive HNSCC cell lines to identify DNA methylation markers that predict the tumors sensitivity to radiotherapy. It must be noted that employing TCGA data to this study is more difficult because the radiotherapy data in the TCGA database is less well defined and doesn't allow for precise determination of the tumors radiotherapy sensitivity.

As we have shown in this thesis, no expected universal predictive marker is available for predicting the nodal status of OSCC. It is therefore expected that a gene panel consisting of various DNA methylation biomarkers as well as other biomarkers, is a more feasible approach for clinical application. The size of this panel is yet unclear. In the literature a well-defined mRNA expression profile is described and this study consists of 696 genes to achieve a negative predictive value of 89% for determining nodal spread. The size of the panel greatly determines the DNA methylation detection technique that can be used for diagnostics. When assuming a gene panel of 228 genes a multiplexing on for example on the Ion Torrent S5 could be employed or upcoming techniques such as digitalMLPA with allows multiplexing of up 1500 regions of interest would be greatly suited for translating a gene panel to the clinic [444]. Using either of these NGS based methods would be much more expensive than for example MSP or Q-MSP, because MSP is most likely not quantitative enough to be applicable to the OSCC methylome that seems highly variable and requires strict cut-offs. This might make such a large desirable gene panel unsuited for large spread clinical application.

By developing a DNA methylation marker panel that for example focusses on improving on the major downsides of SLNB by providing a non-invasive and high positive predictive panel might greatly contribute to a clinical modality consisting of both SLNB and a methylation biomarker. Potentially, in the future the TCGA 450k data could be used to construct a DNA methylation gene signature predictive for LN-metastasis.

7

REFERENCES

[1] B. A. Van Dijk et al., "Rare cancers of the head and neck area in Europe," Eur. J. Cancer, vol. 48, no. 6, pp. 783–796, Apr. 2012.

[2] W. Garavello, E. Lucenteforte, C. Bosetti, and C. La Vecchia, "The role of foods and nutrients on oral and pharyngeal cancer risk," Minerva Stomatol, vol. 58, no. 1–2, pp. 25–34, 2009.

[3] B. J. M. Braakhuis et al., "A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications," Cancer Res., vol. 63, no. 8, pp. 1727–1730, Apr. 2003.

[4] S. Marur, G. D'Souza, W. H. Westra, and A. A. Forastiere, "HPV-associated head and neck cancer: a virus-related cancer epidemic," Lancet Oncol., vol. 11, no. 8, pp. 781–789, Aug. 2010.

[5] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," CA. Cancer J. Clin., vol. 65, no. 2, pp. 87–108, Mar. 2015.

[6] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," CA. Cancer J. Clin., vol. 66, no. 1, pp. 7–30, Jan. 2016.

[7] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2012," CA. Cancer J. Clin., vol. 62, no. 1, pp. 10–29, Jan. 2012.

[8] T. Hasegawa et al., "Risk Factors Associated with Distant Metastasis in Patients with Oral Squamous Cell Carcinoma," Otolaryngol. Neck Surg., vol. 152, no. 6, pp. 1053–1060, Jun. 2015.

[9] S. Irani, "Distant metastasis from oral cancer: A review and molecular biologic aspects.," J. Int. Soc. Prev. Community Dent., vol. 6, no. 4, pp. 265–71, 2016.

[10] C. R. Leemans, R. Tiwari, J. P. J. Nauta, I. Van der Waal, and G. B. Snow, "Regional lymph node involvement and its significance in the development of distant metastases in head and neck carcinoma," Cancer, vol. 71, no. 2, pp. 452–456, Jan. 1993.

[11] A. Yasui, Y. Okada, I. Mataga, and M. Katagiri, "An Analysis of Distant Metastases in Oral Squamous Cell Carcinoma," 2010.

[12] R. Murakami et al., "Prognostic impact of the level of nodal involvement: retrospective analysis of patients with advanced oral squamous cell carcinoma," Br. J. Oral Maxillofac. Surg., vol. 55, no. 1, pp. 50–55, Jan. 2017.

[13] S. R. Larsen, J. Johansen, J. A. Sørensen, and A. Krogdahl, "The prognostic significance of histological features in oral squamous cell carcinoma.," J. Oral Pathol. Med., vol. 38, no. 8, pp. 657–62, Sep. 2009.

[14] E. W. Odell et al., "The prognostic value of individual histologic grading parameters in small lingual squamous cell carcinoma. The importance of the pattern of invasion.," Cancer, vol. 74, no. 3, pp. 789–94, Aug. 1994.

[15] M. Pentenero, S. Gandolfo, and M. Carrozzo, "Importance of tumor thickness and depth of invasion in nodal involvement and prognosis of oral squamous cell carcinoma: a review of the literature.," Head Neck, vol. 27, no. 12, pp. 1080–91, Dec. 2005.

[16] M. Suzuki et al., "Clinicopathological factors related to cervical lymph node metastasis in a patient with carcinoma of the oral floor," Acta Otolaryngol., vol. 127, no. sup559, pp. 129–135, Jan. 2007.

[17] J. A. Woolgar, S. Rogers, C. R. West, R. D. Errington, J. S. Brown, and E. D. Vaughan, "Survival and patterns of recurrence in 200 oral cancer patients treated by radical surgery and neck dissection.," Oral Oncol., vol. 35, no. 3, pp. 257–65, May 1999.

[18] L. J. Melchers et al., "Tumour infiltration depth ≥4 mm is an indication for an elective neck dissection in pT1cN0 oral squamous cell carcinoma.," Oral Oncol., vol. 48, no. 4, pp. 337–42, Apr. 2012.

[19] W. M. Lydiatt et al., "Head and neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual," CA. Cancer J. Clin., vol. 67, no. 2, pp. 122–137, Mar. 2017.

[20] S. B. Edge and American Joint Committee on Cancer., AJCC cancer staging manual. Springer, 2010.

[21] K. Omura, "Current status of oral cancer treatment strategies: surgical treatments for oral squamous cell carcinoma," Int. J. Clin. Oncol., vol. 19, no. 3, pp. 423–430, Jun. 2014.

[22] B. J. M. Braakhuis, C. R. Leemans, and O. Visser, "Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011.," Oral Oncol., vol. 50, no. 7, pp. 670–5, Jul. 2014.

[23] R. de Bree et al., "Advances in diagnostic modalities to detect occult lymph node metastases in head and neck squamous cell carcinoma.," Head Neck, vol. 37, no. 12, pp. 1829–39, Dec. 2015.

[24] T. M. Govers et al., "An international comparison of the management of the neck in early oral squamous cell carcinoma in the Netherlands, UK, and USA.," J. Craniomaxillofac. Surg., vol. 44, no. 1, pp. 62–9, Jan. 2016.

[25] K. Boeve et al., "High sensitivity and negative predictive value of sentinel lymph node biopsy in a retrospective early stage oral cavity cancer cohort in the Northern Netherlands," Clin. Otolaryngol., Mar. 2018.

[26] C. F. Thompson, M. A. St. John, G. Lawson, T. Grogan, D. Elashoff, and A. H. Mendelsohn, "Diagnostic value of sentinel lymph node biopsy in head and neck cancer: a meta-analysis," Eur. Arch. Oto-Rhino-Laryngology, vol. 270, no. 7, pp. 2115–2122, Jul. 2013.

[27] S. J. Stoeckli, T. Huebner, G. F. Huber, and M. A. Broglie, "Technique for reliable sentinel node biopsy in squamous cell carcinomas of the floor of mouth," Head Neck, vol. 38, no. 9, pp. 1367–1372, Sep. 2016.

[28] C. Bluemel, D. Rubello, P. M. Colletti, R. de Bree, and K. Herrmann, "Sentinel lymph node biopsy in oral and oropharyngeal squamous cell carcinoma: current status and unresolved challenges," Eur. J. Nucl. Med. Mol. Imaging, vol. 42, no. 9, pp. 1469–1480, Aug. 2015.

[29] C. Schilling et al., "Sentinel European Node Trial (SENT): 3-year results of sentinel node biopsy in oral cancer," Eur. J. Cancer, vol. 51, no. 18, pp. 2777–2784, Dec. 2015.

[30] C. Schilling et al., "Sentinel lymph node biopsy for oral squamous cell carcinoma. Where are we now?," Br. J. Oral Maxillofac. Surg., vol. 55, no. 8, pp. 757–762, Oct. 2017.

[31] K. Boeve et al., "Lymphatic drainage patterns of oral maxillary tumors: Approachable locations of sentinel lymph nodes mainly at the cervical neck level," Head Neck, vol. 39, no. 3, pp. 486–491, Mar. 2017.

[32] F. A. Sawair, C. R. Irwin, D. J. Gordon, A. G. Leonard, M. Stephenson, and S. S. Napier, "Invasive front grading: reliability and usefulness in the management of oral squamous cell carcinoma.," J. Oral Pathol. Med., vol. 32, no. 1, pp. 1–9, Jan. 2003.

[33] M. Brandwein-Gensler et al., "Validation of the histologic risk model in a new cohort of patients with head and neck squamous cell carcinoma.," Am. J. Surg. Pathol., vol. 34, no. 5, pp. 676–88, May 2010.

[34] Y.C. Chang, S. Nieh, S.-F. Chen, S.-W. Jao, Y.-L. Lin, and E. Fu, "Invasive pattern grading score designed as an independent prognostic indicator in oral squamous cell carcinoma," Histopathology, vol. 57, no. 2, pp. 295–303, Aug. 2010.

[35] M. G. J. Heerema, L. J. Melchers, J. L. N. Roodenburg, E. Schuuring, G. H. de Bock, and B. van der Vegt, "Reproducibility and prognostic value of pattern of invasion scoring in low-stage oral squamous cell carcinoma," Histopathology, vol. 68, no. 3, pp. 388–397, Feb. 2016.

[36] R. L. Foote et al., "NCCN Clinical Practice Guidelines in Oncology for HNSCC," 2019.

[37] T. Hamada et al., "DF3 epitope expression on MUC1 mucin is associated with tumor aggressiveness, subsequent lymph node metastasis, and poor prognosis in patients with oral squamous cell carcinoma," Cancer, vol. 118, no. 21, pp. 5251–5264, Nov. 2012.

[38] Y. Michifuri et al., "High expression of ALDH1 and SOX2 diffuse staining pattern of oral squamous cell carcinomas correlates to lymph node metastasis," Pathol. Int., vol. 62, no. 10, pp. 684–689, Oct. 2012.

[39] A. Rahmani, M. Alzohairy, A. Y. Babiker, M. A. Rizvi, and H. G. Elkarimahmad, "Clinicopathological significance of PTEN and bcl2 expressions in oral squamous cell carcinoma.," Int. J. Clin. Exp. Pathol., vol. 5, no. 9, pp. 965–71, 2012.

[40] H. Xie, S. Huang, W. Li, H. Zhao, T. Zhang, and D. Zhang, "Upregulation of Src homology phosphotyrosyl phosphatase 2 (Shp2) expression in oral cancer and knockdown of Shp2 expression inhibit tumor cell viability and invasion in vitro," Oral Surgery, Oral Med. Oral Pathol. Oral Radiol., vol. 117, pp. 234–242, 2014.

[41] M. T. Brazão-Silva et al., "Metallothionein gene expression is altered in oral cancer and may predict metastasis and patient outcomes," Histopathology, vol. 67, no. 3, pp. 358–367, Sep. 2015.

[42] J. P. Oliveira-Costa et al., "Gene expression patterns through oral squamous cell carcinoma development: PD-L1 expression in primary tumor and circulating tumor cells," Oncotarget, vol. 6, no. 25, pp. 20902–20, Aug. 2015.

[43] J. Y. Tang et al., "Overexpression of Autophagy-Related 16-Like 1 in Patients with Oral Squamous Cell Carcinoma," Pathol. Oncol. Res., vol. 21, no. 2, pp. 301–305, Apr. 2015.

[44] M. Grimm et al., "ABCB5 expression and cancer stem cell hypothesis in oral squamous cell carcinoma.," Eur. J. Cancer, vol. 48, no. 17, pp. 3186–97, Nov. 2012.

[45] A. Wushou et al., "Correlation of increased twist with lymph node metastasis in patients with oral squamous cell

carcinoma.," J. Oral Maxillofac. Surg., vol. 70, no. 6, pp. 1473–9, Jun. 2012.

[46] M. P. Foschini et al., "Podoplanin and E-cadherin Expression in Preoperative Incisional Biopsies of Oral Squamous Cell Carcinoma Is Related to Lymph Node Metastases," Int. J. Surg. Pathol., vol. 21, no. 2, pp. 133–141, Apr. 2013.

[47] M. Kono, M. Watanabe, H. Abukawa, O. Hasegawa, T. Satomi, and D. Chikazu, "Cyclo-Oxygenase−2 Expression Is Associated With Vascular Endothelial Growth Factor C Expression and Lymph Node Metastasis in Oral Squamous Cell Carcinoma," J. Oral Maxillofac. Surg., vol. 71, no. 10, pp. 1694–1702, Oct. 2013.

[48] M. Nagata et al., "ITGA3 and ITGB4 expression biomarkers estimate the risks of locoregional and hematogenous dissemination of oral squamous cell carcinoma.," BMC Cancer, vol. 13, no. 1, p. 410, Sep. 2013.

[49] L. J. Melchers et al., "Lack of claudin-7 is a strong predictor of regional recurrence in oral and oropharyngeal squamous cell carcinoma," Oral Oncol., vol. 49, no. 10, pp. 998–1005, Oct. 2013.

[50] Y. Goto et al., "Possible involvement of ΔNp63 downregulation in the invasion and metastasis of oral squamous cell carcinoma via induction of a mesenchymal phenotype.," Clin. Exp. Metastasis, vol. 31, no. 3, pp. 293–306, Mar. 2014.

[51] C. H. Hsin et al., "High Level of Plasma Matrix Metalloproteinase-11 Is Associated with Clinicopathological Characteristics in Patients with Oral Squamous Cell Carcinoma," PLoS One, vol. 9, no. 11, p. e113129, Nov. 2014.

[52] Y. Li, J. Zhang, and S. Hong, "ANO1 as a marker of oral squamous cell carcinoma and silencing ANO1 suppresses migration of human SCC-25 cells.," Med. Oral Patol. Oral Cir. Bucal, vol. 19, no. 4, pp. e313-9, Jul. 2014.

[53] S. Magnussen, O. G. Rikardsen, E. Hadler-Olsen, L. Uhlin-Hansen, S. E. Steigen, and G. Svineng, "Urokinase plasminogen activator receptor (uPAR) and plasminogen activator inhibitor-1 (PAI-1) are potential predictive biomarkers in early stage oral squamous cell carcinomas (OSCC).," PLoS One, vol. 9, no. 7, p. e101895, Jul. 2014.

[54] J. Natarajan, K. Hunter, V. S. Mutalik, and R. Radhakrishnan, "Overexpression of S100A4 as a biomarker of metastasis and recurrence in oral squamous cell carcinoma.," J. Appl. Oral Sci., vol. 22, no. 5, pp. 426–33.

[55] A. A. Byatnal, A. A. Byatnal, S. Sen, V. Guddattu, and M. C. Solomon, "Cyclooxygenase-2 − An Imperative Prognostic Biomarker in Oral Squamous Cell Carcinoma- An Immunohistochemical Study," Pathol. Oncol. Res., vol. 21, no. 4, pp. 1123–1131, Sep. 2015.

[56] Y. P. Hsu et al., "Serum markers of CYFRA 21-1 and C-reactive proteins in oral squamous cell carcinoma.," World J. Surg. Oncol., vol. 13, no. 1, p. 253, Aug. 2015.

[57] Y. H. Ni, L. Ding, X. F. Huang, Y. Dong, Q. G. Hu, and Y. Y. Hou, "Microlocalization of CD68+ tumor-associated macrophages in tumor stroma correlated with poor clinical outcomes in oral squamous cell carcinoma patients.," Tumour Biol., vol. 36, no. 7, pp. 5291–8, Jul. 2015.

[58] K. Sappayatosok and E. Phattarataratip, "Overexpression of Claudin-1 is Associated with Advanced Clinical Stage and Invasive Pathologic Characteristics of Oral Squamous Cell

Carcinoma.," Head Neck Pathol., vol. 9, no. 2, pp. 173–80, Jun. 2015.

[59] H. Zhang et al., "CMTM3 inhibits cell growth and migration and predicts favorable survival in oral squamous cell carcinoma.," Tumour Biol., vol. 36, no. 10, pp. 7849–58, Sep. 2015.

[60] J. Zhou, D. Tao, Q. Xu, Z. Gao, and D. Tang, "Expression of E-cadherin and vimentin in oral squamous cell carcinoma." Int. J. Clin. Exp. Pathol., vol. 8, no. 3, pp. 3150–4, 2015.

[61] C. W. Lin et al., "Role of lipocalin 2 and its complex with matrix metalloproteinase-9 in oral cancer.," Oral Dis., vol. 18, no. 8, pp. 734–40, Nov. 2012.

[62] K. Miyaguchi et al., "Loss of NKX3-1 as a potential marker for an increased risk of occult lymph node metastasis and poor prognosis in oral squamous cell carcinoma.," Int. J. Oncol., vol. 40, no. 6, pp. 1907–14, Jun. 2012.

[63] D. Sartini et al., "Analysis of tissue and salivary nicotinamide N-methyltransferase in oral squamous cell carcinoma: basis for the development of a noninvasive diagnostic test for early-stage disease.," Biol. Chem., vol. 393, no. 6, pp. 505–11, May 2012.

[64] H. M. Wu, W. Cao, D. Ye, G. X. Ren, Y. N. Wu, and W. Guo, "Contactin 1 (CNTN1) expression associates with regional lymph node metastasis and is a novel predictor of prognosis in patients with oral squamous cell carcinoma.," Mol. Med. Rep., vol. 6, no. 2, pp. 265–70, Aug. 2012.

[65] S. Ishige et al., "Decreased expression of kallikrein-related peptidase 13: possible contribution to metastasis of human oral cancer.," Mol. Carcinog., vol. 53, no. 7, pp. 557–65, Jul. 2014.

[66] T. Sasahira et al., "Transport and Golgi organisation protein 1 is a novel tumour progressive factor in oral squamous cell carcinoma.," Eur. J. Cancer, vol. 50, no. 12, pp. 2142–51, Aug. 2014.

[67] M. Weber et al., "Small oral squamous cell carcinomas with nodal lymphogenic metastasis show increased infiltration of M2 polarized macrophages--an immunohistochemical analysis.," J. Craniomaxillofac. Surg., vol. 42, no. 7, pp. 1087–94, Oct. 2014.

[68] F. Xia et al., "Glucose-regulated protein 78 and heparanase expression in oral squamous cell carcinoma: correlations and prognostic significance.," World J. Surg. Oncol., vol. 12, no. 1, p. 121, Apr. 2014.

[69] W. T. Lin, T. M. Shieh, L. C. Yang, T. Y. Wang, M. Y. Chou, and C. C. Yu, "Elevated Lin28B expression is correlated with lymph node metastasis in oral squamous cell carcinomas.," J. Oral Pathol. Med., vol. 44, no. 10, pp. 823–30, Nov. 2015.

[70] J. S. Yang, C. W. Lin, C. Y. Chuang, S. C. Su, S. H. Lin, and S. F. Yang, "Carbonic anhydrase IX overexpression regulates the migration and progression in oral squamous cell carcinoma.," Tumour Biol., vol. 36, no. 12, pp. 9517–24, Dec. 2015.

[71] J. W. Kim et al., "Prognostic value of glucosylceramide synthase and P-glycoprotein expression in oral cavity cancer.," Int. J. Clin. Oncol., vol. 21, no. 5, pp. 883–889, Oct. 2016.

[72] L. Lin et al., "Interleukin-37 expression and its potential role in oral leukoplakia and oral squamous cell carcinoma.," Sci. Rep., vol. 6, no. 1, p. 26757, Jul. 2016.

[73] W. J. Shin et al., "KiSS-1 expression in oral squamous cell carcinoma and its prognostic significance.," APMIS, vol. 124, no. 4, pp. 291–8, Apr. 2016.

[74] H. Xu et al., "Serum miR-483-5p: a novel diagnostic and prognostic biomarker for patients with oral squamous cell carcinoma.," Tumour Biol., vol. 37, no. 1, pp. 447–53, Jan. 2016.

[75] S. Blatt et al., "Biomarkers in diagnosis and therapy of oral squamous cell carcinoma: A review of the literature," J. Cranio-Maxillofacial Surg., vol. 45, no. 5, pp. 722–730, May 2017.

[76] D. G. Altman, L. M. McShane, W. Sauerbrei, and S. E. Taube, "Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration," BMC Med., vol. 10, no. 1, p. 51, Dec. 2012.

[77] M. Giefing et al., "Moving towards personalised therapy in head and neck squamous cell carcinoma through analysis of next generation sequencing data.," Eur. J. Cancer, vol. 55, pp. 147–57, Mar. 2016.

[78] S. R. van Hooff et al., "Validation of a gene expression signature for assessment of lymph node metastasis in oral squamous cell carcinoma," J. Clin. Oncol., vol. 30, no. 33, pp. 4104–4110, Nov. 2012.

[79] R. Noorlag et al., "Amplification and protein overexpression of cyclin D1: Predictor of occult nodal metastasis in early oral cancer," Head Neck, vol. 39, no. 2, pp. 326–333, Feb. 2017.

[80] A. C. Chiang and J. Massagué, "Molecular Basis of Metastasis," N. Engl. J. Med., vol. 359, no. 26, pp. 2814–2823, Dec. 2008.

[81] S. Valastyan and R. A. Weinberg, "Tumor metastasis: molecular insights and evolving paradigms.," Cell, vol. 147, no. 2, pp. 275–92, Oct. 2011.

[82] L. Tonella, M. Giannoccaro, S. Alfieri, S. Canevari, and L. De Cecco, "Gene Expression Signatures for Head and Neck Cancer Patient Stratification: Are Results Ready for Clinical Application?," Curr. Treat. Options Oncol., vol. 18, no. 5, p. 32, May 2017.

[83] P. Roepman et al., "An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas," Nat. Genet., vol. 37, no. 2, pp. 182–186, Feb. 2005.

[84] C. C. Yu et al., "Bmi-1 Regulates Snail Expression and Promotes Metastasis Ability in Head and Neck Squamous Cancer-Derived ALDH1 Positive Cells.," J. Oncol., vol. 2011, Sep. 2011.

[85] S. Kimura et al., "Expression of microRNAs in squamous cell carcinoma of human head and neck and the esophagus: miR-205 and miR-21 are specific markers for HNSCC and ESCC.," Oncol. Rep., vol. 23, no. 6, pp. 1625–33, Jun. 2010.

[86] S. K. Zaidi et al., "Architectural Epigenetics: Mitotic Retention of Mammalian Transcriptional Regulatory Information," Mol. Cell. Biol., vol. 30, no. 20, pp. 4758–4766, Oct. 2010.

[87] C. H. Waddington, "The epigenotype," Endeavour, vol. 1, pp. 18–20, 1942.

[88] J. D. Watson and F. H. Crick, "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.," Nature, vol. 171, no. 4356, pp. 737–8, Apr. 1953.

[89] A. Razin, A. Bird, and P. Pitha, "CpG methylation, chromatin structure and gene silencing-a three-way connection.," EMBO J., vol. 17, no. 17, pp. 4905–8, Sep. 1998.

[90] A. Shilatifard, "Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression," Annu. Rev. Biochem., vol. 75, no. 1, pp. 243–269, Jun. 2006.

[91] K. C. Wang et al., "A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression.," Nature, vol. 472, no. 7341, pp. 120–4, Apr. 2011.

[92] A. M. Khalil et al., "Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression.," Proc. Natl. Acad. Sci. U. S. A., vol. 106, no. 28, pp. 11667–72, Jul. 2009.

[93] S. A. Weiner and A. L. Toth, "Epigenetics in social insects: a new direction for understanding the evolution of castes.," Genet. Res. Int., vol. 2012, p. 609810, 2012.

[94] S. B. Baylin and P. A. Jones, "A decade of exploring the cancer epigenome - biological and translational implications," Nat. Rev., vol. 11, no. 10, pp. 726–734, Sep. 2011.

[95] W. M. Rideout Iii, G. A. Coetzee, A. F. Olumi, and P. A. Jones, "5-Methylcytosine as an Endogenous Mutagen in the Human LDL Receptor and p53 Genes," Source Sci. New Ser., vol. 249, no. 4974, pp. 1288–1290, 1990.

[96] G. P. Pfeifer, "Environmental exposures and mutational patterns of cancer genomes.," Genome Med., vol. 2, no. 8, p. 54, Aug. 2010.

[97] M. Esteller, "Epigenetics in cancer," N. Engl. J. Med., vol. 358, no. 11, pp. 1148–1159, Mar. 2008.

[98] R. Singal and G. D. Ginder, "DNA methylation.," Blood, vol. 93, no. 12, pp. 4059–70, Jun. 1999.

[99] J. Boyes and A. Bird, "DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein.," Cell, vol. 64, no. 6, pp. 1123–34, Mar. 1991.

[100] M. Jackson, L. Marks, G. H. W. May, and J. B. Wilson, "The genetic basis of disease.," Essays Biochem., vol. 62, no. 5, pp. 643–723, Dec. 2018.

[101] P. A. Jones, J.-P. J. Issa, and S. Baylin, "Targeting the cancer epigenome for therapy," Nat. Rev. Genet., vol. 17, no. 10, pp. 630–641, Oct. 2016.

[102] Z. Siegfried, S. Eden, M. Mendelsohn, X. Feng, B.-Z. Tsuberi, and H. Cedar, "DNA methylation represses transcription in vivo," Nat. Genet., vol. 22, no. 2, pp. 203–206, Jun. 1999.

[103] M. Ehrlich, "DNA hypomethylation in cancer cells," Epigenomics, vol. 1, no. 2, pp. 239–259, Dec. 2009.

[104] B. Aiissani and G. Bernardi, "CpG islands: features and distribution in the genomes of vertebrates," Gene, vol. 106, pp. 173–183, 1991.

[105] M. Ehrlich et al., "Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells," Nucleic Acids Res., vol. 10, no. 8, pp. 2709–2721, Apr. 1982.

[106] A. P. Bird, "CpG-rich islands and the function of DNA methylation.," Nature, vol. 321, no. 6067, pp. 209–13, 1986.

[107] A. Bird, "DNA methylation patterns and epigenetic memory," Genes Dev., vol. 16, no. 1, pp. 6–21, Jan. 2002.

[108] T. H. Bestor, "Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain.," EMBO J., vol. 11, no. 7, pp. 2611–7, Jul. 1992.

[109] S. Pradhan, A. Bacolla, R. D. Wells, and R. J. Roberts, "Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation.," J. Biol. Chem., vol. 274, no. 46, pp. 33002–10, Nov. 1999.

[110] W. Reik and J. Walter, "Genomic imprinting: parental influence on the genome," Nat. Rev. Genet., vol. 2, no. 1, pp. 21–32, Jan. 2001.

[111] M. F. Fraga et al., "A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors.," Cancer Res., vol. 64, no. 16, pp. 5527–34, Aug. 2004.

[112] M. Esteller et al., "Promoter Hypermethylation and BRCA1 Inactivation in Sporadic Breast and Ovarian Tumors," J. Natl. Cancer Inst., vol. 92, no. 7, pp. 564–569, Apr. 2000.

[113] L. Van Neste, J. G. Herman, G. Otto, J. W. Bigley, J. I. Epstein, and W. Van Criekinge, "The Epigenetic promise for prostate cancer diagnosis," Prostate, vol. 72, no. 11, pp. 1248–1261, Aug. 2012.

[114] N. Thon, S. Kreth, and F.-W. Kreth, "Personalized treatment strategies in glioblastoma: MGMT promoter methylation status.," Onco. Targets. Ther., vol. 6, pp. 1363–72, Sep. 2013.

[115] M. E. Hegi et al., "MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma," N. Engl. J. Med., vol. 352, no. 10, pp. 997–1003, Mar. 2005.

[116] W. Wick et al., "NOA-04 Randomized Phase III Trial of Sequential Radiochemotherapy of Anaplastic Glioma With Procarbazine, Lomustine, and Vincristine or Temozolomide," J. Clin. Oncol., vol. 27, no. 35, pp. 5874–5880, Dec. 2009.

[117] S. Lalezari et al., "Combined analysis of O6-methylguanine-DNA methyltransferase protein expression and promoter methylation provides optimized prognostication of glioblastoma outcome.," Neuro. Oncol., vol. 15, no. 3, pp. 370–81, Mar. 2013.

[118] J. R. Graff, E. Gabrielson, H. Fujii, S. B. Baylin, and J. G. Herman, "Methylation patterns of the E-cadherin 5′ CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression.," J. Biol. Chem., vol. 275, no. 4, pp. 2727–32, Jan. 2000.

[119] F. J. Carmona et al., "Epigenetic disruption of cadherin-11 in human cancer metastasis.," J. Pathol., vol. 228, no. 2, pp. 230–40, Oct. 2012.

[120] A. Lujambio et al., "A microRNA DNA methylation signature for human cancer metastasis.," Proc. Natl. Acad. Sci. U. S. A., vol. 105, no. 36, pp. 13556–61, 2008.

[121] X. Hao et al., "DNA methylation markers for diagnosis and prognosis of common cancers.," Proc. Natl. Acad. Sci. U. S. A., vol. 114, no. 28, pp. 7414–7419, Jul. 2017.

[122] P. V Jithesh et al., "The epigenetic landscape of oral squamous cell carcinoma," Br. J. Cancer, vol. 108, no. 2, pp. 370–379, Feb. 2013.

[123] J. A. Gasche and A. Goel, "Epigenetic mechanisms in oral carcinogenesis," Futur. Oncol., vol. 8, no. 11, pp. 1407–1425, Nov. 2012.

[124] I. González-Ramírez, C. García-Cuellar, Y. Sánchez-Pérez, and M. Granados-García, "DNA methylation in oral squamous cell carcinoma: molecular mechanisms and clinical implications," Oral Dis., vol. 17, no. 8, pp. 771–778, Nov. 2011.

[125] M. Mascolo et al., "Epigenetic Disregulation in Oral Cancer," Int. J. Mol. Sci., vol. 13, no. 2, pp. 2331–2353, Feb. 2012.

[126] Y. Li et al., "Investigation of tumor suppressing function of CACNA2D3 in esophageal squamous cell carcinoma," PLoS One, vol. 8, no. 4, p. e60027, Apr. 2013.

[127] X. Xing et al., "The prognostic value of CDKN2A hypermethylation in colorectal cancer: a meta-analysis.," Br. J. Cancer, vol. 108, no. 12, pp. 2542–8, Jun. 2013.

[128] A. L. Carvalho et al., "Detection of promoter hypermethylation in salivary rinses as a biomarker for head and neck squamous cell carcinoma surveillance," Clin. Cancer Res., vol. 17, no. 14, pp. 4782–4789, Jul. 2011.

[129] E. Heitzer, I. S. Haque, C. E. S. Roberts, and M. R. Speicher, "Current and future perspectives of liquid biopsies in genomics-driven oncology," Nat. Rev. Genet., vol. 20, no. 2, pp. 71–88, Feb. 2019.

[130] W. A. Palmisano et al., "Predicting Lung Cancer by Detecting Aberrant Promoter Methylation in Sputum," CANCER Res., vol. 60, pp. 5954–5958, 2000.

[131] P. W. Laird, "Principles and challenges of genomewide DNA methylation analysis," Nat. Rev., vol. 11, no. 3, pp. 191–203, Mar. 2010.

[132] J. Kaput and T. W. Sneider, "Methylation of somatic vs germ cell DNAs analyzed by restriction endonuclease digestions.," Nucleic Acids Res., vol. 7, no. 8, pp. 2303–22, Dec. 1979.

[133] S. H. Cross, J. A. Charlton, X. Nan, and A. P. Bird, "Purification of CpG islands using a methylated DNA binding column," Nat. Genet., vol. 6, no. 3, pp. 236–244, Mar. 1994.

[134] M. Frommer et al., "A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.," Proc. Natl. Acad. Sci. U. S. A., vol. 89, no. 5, pp. 1827–31, Mar. 1992.

[135] E. Ballestar et al., "Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer," EMBO J., vol. 22, no. 23, pp. 6335–6345, Dec. 2003.

[136] K. Mensaert, S. Denil, G. Trooskens, W. Van Criekinge, O. Thas, and T. De Meyer, "Next-generation technologies and data analytical approaches for epigenomics," Environ. Mol. Mutagen., vol. 55, pp. 155–170, Dec. 2013.

[137] E. Olkhov-Mitsel and B. Bapat, "Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers," Cancer Med., vol. 1, no. 2, pp. 237–260, Oct. 2012.

[138] J. G. Herman, J. R. Graff, S. Myöhänen, B. D. Nelkin, and S. B. Baylin, "Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.," Proc. Natl. Acad. Sci. U. S. A., vol. 93, no. 18, pp. 9821–6, Sep. 1996.

[139] G. B. Wisman et al., "Assessment of gene promoter hypermethylation for detection of cervical neoplasia," Int. J. cancer.Journal Int. du cancer, vol. 119, no. 8, pp. 1908–1914, Oct. 2006.

[140] S. Colella, L. Shen, K. A. Baggerly, J. P. Issa, and R. Krahe, "Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites," Biotechniques, vol. 35, no. 1, pp. 146–150, Jul. 2003.

[141] Z. Xiong and P. W. Laird, "COBRA: a sensitive and quantitative DNA methylation assay," Nucleic Acids Res., vol. 25, no. 12, 1997.

[142] M. Ongenaert et al., "Discovery of DNA methylation markers in cervical cancer using relaxation ranking," BMC Med. Genomics, vol. 1, no. 1, p. 57, Dec. 2008.

[143] F. Gautier, H. Bünemann, and L. Grotjahn, "Analysis of calf-thymus satellite DNA: evidence for specific methylation of cytosine in C-G sequences.," Eur. J. Biochem., vol. 80, no. 1, pp. 175–83, Oct. 1977.

[144] C. Waalwijk and R. A. Flavell, "DNA methylation at a CCGG sequence in the large intron of the rabbit beta-globin gene: tissue-specific variations.," Nucleic Acids Res., vol. 5, no. 12, pp. 4631–4, Dec. 1978.

[145] E. M. Southern, "Detection of specific sequences among DNA fragments separated by gel electrophoresis," J. Mol. Biol., vol. 98, no. 3, pp. 503–517, Nov. 1975.

[146] A. Baumer, U. Wiedemann, M. Hergersberg, and A. Schinzel, "A novel MSP/DHPLC method for the investigation of the methylation status of imprinted genes enables the molecular detection of low cell mosaicisms.," Hum. Mutat., vol. 17, no. 5, pp. 423–30, May 2001.

[147] Simonetta Friso, Sang-Woon Choi, and Gregory G. Dolnikowski, and J. Selhub, "A Method to Assess Genomic DNA Methylation Using High-Performance Liquid Chromatography/Electrospray Ionization Mass Spectrometry," 2002.

[148] R. Lister et al., "Human DNA methylomes at base resolution show widespread epigenomic differences," Nature, vol. 462, no. 7271, pp. 315–322, Nov. 2009.

[149] A. Meissner et al., "Genome-scale DNA methylation maps of pluripotent and differentiated cells," Nature, vol. 454, no. 7205, pp. 766–770, Aug. 2008.

[150] T. Rauch and G. P. Pfeifer, "Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer," Lab. Invest., vol. 85, no. 9, pp. 1172–1180, Sep. 2005.

[151] T. De Meyer et al., "Quality evaluation of methyl binding domain based kits for enrichment DNA-methylation sequencing," PLoS One, vol. 8, no. 3, p. e59068, 2013.

[152] T. A. Down et al., "A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis.," Nat. Biotechnol., vol. 26, no. 7, pp. 779–85, Jul. 2008.

[153] F. V Jacinto, E. Ballestar, and M. Esteller, "Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome.," Biotechniques, vol. 44, no. 1, p. 35, 37, 39 passim, Jan. 2008.

[154] M. Bibikova et al., "High density DNA methylation array with single CpG site resolution," Genomics, vol. 98, no. 4, pp. 288–295, Oct. 2011.

[155] D. J. Weisenberger, D. Van, D. Berg, F. Pan, B. P. Berman, and P. W. Laird, "Comprehensive DNA Methylation Analysis on the Illumina ® Infinium ® Assay Platform."

[156] M. Weber et al., "Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells," Nat. Genet., vol. 37, no. 8, pp. 853–862, Aug. 2005.

[157] W.-S. Yong, F.-M. Hsu, and P.-Y. Chen, "Profiling genome-wide DNA methylation," Epigenetics Chromatin, vol. 9, no. 1, p. 26, Dec. 2016.

[158] S. G. Houghton and F. R. Cockerill, "Real-time PCR: Overview

and applications," Surgery, vol. 139, no. 1, pp. 1–5, Jan. 2006.

[159] S. Kurdyukov and M. Bullock, "DNA Methylation Analysis: Choosing the Right Method.," Biology (Basel)., vol. 5, no. 1, Jan. 2016.

[160] T. M. Govers et al., "Management of the N0neck in early stage oral squamous cell cancer: A modeling study of the cost-effectiveness," Oral Oncol., vol. 49, no. 8, pp. 771–777, Aug. 2013.

[161] J. U. Guo, Y. Su, C. Zhong, G. Ming, and H. Song, "Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain," Cell, vol. 145, no. 3, pp. 423–434, Apr. 2011.

[162] K. Williams et al., "TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity," Nature, vol. 473, no. 7347, pp. 343–348, May 2011.

[163] L. J. Melchers et al., "Detection of HPV-associated oropharyngeal tumours in a 16-year cohort: more than meets the eye," Br. J. Cancer, vol. 112, no. 8, pp. 1349–1357, Apr. 2015.

[164] R. Siegel, J. Ma, Z. Zou, and A. Jemal, "Cancer statistics, 2014," CA. Cancer J. Clin., vol. 64, no. 1, pp. 9–29, Jan. 2014.

[165] S. Demokan et al., "KIF1A and EDNRB are differentially methylated in primary HNSCC and salivary rinses," Int. J. cancer.Journal Int. du cancer, vol. 127, no. 10, pp. 2351–2359, Nov. 2010.

[166] R. Guerrero-Preston et al., "NID2 and HOXA9 promoter hypermethylation as biomarkers for prevention and early detection in oral cavity squamous cell carcinoma tissues and saliva," Cancer Prev. Res. (Phila)., vol. 4, no. 7, pp. 1061–1072, Jul. 2011.

[167] W. Sun et al., "Detection of TIMP3 promoter hypermethylation in salivary rinse as an independent predictor of local recurrence-free survival in head and neck cancer," Clin. Cancer Res., vol. 18, no. 4, pp. 1082–1091, Feb. 2012.

[168] S. S. Nair et al., "Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias," Epigenetics, vol. 6, no. 1, pp. 34–44, Jan. 2011.

[169] P. G. Arduino et al., "Clinical and histopathologic independent prognostic factors in oral squamous cell carcinoma: a retrospective study of 334 cases.," J. Oral Maxillofac. Surg., vol. 66, no. 8, pp. 1570–9, Aug. 2008.

[170] L. J. Liao, W. C. Lo, W. L. Hsu, C. T. Wang, and M. S. Lai, "Detection of cervical lymph node metastasis in head and neck cancer patients with clinically N0 neck-a meta-analysis comparing different imaging modalities," BMC Cancer, vol. 12, p. 236, Jun. 2012.

[171] M. S. Vorburger et al., "Validity of frozen section in sentinel lymph node biopsy for the staging in oral and oropharyngeal squamous cell carcinoma.," J. Surg. Oncol., vol. 106, no. 7, pp. 816–9, Dec. 2012.

[172] A. Terada et al., "Sentinel lymph node radiolocalization in clinically negative neck oral cancer.," Head Neck, vol. 28, no. 2, pp. 114–20, Feb. 2006.

[173] M. Esteller, "Cancer epigenomics: DNA methylomes and histone-modification maps.," Nat. Rev. Genet., vol. 8, no. 4,

pp. 286–98, Apr. 2007.

[174] F. Roossink, S. de Jong, G. B. Wisman, A. G. van der Zee, and E. Schuuring, "DNA hypermethylation biomarkers to predict response to cisplatin treatment, radiotherapy or chemoradiation: the present state of art," Cell. Oncol. (Dordr)., vol. 35, no. 4, pp. 231–241, Aug. 2012.

[175] D. J. Smiraglia et al., "Differential targets of CpG island hypermethylation in primary and metastatic head and neck squamous cell carcinoma (HNSCC).," J. Med. Genet., vol. 40, no. 1, pp. 25–33, Jan. 2003.

[176] M. Rius and F. Lyko, "Epigenetic cancer therapy: rationales, targets and drugs.," Oncogene, vol. 31, no. 39, pp. 4257–65, Sep. 2012.

[177] A. Cock-Rada and J. B. Weitzman, "The methylation landscape of tumour metastasis.," Biol. cell, vol. 105, no. 2, pp. 73–90, Feb. 2013.

[178] L. S. Kristensen and L. L. Hansen, "PCR-based methods for detecting single-locus DNA methylation biomarkers in cancer diagnostics, prognostics, and response to treatment.," Clin. Chem., vol. 55, no. 8, pp. 1471–83, Aug. 2009.

[179] F. V Jacinto and M. Esteller, "MGMT hypermethylation: a prognostic foe, a predictive friend.," DNA Repair (Amst)., vol. 6, no. 8, pp. 1155–60, Aug. 2007.

[180] S. Borges et al., "Pharmacologic reversion of epigenetic silencing of the PRKD1promoter blocks breast tumor cell invasion and metastasis," Breast Cancer Res., vol. 15, no. 2, p. R66, Apr. 2013.

[181] C. E. Schmalbach et al., "Molecular Profiling and the Identification of Genes Associated With Metastatic Oral Cavity/Pharynx Squamous Cell Carcinoma," Arch. Otolaryngol. Neck Surg., vol. 130, no. 3, p. 295, Mar. 2004.

[182] E. Méndez et al., "A genetic expression profile associated with oral cancer identifies a group of patients at high risk of poor survival.," Clin. Cancer Res., vol. 15, no. 4, pp. 1353–61, Feb. 2009.

[183] R. K. O'Donnell et al., "Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity.," Oncogene, vol. 24, no. 7, pp. 1244–51, Feb. 2005.

[184] M. G. Noordhuis et al., "Involvement of the TGF-beta and beta-catenin pathways in pelvic lymph node metastasis in early-stage cervical cancer.," Clin. Cancer Res., vol. 17, no. 6, pp. 1317–30, Mar. 2011.

[185] N. F. Erdem, E. R. Carlson, and D. A. Gerard, "Characterization of gene expression profiles of 3 different human oral squamous cell carcinoma cell lines with different invasion and metastatic capacities.," J. Oral Maxillofac. Surg., vol. 66, no. 5, pp. 918–27, May 2008.

[186] T. A. Martin, R. E. Mansel, and W. G. Jiang, "Loss of occludin leads to the progression of human breast cancer," Int. J. Mol. Med., vol. 26, no. 5, pp. 723–34, Sep. 2010.

[187] D. Blanco et al., "Altered expression of adhesion molecules and epithelial–mesenchymal transition in silica-induced rat lung carcinogenesis," Lab. Investig., vol. 84, no. 8, pp. 999–1012, Aug. 2004.

[188] A. Perez et al., "CD44 interacts with EGFR and promotes head and neck squamous cell carcinoma initiation and progression.," Oral Oncol., vol. 49, no. 4, pp. 306–13, Apr.

2013.

[189] G. Šupić, R. Kozomara, M. Branković-Magić, N. Jović, and Z. Magić, "Gene hypermethylation in tumor tissue of advanced oral squamous cell carcinoma patients," Oral Oncol., vol. 45, no. 12, pp. 1051–1057, Dec. 2009.

[190] R. P. Dikshit et al., "Hypermethylation, risk factors, clinical characteristics, and survival in 235 patients with laryngeal and hypopharyngeal cancers," Cancer, vol. 110, no. 8, pp. 1745–1751, Oct. 2007.

[191] S. K. Puri, L. Si, C. Y. Fan, and E. Hanna, "Aberrant promoter hypermethylation of multiple genes in head and neck squamous cell carcinoma," Am. J. Otolaryngol., vol. 26, no. 1, pp. 12–17, Jan. 2005.

[192] H. M. Tawfik, N. M. R. A. El-Maqsoud, B. H. A. A. Hak, and Y. M. El-Sherbiny, "Head and neck squamous cell carcinoma: mismatch repair immunohistochemistry and promoter hypermethylation of hMLH1 gene," Am. J. Otolaryngol., Feb. 2011.

[193] R. A. Irizarry et al., "The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores," Nat. Genet., vol. 41, no. 2, pp. 178–186, Feb. 2009.

[194] S. L. Rosas et al., "Promoter hypermethylation patterns of p16, O6-methylguanine-DNA-methyltransferase, and death-associated protein kinase in tumors and saliva of head and neck cancer patients," Cancer Res., vol. 61, no. 3, pp. 939–942, Feb. 2001.

[195] M. Esteller, S. R. Hamilton, P. C. Burger, S. B. Baylin, and J. G. Herman, "Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia.," Cancer Res., vol. 59, no. 4, pp. 793–7, Feb. 1999.

[196] J. J. M. van Dongen et al., "Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936," Leukemia, vol. 17, no. 12, pp. 2257–2317, Dec. 2003.

[197] R. C. Grafstrom, A. E. Pegg, B. F. Trump, and C. C. Harris, "O6-alkylguanine-DNA alkyltransferase activity in normal human tissues and cells.," Cancer Res., vol. 44, no. 7, pp. 2855–7, Jul. 1984.

[198] M. J. Rodríguez, A. Acha, M. T. Ruesga, C. Rodríguez, J. M. Rivera, and J. M. Aguirre, "Loss of expression of DNA repair enzyme MGMT in oral leukoplakia and early oral squamous cell carcinoma. A prognostic tool?," Cancer Lett., vol. 245, no. 1–2, pp. 263–268, Jan. 2007.

[199] M. Sawhney et al., "MGMT expression in oral precancerous and cancerous lesions: Correlation with progression, nodal metastasis and poor prognosis," Oral Oncol., vol. 43, no. 5, pp. 515–522, May 2007.

[200] S.-H. Huang, H.-S. Lee, K. Mar, D.-D. Ji, M.-S. Huang, and K.-T. Hsia, "Loss expression of O6-methylguanine DNA methyltransferase by promoter hypermethylation and its relationship to betel quid chewing in oral squamous cell carcinoma.," Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod., vol. 109, no. 6, pp. 883–9, Jun. 2010.

[201] W. T. Chen, W. C. Hung, W. Y. Kang, Y. C. Huang, and C. Y. Chai, "Urothelial carcinomas arising in arsenic-contaminated areas are associated with hypermethylation of the gene promoter of the death-associated protein kinase," Histopathology, vol. 51, no. 6, pp. 785–792, Dec. 2007.

[202] K. Kawaguchi et al., "Death-associated protein kinase (DAP kinase) alteration in soft tissue leiomyosarcoma: Promoter methylation or homozygous deletion is associated with a loss of DAP kinase expression.," Hum. Pathol., vol. 35, no. 10, pp. 1266–71, Oct. 2004.

[203] P.-F. Su, W.-L. Huang, H.-T. Wu, C.-H. Wu, T.-Y. Liu, and S.-Y. Kao, "p16(INK4A) promoter hypermethylation is associated with invasiveness and prognosis of oral squamous cell carcinoma in an age-dependent manner.," Oral Oncol., vol. 46, no. 10, pp. 734–9, Oct. 2010.

[204] V. T. Cao et al., "The correlation and prognostic significance of MGMT promoter methylation and MGMT protein in glioblastomas.," Neurosurgery, vol. 65, no. 5, p. 866–75; discussion 875, Nov. 2009.

[205] I. Silva de Meneses, R. Reis de Souza, V. de Lourdes Sierpe Jeraldo, D. Rodrigues Ribeiro Cavalcante, F. Prado Reis, and R. Luiz Cavalcanti de Albuquerque Júnior, "Death-Associated Protein Kinase is Underexpressed in High-Grade Oral Squamous Cell Carcinoma," 2010.

[206] M. J. Worsham, K. M. Chen, T. Ghanem, J. K. Stephen, and G. Divine, "Epigenetic modulation of signal transduction pathways in HPV-associated HNSCC," Otolaryngol. Head. Neck Surg., vol. 149, no. 3, pp. 409–416, Sep. 2013.

[207] M. Rastetter et al., "Frequent intra-tumoural heterogeneity of promoter hypermethylation in malignant melanoma.," Histol. Histopathol., vol. 22, no. 9, pp. 1005–15, 2007.

[208] B. J. Braakhuis, R. H. Brakenhoff, and C. R. Leemans, "Gene expression profiling in head and neck squamous cell carcinoma," Curr. Opin. Otolaryngol. Head Neck Surg., vol. 18, no. 2, pp. 67–71, Apr. 2010.

[209] D. Gozuacik and A. Kimchi, "DAPk protein family and cancer," Autophagy, vol. 2, no. 2. pp. 74–79, 2006.

[210] M. Sanchez-Cespedes et al., "Gene promoter hypermethylation in tumors and serum of head and neck cancer patients.," Cancer Res., vol. 60, no. 4, pp. 892–5, Feb. 2000.

[211] W. J. Wang, J. C. Kuo, C. C. Yao, and R. H. Chen, "DAP-kinase induces apoptosis by suppressing integrin activity and disrupting matrix survival signals.," J. Cell Biol., vol. 159, no. 1, pp. 169–79, Oct. 2002.

[212] J. C. Kuo, W. J. Wang, C. C. Yao, P. R. Wu, and R. H. Chen, "The tumor suppressor DAPK inhibits cell motility by blocking the integrin-mediated polarity pathway.," J. Cell Biol., vol. 172, no. 4, pp. 619–31, Feb. 2006.

[213] R. A. Olson, P. K. Brastianos, and D. A. Palma, "Prognostic and predictive value of epigenetic silencing of MGMT in patients with high grade gliomas: a systematic review and meta-analysis.," J. Neurooncol., vol. 105, no. 2, pp. 325–35, Nov. 2011.

[214] S. Zeilinger et al., "Tobacco smoking leads to extensive genome-wide changes in DNA methylation," PLoS One, vol. 8, no. 5, p. e63812, May 2013.

[215] M. Christmann and B. Kaina, "O(6)-methylguanine-DNA methyltransferase (MGMT): impact on cancer risk in response to tobacco smoke.," Mutat. Res., vol. 736, no. 1–2, pp. 64–74, Aug. 2012.

[216] P. Lassen, "The role of Human papillomavirus in head and neck cancer and the impact on radiotherapy outcome," Radiother. Oncol., vol. 95, no. 3, pp. 371–380, Jun. 2010.

[217] D. Weiss, T. Basel, F. Sachse, A. Braeuninger, and C. Rudack, "Promoter methylation of cyclin A1 is associated with human papillomavirus 16 induced head and neck squamous cell carcinoma independently of p53 mutation," Mol. Carcinog., vol. 50, no. 9, pp. 680–688, Sep. 2011.

[218] B. Ingold, P. Schraml, F. L. Heppner, and H. Moch, "Homogeneous MGMT immunoreactivity correlates with an unmethylated MGMT promoter status in brain metastases of various solid tumors.," PLoS One, vol. 4, no. 3, p. e4775, Mar. 2009.

[219] L. W. T. Alkureishi et al., "Sentinel node biopsy in head and neck squamous cell cancer: 5-year follow-up of a European multicenter trial.," Ann. Surg. Oncol., vol. 17, no. 9, pp. 2459–64, Sep. 2010.

[220] C. Bock et al., "Quantitative comparison of genome-wide DNA methylation mapping technologies," Nat. Biotechnol., vol. 28, no. 10, pp. 1106–1114, Oct. 2010.

[221] A. Ferlito et al., "Prognostic significance of microscopic and macroscopic extracapsular spread from metastatic tumor in the cervical lymph nodes," Oral Oncol., vol. 38, no. 8, pp. 747–751, Dec. 2002.

[222] M. Magnano et al., "Prognostic factors of cervical lymph node metastasis in head and neck squamous cell carcinoma," Tumori, vol. 83, no. 6, pp. 922–926, 1997.

[223] G. Mamelle, J. Pampurik, B. Luboinski, R. Lancar, A. Lusinchi, and J. Bosq, "Lymph node prognostic factors in head and neck squamous cell carcinomas," Am. J. Surg., vol. 168, no. 5, pp. 494–498, Nov. 1994.

[224] C. G. Gourin, B. T. Conger, E. S. Porubsky, W. C. Sheils, P. A. Bilodeau, and T. A. Coleman, "The effect of occult nodal metastases on survival and regional control in patients with head and neck squamous cell carcinoma," Laryngoscope, vol. 118, no. 7, pp. 1191–1194, Jul. 2008.

[225] N. Azad, C. A. Zahnow, C. M. Rudin, and S. B. Baylin, "The future of epigenetic therapy in solid tumours--lessons from the past," Nat. Rev. Oncol., vol. 10, no. 5, pp. 256–266, May 2013.

[226] N. Ahuja, H. Easwaran, and S. B. Baylin, "Harnessing the potential of epigenetic therapy to target solid tumors," J. Clin. Invest., vol. 124, no. 1, pp. 56–63, Jan. 2014.

[227] M. Esteller et al., "Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents," N. Engl. J. Med., vol. 343, no. 19, pp. 1350–1354, Nov. 2000.

[228] H. Heyn and M. Esteller, "DNA methylation profiling in the clinic: applications and challenges," Nat. Rev., vol. 13, no. 10, pp. 679–692, Oct. 2012.

[229] M. J. Kwon et al., "TWIST1 promoter methylation is associated with prognosis in tonsillar squamous cell carcinoma," Hum. Pathol., vol. 44, no. 9, pp. 1722–1729, Sep. 2013.

[230] T. Gao et al., "H19 DMR methylation correlates to the progression of esophageal squamous cell carcinoma through IGF2 imprinting pathway," Clin. Transl. Oncol., vol. 16, pp. 410–417, Aug. 2013.

[231] S. Pierini et al., "Promoter hypermethylation of CDKN2A, MGMT, MLH1 and DAPK genes in laryngeal squamous cell carcinoma and their associations with clinical profiles of the patients," Head Neck, vol. 36, pp. 1103–1108, Jun. 2013.

[232] A. B. Brinkman, F. Simmer, K. Ma, A. Kaan, J. Zhu, and H. G. Stunnenberg, "Whole-genome DNA methylation profiling using MethylCap-seq," Methods, vol. 52, no. 3, pp. 232–236, Nov. 2010.

[233] S. J. Smeets et al., "A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen," Int. J. cancer.Journal Int. du cancer, vol. 121, no. 11, pp. 2465–2472, Dec. 2007.

[234] N. Reesink-Peters et al., "Detecting cervical cancer by quantitative promoter hypermethylation assay on cervical scrapings: a feasibility study," Mol. Cancer Res., vol. 2, no. 5, pp. 289–295, May 2004.

[235] N. Yang et al., "Methylation markers for CCNA1 and C13ORF18 are strongly associated with high-grade cervical intraepithelial neoplasia and cervical cancer in cervical scrapings," Cancer Epidemiol. Biomarkers Prev., vol. 18, no. 11, pp. 3000–3007, Nov. 2009.

[236] M. H. Oonk et al., "Identification of inguinofemoral lymph node metastases by methylation markers in vulvar cancer," Gynecol. Oncol., vol. 125, no. 2, pp. 352–357, May 2012.

[237] E. J. Lee et al., "Targeted bisulfite sequencing by solution hybrid selection and massively parallel sequencing," Nucleic Acids Res., vol. 39, no. 19, p. e127, Oct. 2011.

[238] R. C. Team, R. Ihaka, R. Gentleman, and R. R Development Core Team, "R: A Language and Environment for Statistical Computing," R Foundation for Statistical Computing, 2011. [Online]. Available: http://www.r-project.org.

[239] T. J. Hardcastle and K. A. Kelly, "baySeq: empirical Bayesian methods for identifying differential expression in sequence count data," BMC Bioinformatics, vol. 11, pp. 422–435, Aug. 2010.

[240] Y. Nagai et al., "Clinical significance of Wnt-induced secreted protein-1 (WISP-1/CCN4) in esophageal squamous cell carcinoma," Anticancer Res., vol. 31, no. 3, pp. 991–997, Mar. 2011.

[241] M. Uhlen et al., "Towards a knowledge-based Human Protein Atlas," Nat. Biotechnol., vol. 28, no. 12, pp. 1248–1250, Dec. 2010.

[242] M. E. Price et al., "Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array," Epigenetics Chromatin, vol. 6, no. 1, p. 4, Mar. 2013.

[243] R. C. Gentleman et al., "Bioconductor: open software development for computational biology and bioinformatics," Genome Biol., vol. 5, no. 10, p. R80, 2004.

[244] G. K. Smyth, "Linear models and empirical bayes methods for assessing differential expression in microarray experiments," Stat. Appl. Genet. Mol. Biol., vol. 3, p. Article3, 2004.

[245] W. J. Kent et al., "The human genome browser at UCSC," Genome Res., vol. 12, no. 6, pp. 996–1006, Jun. 2002.

[246] J. Sandoval et al., "Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome," Epigenetics, vol. 6, no. 6, pp. 692–702, Jun. 2011.

[247] T. T. Huang et al., "Epigenetic deregulation of the anaplastic lymphoma kinase gene modulates mesenchymal

characteristics of oral squamous cell carcinomas," Carcinogenesis, vol. 34, no. 8, pp. 1717–1727, Aug. 2013.

[248] Q. Sun et al., "Dysregulated miR-363 affects head and neck cancer invasion and metastasis by targeting podoplanin," Int. J. Biochem. Cell Biol., vol. 45, no. 3, pp. 513–520, Mar. 2013.

[249] G. Supic, R. Kozomara, N. Jovic, K. Zeljic, and Z. Magic, "Hypermethylation of RUNX3 but not WIF1 gene and its association with stage and nodal status of tongue cancers," Oral Dis., vol. 17, no. 8, pp. 794–800, Nov. 2011.

[250] H. De Schutter, H. Geeraerts, E. Verbeken, and S. Nuyts, "Promoter methylation of TIMP3 and CDH1 predicts better outcome in head and neck squamous cell carcinoma treated by radiotherapy only," Oncol. Rep., vol. 21, no. 2, pp. 507–513, Feb. 2009.

[251] S. M. Langevin et al., "MicroRNA-137 promoter methylation is associated with poorer overall survival in patients with squamous cell carcinoma of the head and neck," Cancer, vol. 117, no. 7, pp. 1454–1462, Apr. 2011.

[252] L. M. Arantes, A. C. de Carvalho, M. E. Melendez, A. L. Carvalho, and E. M. Goloni-Bertollo, "Methylation as a biomarker for head and neck cancer," Oral Oncol., vol. 50, pp. 587–592, Mar. 2014.

[253] P. P. Chen et al., "Expression of Cyr61, CTGF, and WISP-1 correlates with clinical features of lung cancer," PLoS One, vol. 2, no. 6, p. e534, Jun. 2007.

[254] S. R. Davies, M. L. Davies, A. Sanders, C. Parr, J. Torkington, and W. G. Jiang, "Differential expression of the CCN family member WISP-1, WISP-2 and WISP-3 in human colorectal cancer and the prognostic implications," Int. J. Oncol., vol. 36, no. 5, pp. 1129–1136, May 2010.

[255] D. Pennica et al., "WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors," Proc. Natl. Acad. Sci. U. S. A., vol. 95, no. 25, pp. 14717–14722, Dec. 1998.

[256] H. Shao, L. Cai, J. M. Grichnik, A. S. Livingstone, O. C. Velazquez, and Z. J. Liu, "Activation of Notch1 signaling in stromal fibroblasts inhibits melanoma growth by upregulating WISP-1," Oncogene, vol. 30, no. 42, pp. 4316–4326, Oct. 2011.

[257] C. Tian et al., "Overexpression of connective tissue growth factor WISP-1 in Chinese primary rectal cancer patients," World J. Gastroenterol., vol. 13, no. 28, pp. 3878–3882, Jul. 2007.

[258] D. Xie, K. Nakachi, H. Wang, R. Elashoff, and H. P. Koeffler, "Elevated levels of connective tissue growth factor, WISP-1, and CYR61 in primary breast cancers associated with more advanced features," Cancer Res., vol. 61, no. 24, pp. 8917–8923, Dec. 2001.

[259] L. Xu, R. B. Corcoran, J. W. Welsh, D. Pennica, and A. J. Levine, "WISP-1 is a Wnt-1- and beta-catenin-responsive oncogene," Genes Dev., vol. 14, no. 5, pp. 585–595, Mar. 2000.

[260] K. L. Richards et al., "Genome-wide hypomethylation in head and neck cancer is more pronounced in HPV-negative tumors and is associated with genomic instability," PLoS One, vol. 4, no. 3, p. e4941, 2009.

[261] K. M. Pattani et al., "MAGEB2 is activated by promoter demethylation in head and neck squamous cell carcinoma," PLoS One, vol. 7, no. 9, p. e45534, 2012.

[262] R. Warta et al., "Reduced promoter methylation and increased expression of CSPG4 negatively influences survival of HNSCC patients," Int. J. cancer.Journal Int. du cancer, Apr. 2014.

[263] A. Murata et al., "IGF2 DMR0 Methylation, Loss of Imprinting, and Patient Prognosis in Esophageal Squamous Cell Carcinoma," Ann. Surg. Oncol., no. 21, pp. 1166–1174, Dec. 2013.

[264] C. Puttipanyalears, K. Subbalekha, A. Mutirangura, and N. Kitkumthorn, "Alu hypomethylation in smoke-exposed epithelia and oral squamous carcinoma," Asian Pac. J. Cancer Prev., vol. 14, no. 9, pp. 5495–5501, 2013.

[265] N. Kitkumthorn, S. Keelawat, P. Rattanatanyong, and A. Mutirangura, "LINE-1 and Alu methylation patterns in lymph node metastases of head and neck cancers," Asian Pac. J. Cancer Prev., vol. 13, no. 9, pp. 4469–4475, 2012.

[266] J. Jasielec, V. Saloura, and L. A. Godley, "The mechanistic role of DNA methylation in myeloid leukemogenesis," Leukemia, vol. 28, no. 9, pp. 1765–1773, Sep. 2014.

[267] C. C. Chen and L. F. Lau, "Functions and mechanisms of action of CCN matricellular proteins," Int. J. Biochem. Cell Biol., vol. 41, no. 4, pp. 771–783, Apr. 2009.

[268] J. Song, I. Chang, Z. Chen, M. Kang, and C. Y. Wang, "Characterization of side populations in HNSCC: highly invasive, chemoresistant and abnormal Wnt signaling," PLoS One, vol. 5, no. 7, p. e11456, Jul. 2010.

[269] L. Lo Muzio, "A possible role for the WNT-1 pathway in oral carcinogenesis," Crit. Rev. Oral Biol. Med., vol. 12, no. 2, pp. 152–165, 2001.

[270] F. Su, M. Overholtzer, D. Besser, and A. J. Levine, "WISP-1 attenuates p53-mediated apoptosis in response to DNA damage through activation of the Akt kinase," Genes Dev., vol. 16, no. 1, pp. 46–57, Jan. 2002.

[271] Y. P. Mosse, A. Wood, and J. M. Maris, "Inhibition of ALK signaling for cancer therapy," Clin. Cancer Res., vol. 15, no. 18, pp. 5609–5614, Sep. 2009.

[272] M. K. Layland, D. G. Sessions, and J. Lenox, "The influence of lymph node metastasis in the treatment of squamous cell carcinoma of the oral cavity, oropharynx, larynx, and hypopharynx: N0 versus N+.," Laryngoscope, vol. 115, no. 4, pp. 629–39, Apr. 2005.

[273] C. K. (eds). Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, "SEER Cancer Statistics Review (CSR), 1975-2012," 2014. [Online]. Available: http://seer.cancer.gov/csr/1975_2012/.

[274] R. J. Baatenburg de Jong, R. J. Rongen, J. S. Laméris, M. Harthoorn, C. D. Verwoerd, and P. Knegt, "Metastatic neck disease. Palpation vs ultrasound examination.," Arch. Otolaryngol. Head. Neck Surg., vol. 115, no. 6, pp. 689–90, Jun. 1989.

[275] H. Schöder et al., "18F-FDG PET/CT for detecting nodal metastases in patients with oral cancer staged N0 by clinical examination and CT/MRI.," J. Nucl. Med., vol. 47, no. 5, pp. 755–62, May 2006.

[276] H. R. Chu, J. H. Kim, D. Y. Yoon, H. S. Hwang, and Y.-S. Rho, "Additional diagnostic value of (18)F-FDG PET-CT in detecting retropharyngeal nodal metastases.," Otolaryngol. Head. Neck Surg., vol. 141, no. 5, pp. 633–8, Nov. 2009.

[277] P. Zhang, J. Wang, W. Gao, B. Z. Yuan, J. Rogers, and E. Reed, "CHK2 kinase expression is down-regulated due to promoter methylation in non-small cell lung cancer," Mol. Cancer, vol. 3, p. 14, May 2004.

[278] L. J. Melchers et al., "Identification of methylation markers for the prediction of nodal metastasis in oral and oropharyngeal squamous cell carcinoma," Epigenetics, vol. 10, no. 9, pp. 850–860, Jan. 2015.

[279] M. J. A. M. Clausen et al., "Identification and validation of WISP1 as an epigenetic regulator of metastasis in oral squamous cell carcinoma.," Genes. Chromosomes Cancer, vol. 55, no. 1, pp. 45–59, Sep. 2016.

[280] C. B. Yoo and P. A. Jones, "Epigenetic therapy of cancer: past, present and future.," Nat. Rev. Drug Discov., vol. 5, no. 1, pp. 37–50, Jan. 2006.

[281] P. Roepman, P. Kemmeren, L. F. Wessels, P. J. Slootweg, and F. C. Holstege, "Multiple robust signatures for detecting lymph node metastasis in head and neck cancer," Cancer Res., vol. 66, no. 4, pp. 2361–2366, Feb. 2006.

[282] J. J. Eijsink et al., "A four-gene methylation marker panel as triage test in high-risk human papillomavirus positive patients," Int. J. cancer.Journal Int. du cancer, vol. 130, no. 8, pp. 1861–1869, Apr. 2012.

[283] B. Langmead, C. Trapnell, M. Pop, and S. L. Salzberg, "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome," Genome Biol., vol. 10, no. 3, p. R25–2009–10–3–r25. Epub 2009 Mar 4, 2009.

[284] P. Du, W. A. Kibbe, and S. M. Lin, "lumi: a pipeline for processing Illumina microarray," Bioinformatics, vol. 24, no. 13, pp. 1547–1548, Jul. 2008.

[285] M. E. Ritchie et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," Nucleic Acids Res., Jan. 2015.

[286] E. Cerami et al., "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," Cancer Discov., vol. 2, no. 5, pp. 401–404, May 2012.

[287] J. Gao et al., "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal," Sci. Signal., vol. 6, no. 269, p. pl1, Apr. 2013.

[288] R. C. Team, "R: A Language and Environment for Statistical Computing," 2014. .

[289] N. Wong and X. Wang, "miRDB: an online resource for microRNA target prediction and functional annotations," Nucleic Acids Res., vol. 43, no. Database issue, pp. D146-52, Jan. 2015.

[290] J. M. Cheng, M. Ding, A. Aribi, P. Shah, and K. Rao, "Loss of RAB25 expression in breast cancer," Int. J. cancer.Journal Int. du cancer, vol. 118, no. 12, pp. 2957–2964, Jun. 2006.

[291] M. Tong et al., "Rab25 is a tumor suppressor gene with antiangiogenic and anti-invasive activities in esophageal squamous cell carcinoma," Cancer Res., vol. 72, no. 22, pp. 6024–6035, Nov. 2012.

[292] M. Tellez-Gabriel et al., "High RAB25 expression is associated with good clinical outcome in patients with locally advanced head and neck squamous cell carcinoma," Cancer Med., vol. 2, no. 6, pp. 950–963, Dec. 2013.

[293] S. Mitra, K. W. Cheng, and G. B. Mills, "Rab25 in cancer: a brief update," Biochem. Soc. Trans., vol. 40, no. 6, pp. 1404–1408, Dec. 2012.

[294] C. E. Chua and B. L. Tang, "The role of the small GTPase Rab31 in cancer," J. Cell. Mol. Med., vol. 19, no. 1, pp. 1–10, Jan. 2015.

[295] K. W. Cheng, J. P. Lahad, J. W. Gray, and G. B. Mills, "Emerging role of RAB GTPases in cancer and human disease," Cancer Res., vol. 65, no. 7, pp. 2516–2519, Apr. 2005.

[296] C. Recchi and M. C. Seabra, "Novel functions for Rab GTPases in multiple aspects of tumour progression," Biochem. Soc. Trans., vol. 40, no. 6, pp. 1398–1403, Dec. 2012.

[297] K. W. Cheng et al., "The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers," Nat. Med., vol. 10, no. 11, pp. 1251–1256, Nov. 2004.

[298] J. Zhang et al., "Overexpression of Rab25 contributes to metastasis of bladder cancer through induction of epithelial-mesenchymal transition and activation of Akt/GSK-3beta/Snail signaling," Carcinogenesis, vol. 34, no. 10, pp. 2401–2408, Oct. 2013.

[299] C. Cao, C. Lu, J. Xu, J. Zhang, J. Zhang, and M. Li, "Expression of Rab25 correlates with the invasion and metastasis of gastric cancer," Chin. J. Cancer Res., vol. 25, no. 2, pp. 192–199, Apr. 2013.

[300] Y. X. Yin et al., "Increased expression of Rab25 in breast cancer correlates with lymphatic metastasis," Tumour Biol., vol. 33, no. 5, pp. 1581–1587, Oct. 2012.

[301] P. Amornphimoltham et al., "Rab25 Regulates Invasion and Metastasis in Head and Neck Cancer," Clin. Cancer Res., Feb. 2013.

[302] J. M. Cheng, L. Volk, D. K. Janaki, S. Vyakaranam, S. Ran, and K. A. Rao, "Tumor suppressor function of Rab25 in triple-negative breast cancer," Int. J. cancer.Journal Int. du cancer, vol. 126, no. 12, pp. 2799–2812, Jun. 2010.

[303] K. O. Wrzeszczynski et al., "Identification of tumor suppressors and oncogenes from genomic and epigenetic features in ovarian cancer," PLoS One, vol. 6, no. 12, p. e28503, 2011.

[304] C. Y. Wu, R. C. Tseng, H. S. Hsu, Y. C. Wang, and M. T. Hsu, "Frequent down-regulation of hRAB37 in metastatic tumor by genetic and epigenetic mechanisms in lung cancer," Lung Cancer, vol. 63, no. 3, pp. 360–367, Mar. 2009.

[305] T. Tanaka et al., "Epigenetic silencing of microRNA-373 plays an important role in regulating cell proliferation in colon cancer," Oncol. Rep., vol. 26, no. 5, pp. 1329–1335, Nov. 2011.

[306] M. J. A. M. Clausen et al., "RAB25 expression is epigenetically downregulated in oral and oropharyngeal squamous cell carcinoma with lymph node metastasis.," Epigenetics, vol. 11, no. 9, pp. 653–663, Sep. 2016.

[307] L. J. Melchers et al., "Head and neck squamous cell carcinomas do not express EGFRvIII," Int. J. Radiat. Oncol. Biol. Phys., vol. 90, no. 2, 2014.

[308] C. T. Chone et al., "Impact of immunohistochemistry in sentinel lymph node biopsy in head and neck cancer," Eur. Arch. Oto-Rhino-Laryngology, vol. 270, no. 1, pp. 313–317, Jan. 2013.

[309] "Dutch guideline head and neck cancer," 2015.

[310] M. C. de Jong et al., "Pretreatment microRNA Expression Impacting on Epithelial-to-Mesenchymal Transition Predicts Intrinsic Radiosensitivity in Head and Neck Cancer Cell Lines and Patients," Clin. Cancer Res., vol. 21, no. 24, pp. 5630–5638,

Dec. 2015.

[311] H. J. T. Van Zeeburg et al., "Generation and Molecular Characterization of Head and Neck Squamous Cell Lines of Fanconi Anemia Patients," 2005.

[312] J. H. Gibcus et al., "High-resolution mapping identifies a commonly amplified 11q13.3 region containing multiple genes flanked by segmental duplications.," Hum. Genet., vol. 121, no. 2, pp. 187–201, Apr. 2007.

[313] A. Radonić, S. Thulke, I. M. Mackay, O. Landt, W. Siegert, and A. Nitsche, "Guideline to reference gene selection for quantitative real-time PCR.," Biochem. Biophys. Res. Commun., vol. 313, no. 4, pp. 856–62, Jan. 2004.

[314] K. J. Livak and T. D. Schmittgen, "Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method," Methods, vol. 25, no. 4, pp. 402–408, Dec. 2001.

[315] V. Bovenzi and R. L. Momparler, "Antineoplastic action of 5-aza-2′-deoxycytidine and histone deacetylase inhibitor and their effect on the expression of retinoic acid receptor beta and estrogen receptor alpha genes in breast carcinoma cells.," Cancer Chemother. Pharmacol., vol. 48, no. 1, pp. 71–6, Jul. 2001.

[316] H. Chen, C. Xu, Q. Jin, and Z. Liu, "S100 protein family in human cancer," Am. J. Cancer Res., vol. 4, no. 2, pp. 89–115, Mar. 2014.

[317] J. Markowitz and W. E. Carson 3rd, "Review of S100A9 biology and its role in cancer," Biochim. Biophys. Acta, vol. 1835, no. 1, pp. 100–109, Jan. 2013.

[318] L. Ma, P. Sun, J.-C. Zhang, Q. Zhang, and S.-L. Yao, "Proinflammatory effects of S100A8/A9 via TLR4 and RAGE signaling pathways in BV-2 microglial cells.," Int. J. Mol. Med., vol. 40, no. 1, pp. 31–38, Jul. 2017.

[319] Y. H. Xiao et al., "Identification of GLIPR1 tumor suppressor as methylation-silenced gene in acute myeloid leukemia by microarray analysis," J. Cancer Res. Clin. Oncol., vol. 137, no. 12, pp. 1831–1840, Dec. 2011.

[320] Z. Li et al., "Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias," Cancer Res., vol. 69, no. 3, pp. 1109–1116, Feb. 2009.

[321] J. H. Choi et al., "Identification of S100A8 and S100A9 as negative regulators for lymph node metastasis of gastric adenocarcinoma," Histol. Histopathol., vol. 27, no. 11, pp. 1439–1448, Nov. 2012.

[322] G. A. Jeon et al., "Global gene expression profiles of human head and neck squamous carcinoma cell lines," Int. J. cancer. Journal Int. du cancer, vol. 112, no. 2, pp. 249–258, Nov. 2004.

[323] M. Roesch Ely et al., "Transcript and proteome analysis reveals reduced expression of calgranulins in head and neck squamous cell carcinoma," Eur. J. Cell Biol., vol. 84, no. 2–3, pp. 431–444, Mar. 2005.

[324] C. Melle et al., "A technical triade for proteomic identification and characterization of cancer biomarkers," Cancer Res., vol. 64, no. 12, pp. 4099–4104, Jun. 2004.

[325] H. E. Gonzalez et al., "Identification of 9 genes differentially expressed in head and neck squamous cell carcinoma," Arch. Otolaryngol. Head. Neck Surg., vol. 129, no. 7, pp. 754–759, Jul. 2003.

[326] A. Luo et al., "Discovery of Ca2+-relevant and differentiation-associated genes downregulated in esophageal squamous cell carcinoma using cDNA microarray," Oncogene, vol. 23, no. 6, pp. 1291–1299, Feb. 2004.

[327] N. J. Silveira et al., "Searching for molecular markers in head and neck squamous cell carcinomas (HNSCC) by statistical and bioinformatic analysis of larynx-derived SAGE libraries," BMC Med. Genomics, vol. 1, p. 56, Nov. 2008.

[328] X. Dong et al., "Modeling gene expression using chromatin features in various cellular contexts," Genome Biol., vol. 13, no. 9, pp. R53-2012-13-9-r53, Jun. 2012.

[329] J. T. Bell et al., "DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines," Genome Biol., vol. 12, no. 1, p. R10–2011–12–1–r10. Epub 2011 Jan 20, 2011.

[330] R. G. Urdinguio et al., "Mecp2-null mice provide new neuronal targets for Rett syndrome," PLoS One, vol. 3, no. 11, p. e3669, 2008.

[331] D. K. Rhee, S. H. Park, and Y. K. Jang, "Molecular signatures associated with transformation and progression to breast cancer in the isogenic MCF10 model," Genomics, vol. 92, no. 6, pp. 419–428, Dec. 2008.

[332] R. A. Harris et al., "DNA methylation-associated colonic mucosal immune and defense responses in treatment-naive pediatric ulcerative colitis," Epigenetics, vol. 9, no. 8, Jun. 2014.

[333] O. Y. Dokun, A. R. Florl, H. H. Seifert, I. Wolff, and W. A. Schulz, "Relationship of SNCG, S100A4, S100A9 and LCN2 gene expression and DNA methylation in bladder cancer," Int. J. cancer.Journal Int. du cancer, vol. 123, no. 12, pp. 2798–2807, Dec. 2008.

[334] M. Ichikawa, R. Williams, L. Wang, T. Vogl, and G. Srikrishna, "S100A8/A9 activate key genes and pathways in colon tumor progression," Mol. Cancer Res., vol. 9, no. 2, pp. 133–148, Feb. 2011.

[335] S. Ghavami, C. Kerkhoff, M. Los, M. Hashemi, C. Sorg, and F. Karami-Tehrani, "Mechanism of apoptosis induced by S100A8/A9 in colon cancer cell lines: the role of ROS and the effect of metal ions," J. Leukoc. Biol., vol. 76, no. 1, pp. 169–175, Jul. 2004.

[336] S. Ghavami et al., "S100A8/A9 at low concentration promotes tumor cell growth via RAGE ligation and MAP kinase-dependent pathway," J. Leukoc. Biol., vol. 83, no. 6, pp. 1484–1492, Jun. 2008.

[337] O. Turovskaya et al., "RAGE, carboxylated glycans and S100A8/A9 play essential roles in colitis-associated carcinogenesis," Carcinogenesis, vol. 29, no. 10, pp. 2035–2043, Oct. 2008.

[338] A. Hermani, B. De Servi, S. Medunjanin, P. A. Tessier, and D. Mayer, "S100A8 and S100A9 activate MAP kinase and NF-kappaB signaling pathways and trigger translocation of RAGE in human prostate cancer cells," Exp. Cell Res., vol. 312, no. 2, pp. 184–197, Jan. 2006.

[339] S. Hiratsuka, A. Watanabe, H. Aburatani, and Y. Maru, "Tumour-mediated upregulation of chemoattractants and recruitment of myeloid cells predetermines lung metastasis," Nat. Cell Biol., vol. 8, no. 12, pp. 1369–1375, Dec. 2006.

[340] A. Moon et al., "Global gene expression profiling unveils S100A8/A9 as candidate markers in H-ras-mediated human

breast epithelial cell invasion," Mol. Cancer Res., vol. 6, no. 10, pp. 1544–1553, Oct. 2008.

[341] A. Saha, Y. C. Lee, Z. Zhang, G. Chandra, S. B. Su, and A. B. Mukherjee, "Lack of an endogenous anti-inflammatory protein in mice enhances colonization of B16F10 melanoma cells in the lungs," J. Biol. Chem., vol. 285, no. 14, pp. 10822–10831, Apr. 2010.

[342] C. W. Ang et al., "Smad4 loss is associated with fewer S100A8-positive monocytes in colorectal tumors and attenuated response to S100A8 in colorectal and pancreatic cancer cells," Carcinogenesis, vol. 31, no. 9, pp. 1541–1551, Sep. 2010.

[343] S. Hiratsuka et al., "The S100A8-serum amyloid A3-TLR4 paracrine cascade establishes a pre-metastatic phase," Nat. Cell Biol., vol. 10, no. 11, pp. 1349–1355, Nov. 2008.

[344] K. Gumireddy et al., "ID1 Promotes Breast Cancer Metastasis by S100A9 Regulation," Mol. Cancer Res., Jun. 2014.

[345] B. Berschneider and M. Konigshoff, "WNT1 inducible signaling pathway protein 1 (WISP1): a novel mediator linking development and disease," Int. J. Biochem. Cell Biol., vol. 43, no. 3, pp. 306–309, Mar. 2011.

[346] A. D. Kohn and R. T. Moon, "Wnt and calcium signaling: β-Catenin-independent pathways," Cell Calcium, vol. 38, no. 3–4, pp. 439–446, Sep. 2005.

[347] M. Schmidt et al., "Ras-independent activation of the Raf/MEK/ERK pathway upon calcium-induced differentiation of keratinocytes.," J. Biol. Chem., vol. 275, no. 52, pp. 41011–7, Dec. 2000.

[348] B. Chen et al., "S100A9 induced inflammatory responses are mediated by distinct damage associated molecular patterns (DAMP) receptors in vitro and in vivo.," PLoS One, vol. 10, no. 2, p. e0115828, 2015.

[349] D. W. Huang, B. T. Sherman, R. A. Lempicki, W. Huang da, B. T. Sherman, and R. A. Lempicki, "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists," Nucleic Acids Res., vol. 37, no. 1, pp. 1–13, Jan. 2009.

[350] D. W. Huang, B. T. Sherman, R. A. Lempicki, W. Huang da, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," Nat. Protoc., vol. 4, no. 1, pp. 44–57, Jan. 2009.

[351] J. Humeau et al., "Calcium signaling and cell cycle: Progression or death," Cell Calcium, vol. 70, pp. 3–15, Mar. 2018.

[352] M. Xu, A. Seas, M. Kiyani, K. S. Y. Ji, and H. N. Bell, "A temporal examination of calcium signaling in cancer- from tumorigenesis, to immune evasion, and metastasis.," Cell Biosci., vol. 8, p. 25, 2018.

[353] M. C. X. Pinto et al., "Calcium signaling and cell proliferation," Cell. Signal., vol. 27, no. 11, pp. 2139–2149, Nov. 2015.

[354] T. A. Stewart, K. T. D. S. Yapa, and G. R. Monteith, "Altered calcium signaling in cancer cells," Biochim. Biophys. Acta - Biomembr., vol. 1848, no. 10, pp. 2502–2511, Oct. 2015.

[355] F. C. Tsai, G. H. Kuo, S. W. Chang, and P. J. Tsai, "Ca^{2+} Signaling in Cytoskeletal Reorganization, Cell Migration, and Cancer Metastasis," Biomed Res. Int., vol. 2015, pp. 1–13, 2015.

[356] N. Yang et al., "Blockade of store-operated Ca2+ entry inhibits hepatocarcinoma cell migration and invasion by regulating focal adhesion turnover," Cancer Lett., vol. 330, no. 2, pp. 163–169, Apr. 2013.

[357] E. Farahani et al., "Cell adhesion molecules and their relation to (cancer) cell stemness," Carcinogenesis, vol. 35, no. 4, pp. 747–759, Apr. 2014.

[358] C. A. Lyssiotis and A. C. Kimmelman, "Metabolic Interactions in the Tumor Microenvironment.," Trends Cell Biol., vol. 27, no. 11, pp. 863–875, Nov. 2017.

[359] R. E. Dolmetsch, R. S. Lewis, C. C. Goodnow, and J. I. Healy, "Differential activation of transcription factors induced by Ca2+ response amplitude and duration," Nature, vol. 386, no. 6627, pp. 855–858, Apr. 1997.

[360] E. C. Schwarz, B. Qu, and M. Hoth, "Calcium, cancer and killing: The role of calcium in killing cancer cells by cytotoxic T lymphocytes and natural killer cells," Biochim. Biophys. Acta - Mol. Cell Res., vol. 1833, no. 7, pp. 1603–1611, Jul. 2013.

[361] L. Munaron and M. Scianna, "Multilevel complexity of calcium signaling: Modeling angiogenesis.," World J. Biol. Chem., vol. 3, no. 6, pp. 121–6, Jun. 2012.

[362] E. C. Kohn, R. Alessandro, J. Spoonster, R. P. Wersto, and L. A. Liotta, "Angiogenesis: role of calcium-mediated signal transduction.," Proc. Natl. Acad. Sci. U. S. A., vol. 92, no. 5, pp. 1307–11, Feb. 1995.

[363] J. Reckenbeil et al., "Cellular Distribution and Gene Expression Pattern of Metastasin (S100A4), Calgranulin A (S100A8), and Calgranulin B (S100A9) in Oral Lesions as Markers for Molecular Pathology," Cancer Invest., vol. 34, no. 6, pp. 246–254, Jul. 2016.

[364] M. A. Raffat, N. I. Hadi, M. Hosein, S. Mirza, S. Ikram, and Z. Akram, "S100 proteins in oral squamous cell carcinoma," Clin. Chim. Acta, vol. 480, pp. 143–149, May 2018.

[365] H. Kantarjian et al., "Decitabine improves patient outcomes in myelodysplastic syndromes," Cancer, vol. 106, no. 8, pp. 1794–1803, Apr. 2006.

[366] P. Fenaux et al., "Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study.," Lancet. Oncol., vol. 10, no. 3, pp. 223–32, Mar. 2009.

[367] J. F. Linnekamp, R. Butter, R. Spijker, J. P. Medema, and H. W. M. van Laarhoven, "Clinical and biological effects of demethylating agents on solid tumours – A systematic review," Cancer Treat. Rev., vol. 54, pp. 10–23, Mar. 2017.

[368] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang, "Genome engineering using the CRISPR-Cas9 system," Nat. Protoc., vol. 8, no. 11, pp. 2281–2308, Oct. 2013.

[369] X. Xu et al., "A CRISPR-based approach for targeted DNA demethylation," Cell Discov., vol. 2, no. 1, p. 16009, Dec. 2016.

[370] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2013.," CA. Cancer J. Clin., vol. 63, no. 1, pp. 11–30, Jan. 2013.

[371] M. P. Tabor, R. H. Brakenhoff, H. J. Ruijter-Schippers, J. A. Kummer, C. R. Leemans, and B. J. M. Braakhuis, "Genetically altered fields as origin of locally recurrent head and neck cancer: a retrospective study.," Clin. Cancer Res., vol. 10, no. 11, pp. 3607–13, Jun. 2004.

[372] C. R. Leemans, B. J. M. Braakhuis, and R. H. Brakenhoff, "The molecular biology of head and neck cancer," Nat. Rev. Cancer, vol. 11, no. 1, pp. 9–22, Jan. 2011.

[373] A. G. Zygogianni et al., "Oral squamous cell cancer: early

detection and the role of alcohol and smoking," Head Neck Oncol., vol. 3, p. 2, Jan. 2011.

[374] D. H. Maasland, P. A. van den Brandt, B. Kremer, R. A. Goldbohm, and L. J. Schouten, "Alcohol consumption, cigarette smoking and the risk of subtypes of head-neck cancer: results from the Netherlands Cohort Study," BMC Cancer, vol. 14, p. 187, Mar. 2014.

[375] D. Hanahan and R. A. A. Weinberg, "Hallmarks of cancer: the next generation," Cell, vol. 144, no. 5, pp. 646–674, Mar. 2011.

[376] R. de Bree, L. van der Putten, J. Brouwer, J. A. Castelijns, O. S. Hoekstra, and C. R. Leemans, "Detection of locoregional recurrent head and neck cancer after (chemo)radiotherapy using modern imaging.," Oral Oncol., vol. 45, no. 4–5, pp. 386–93, Apr. 2009.

[377] G. Sciandrello, F. Caradonna, M. Mauro, and G. Barbata, "Arsenic-induced DNA hypomethylation affects chromosomal instability in mammalian cells," Carcinogenesis, vol. 25, no. 3, pp. 413–417, Mar. 2004.

[378] A. R. Karpf and S. Matsui, "Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells," Cancer Res., vol. 65, no. 19, pp. 8635–8639, Oct. 2005.

[379] D. Lissa and A. I. Robles, "Methylation analyses in liquid biopsy.," Transl. lung cancer Res., vol. 5, no. 5, pp. 492–504, Oct. 2016.

[380] S. P. Nunes et al., "Cell-Free DNA Methylation of Selected Genes Allows for Early Detection of the Major Cancers in Women.," Cancers (Basel)., vol. 10, no. 10, p. 357, Sep. 2018.

[381] R. Xu et al., "Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma," Nat. Mater., vol. 16, no. 11, pp. 1155–1161, Nov. 2017.

[382] Biobix, Lab of Bioinformatics and Computational Genomics, University of Ghent, Ghent, Belgium, "Map of the Human Methylome." http://www.biobix.be/map-of-the-human-methylome/. 2014.

[383] M. Peng, C. Chen, A. Hulbert, M. V Brock, and F. Yu, "Non-blood circulating tumor DNA detection in cancer.," Oncotarget, vol. 8, no. 40, pp. 69162–69173, Sep. 2017.

[384] J. D. Merker et al., "Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review," J. Clin. Oncol., vol. 36, no. 16, pp. 1631–1641, Jun. 2018.

[385] J. O. Boyle et al., "Gene mutations in saliva as molecular markers for head and neck squamous cell carcinomas," Am. J. Surg., vol. 168, no. 5, pp. 429–432, Nov. 1994.

[386] M. Jung and G. P. Pfeifer, "Aging and DNA methylation," BMC Biol., vol. 13, no. 1, p. 7, Dec. 2015.

[387] C. Qu et al., "Caveolin-1 facilitated KCNA5 expression, promoting breast cancer viability.," Oncol. Lett., vol. 16, no. 4, pp. 4829–4838, Oct. 2018.

[388] K. E. Ryland et al., "Promoter Methylation Analysis Reveals That KCNA5 Ion Channel Silencing Supports Ewing Sarcoma Cell Proliferation.," Mol. Cancer Res., vol. 14, no. 1, pp. 26–34, Jan. 2016.

[389] Y. Wada et al., "Mitogen-inducible SIPA1 is mapped to the conserved syntenic groups of chromosome 19 in mouse and chromosome 11q13.3 centromeric to BCL1 in human.," Genomics, vol. 39, no. 1, pp. 66–73, Jan. 1997.

[390] N. Minato and M. Hattori, "Spa-1 (Sipa1) and Rap signaling in leukemia and cancer metastasis.," Cancer Sci., vol. 100, no. 1, pp. 17–23, Jan. 2009.

[391] T. Takahara et al., "SIPA1 promotes invasion and migration in human oral squamous cell carcinoma by ITGB1 and MMP7.," Exp. Cell Res., vol. 352, no. 2, pp. 357–363, Mar. 2017.

[392] Y. Wang et al., "The meiotic TERB1-TERB2-MAJIN complex tethers telomeres to the nuclear envelope." Nat. Commun., vol. 10, no. 1, p. 564, Dec. 2019.

[393] H. Shibuya, A. Hernández-Hernández, A. Morimoto, L. Negishi, C. Höög, and Y. Watanabe, "MAJIN Links Telomeric DNA to the Nuclear Membrane by Exchanging Telomere Cap," Cell, vol. 163, no. 5, pp. 1252–1266, Nov. 2015.

[394] R. Kuiper et al., "A gene expression signature for high-risk multiple myeloma.," Leukemia, vol. 26, no. 11, pp. 2406–13, Nov. 2012.

[395] S. L. Gerson, "MGMT: its role in cancer aetiology and cancer therapeutics," Nat. Rev. Cancer, vol. 4, no. 4, pp. 296–307, Apr. 2004.

[396] K. Shridhar et al., "DNA methylation markers for oral pre-cancer progression: A critical review," Oral Oncol., vol. 53, pp. 1–9, Feb. 2016.

[397] R. Noorlag et al., "Promoter hypermethylation using 24-gene array in early head and neck cancer: Better outcome in oral than in oropharyngeal cancer," Epigenetics, vol. 9, no. 9, Jul. 2014.

[398] F. K. J. Leusink et al., "Novel diagnostic modalities for assessment of the clinically node-negative neck in oral squamous-cell carcinoma.," Lancet. Oncol., vol. 13, no. 12, pp. e554-61, Dec. 2012.

[399] I. M. Smith, W. K. Mydlarz, S. K. Mithani, and J. A. Califano, "DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage.," Int. J. cancer, vol. 121, no. 8, pp. 1724–8, Oct. 2007.

[400] C. T. Tulisiak, R. A. Harris, and I. Ponomarev, "DNA modifications in models of alcohol use disorders.," Alcohol, vol. 60, pp. 19–30, May 2017.

[401] M. Varela-Rey, A. Woodhoo, M. L. Martinez-Chantar, J. M. Mato, and S. C. Lu, "Alcohol, DNA methylation, and cancer," Alcohol Res., vol. 35, no. 1, pp. 25–35, 2013.

[402] M. Hasegawa, H. H. Nelson, E. Peters, E. Ringstrom, M. Posner, and K. T. Kelsey, "Patterns of gene promoter methylation in squamous cell cancer of the head and neck," Oncogene, vol. 21, no. 27, pp. 4231–4236, Jun. 2002.

[403] K. W. K. Lee and Z. Pausova, "Cigarette smoking and DNA methylation," Front. Genet., vol. 4, p. 132, Jul. 2013.

[404] L. G. Tsaprouni et al., "Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation," Epigenetics, vol. 9, no. 10, pp. 1382–1396, 2014.

[405] H. Liu, Y. Zhou, S. E. Boggs, S. A. Belinsky, and J. Liu, "Cigarette smoke induces demethylation of prometastatic oncogene synuclein-gamma in lung cancer cells by downregulation of DNMT3B.," Oncogene, vol. 26, no. 40, pp. 5900–10, Aug. 2007.

[406] K. L. Ostrow et al., "Cigarette smoke induces methylation of the tumor suppressor gene NISCH," Epigenetics, vol. 8, no. 4, pp. 383–388, Apr. 2013.

[407] Y. Huang et al., "Cigarette smoke induces promoter methylation of single-stranded DNA-binding protein 2 in human esophageal squamous cell carcinoma," Int. J. cancer. Journal Int. du cancer, vol. 128, no. 10, pp. 2261–2273, May 2011.

[408] A. Omuro and L. M. DeAngelis, "Glioblastoma and Other Malignant Gliomas," JAMA, vol. 310, no. 17, p. 1842, Nov. 2013.

[409] R. Edgar, M. Domrachev, and A. E. Lash, "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository," Nucleic Acids Res., vol. 30, no. 1, pp. 207–210, Jan. 2002.

[410] M. Jeong et al., "Large conserved domains of low DNA methylation maintained by Dnmt3a.," Nat. Genet., vol. 46, no. 1, pp. 17–23, Jan. 2014.

[411] E. Calo and J. Wysocka, "Modification of Enhancer Chromatin: What, How, and Why?," Mol. Cell, vol. 49, no. 5, pp. 825–837, Mar. 2013.

[412] S. Moran, C. Arribas, and M. Esteller, "Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences.," Epigenomics, vol. 8, no. 3, pp. 389–99, Mar. 2016.

[413] L. Siggens and K. Ekwall, "Epigenetics, chromatin and genome organization: recent advances from the ENCODE project," J. Intern. Med., vol. 276, no. 3, pp. 201–214, Sep. 2014.

[414] ENCODE Project Consortium, "An integrated encyclopedia of DNA elements in the human genome," Nature, vol. 489, no. 7414, pp. 57–74, Sep. 2012.

[415] M. Lizio et al., "Gateways to the FANTOM5 promoter level mammalian expression atlas," Genome Biol., vol. 16, no. 1, p. 22, Jan. 2015.

[416] R. M. Kohli and Y. Zhang, "TET enzymes, TDG and the dynamics of DNA demethylation.," Nature, vol. 502, no. 7472, pp. 472–9, Oct. 2013.

[417] W. A. Pastor, L. Aravind, and A. Rao, "TETonic shift: biological roles of TET proteins in DNA demethylation and transcription.," Nat. Rev. Mol. Cell Biol., vol. 14, no. 6, pp. 341–56, Jun. 2013.

[418] J. An, A. Rao, and M. Ko, "TET family dioxygenases and DNA demethylation in stem cells and cancers," Nat. Publ. Gr., p. 323, 2017.

[419] Y. Huang and A. Rao, "Connections between TET proteins and aberrant DNA modification in cancer.," Trends Genet., vol. 30, no. 10, pp. 464–74, Oct. 2014.

[420] K. D. Rasmussen and K. Helin, "Role of TET enzymes in DNA methylation, development, and cancer.," Genes Dev., vol. 30, no. 7, pp. 733–50, Apr. 2016.

[421] Y. Huang, W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu, and A. Rao, "The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing," PLoS One, vol. 5, no. 1, p. e8888, Jan. 2010.

[422] S.-G. Jin, S. Kadam, and G. P. Pfeifer, "Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine," Nucleic Acids Res., vol. 38, no. 11, pp. e125–e125, Jun. 2010.

[423] I. Guilleret and J. Benhattar, "Unusual distribution of DNA methylation within the hTERT CpG island in tissues and cell lines," Biochem. Biophys. Res. Commun., vol. 325, no. 3, pp. 1037–1043, Dec. 2004.

[424] I. Guilleret and J. Benhattar, "Demethylation of the human telomerase catalytic subunit (hTERT) gene promoter reduced hTERT expression and telomerase activity and shortened telomeres.," Exp. Cell Res., vol. 289, no. 2, pp. 326–34, Oct. 2003.

[425] F. Court et al., "The PEG13-DMR and brain-specific enhancers dictate imprinted expression within the 8q24 intellectual disability risk locus.," Epigenetics Chromatin, vol. 7, no. 1, p. 5, Mar. 2014.

[426] S. Kurukuti et al., "CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2.," Proc. Natl. Acad. Sci. U. S. A., vol. 103, no. 28, pp. 10684–9, Jul. 2006.

[427] D. Aran, S. Sabato, and A. Hellman, "DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes.," Genome Biol., vol. 14, no. 3, p. R21, Mar. 2013.

[428] A. Marusyk and K. Polyak, "Tumor heterogeneity: causes and consequences.," Biochim. Biophys. Acta, vol. 1805, no. 1, pp. 105–17, Jan. 2010.

[429] B. Chaudhury et al., "Heterogeneity in intratumoral regions with rapid gadolinium washout correlates with estrogen receptor status and nodal metastasis," J. Magn. Reson. Imaging, vol. 42, no. 5, pp. 1421–1430, Nov. 2015.

[430] N. Navin et al., "Inferring tumor progression from genomic heterogeneity.," Genome Res., vol. 20, no. 1, pp. 68–80, Jan. 2010.

[431] D. F. Quail and J. A. Joyce, "Microenvironmental regulation of tumor progression and metastasis.," Nat. Med., vol. 19, no. 11, pp. 1423–37, Nov. 2013.

[432] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge.," Contemp. Oncol. (Poznan, Poland), vol. 19, no. 1A, pp. A68-77, 2015.

[433] T. J. Hudson et al., "International network of cancer genome projects," Nature, vol. 464, no. 7291, pp. 993–998, Apr. 2010.

[434] T. Barrett et al., "NCBI GEO: archive for functional genomics data sets--update," Nucleic Acids Res., vol. 41, no. Database issue, pp. D991-5, Jan. 2013.

[435] M. S. Cline et al., "Exploring TCGA Pan-Cancer Data at the UCSC Cancer Genomics Browser," Sci. Rep., vol. 3, no. 1, p. 2652, Dec. 2013.

[436] F. M. Davis, T. A. Stewart, E. W. Thompson, and G. R. Monteith, "Targeting EMT in cancer: opportunities for pharmacological intervention," Trends Pharmacol. Sci., vol. 35, no. 9, pp. 479–488, Sep. 2014.

[437] N. Prevarskaya, R. Skryma, and Y. Shuba, "Calcium in tumour metastasis: new roles for known actors," Nat. Rev., vol. 11, no. 8, pp. 609–618, Jul. 2011.

[438] G. R. Monteith, D. McAndrew, H. M. Faddy, and S. J. Roberts-Thomson, "Calcium and cancer: targeting Ca2+ transport," Nat. Rev. Cancer, vol. 7, no. 7, pp. 519–530, Jul. 2007.

[439] C. C. Chen, K. Y. Wang, and C. K. J. Shen, "DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases.," J. Biol. Chem., vol. 288, no. 13, pp. 9084–91, Mar. 2013.

[440] J. N. Anastas and R. T. Moon, "WNT signalling pathways as therapeutic targets in cancer," Nat. Rev. Cancer, vol. 13, no. 1, pp. 11–26, Jan. 2013.

[441] T. Zhan, N. Rindtorff, and M. Boutros, "Wnt signaling in cancer," Oncogene, vol. 36, no. 11, pp. 1461–1473, Mar. 2017.

[442] A. Vojta et al., "Repurposing the CRISPR-Cas9 system for targeted DNA methylation," Nucleic Acids Res., vol. 44, no. 12, pp. 5615–5628, Jul. 2016.

[443] T. C. G. A. Network, "Comprehensive genomic characterization of head and neck squamous cell carcinomas," Nature, vol. 517, no. 7536, pp. 576–582, Jan. 2015.

[444] A. Benard et al., "A Novel Multiplex NGS-Based Digital MLPA Assay for Copy Number Detection of 700 Target Sequences in Childhood Acute Lymphoblastic Leukemia," Blood, vol. 128, no. 22, 2016.

ENGLISH SCIENTIFIC SUMMARY

Head and neck cancer is the collective name for cancers that are located in the head-neck (HN) region. Most tumors arise from the mucosa in the upper aerodigestive tract, of which the oral cavity, pharynx and larynx are the most frequent origins. The most common type of these cancers arises from epithelial cells and become squamous cell carcinomas (HNSCC). Important risk factors for HNSCC are tobacco use and alcohol consumption, especially when combined. The frequent exposure of epithelial cells to these factors causes accumulation of (epi)genetic aberrations.

A major subgroup of HNSCC are those tumors located in the oral cavity. In 2012 an estimated 300.400 new cases of cancer of the oral cavity, which consisted of predominantly oral squamous cell carcinomas (OSCC) were diagnosed worldwide. In the same period 145.400 deaths were associated with these same cancers. Despite improvement in diagnosis and therapy, the incidence of these oral cancers has been steadily increasing over the last few years making these tumors an increasing problem.

A dilemma in the treatment of OSCC is the approach of the "clinical negative neck". This term refers to a common problem where there is no evidence for metastases from the OSCC to the lymph nodes in the neck based on clinical assessment and imaging. On the other hand, the behavior of the tumor indicated a substantial risk for very small hard-to-detect (referred to as occult) metastases in the cervical lymph nodes. Moreover, these lymph node metastases in the neck are the most important cause of reduced survival rates for OSCC patients.

To better predict the risk of occult metastases to the neck, some tumors characteristics are used to predict lymph node metastasis such as tumor size, invasive behavior and the pattern of invasion. The best predictors are infiltration depth, perineural and lymphovascular invasion as well as histological differentiation in the pretreatment primary tumor biopsy tissue. Unfortunately, the diagnosis based on these tumor characteristics is not sufficient enough to accurately predict the presence of lymph node metastasis in the neck for all patients. As a result, many patients don't receive the proper treatment for these occult metastases. On the other hand, patients that are incorrectly diagnosed with lymph node metastases can get an unnecessary prophylactic treatment with morbidity due to surgery and radiotherapy.

Other tumor characteristics that predict the presence of lymph node metastases are molecular tumor biomarkers. These markers refer to the status of the tumor cell genetics. DNA aberrations can lead to changes in gene expression which can affect traits such as tumor size as well as biological behavior which can promote tumor cell migration and invasion eventually resulting in tumor metastasis in the lymph nodes. More traditionally diagnosed genetic aberrations include changes in normal gene copy numbers or mutations altering the DNA sequence. While many such pathologic genetic changes influence the process of cancer metastasis, limited understanding of the biological impact of these changes as well as the high costs of biomarker diagnoses have prevented these molecular markers from becoming the silver bullet that clinicians require to tackle the clinical negative neck dilemma.

A relatively recently discovered mechanism of gene expression regulation is epigenetics. Epigenetics means "above genetics" and is a collective term for modifications of the DNA in other areas than the DNA sequence. One such change is DNA methylation which refers to the addition of a methyl-group to nucleotides in the DNA. In general, accumulation of these methyl groups to the nucleotide cytosine in gene promoters causes lowering of the associated genes' expression.

In cancer, DNA methylation of the promoter region of tumor suppressor genes is increased, resulting in the inhibition of their expression. Additionally, DNA stretches between genes is often unmethylated during tumor progression which results in a more unstable DNA structure causing even more pathologic genetic changes. Interestingly, these changes in DNA methylation that promote cancer occur earlier during tumor development and are more common than mutations. Therefore, the inclusion of DNA methylation detection in the primary tumor of specific metastasis-related genes could provide tools for the clinic to better predict the presence of lymph node metastases in the neck of OSCC patients.

In **chapter 2** we used methylation specific PCR (MSP) to validate a selection of previously reported DNA methylation markers associated with N-status in HNSCC. A total of 28 DNA methylation markers were selected and tested on a cohort of 70 early-stage OSCC of a well-established database of patients treated in the University Medical Center Groningen between 1997-2008. For this purpose, we compared the methylation status of these genes by MSP of OSCC patients without lymph node metastases to OSCC patients with histological proven lymph node metastases. Five out of 28 methylation markers (OCLN, CDKN2A, MGMT, MLH1 and DAPK1) were frequently differentially methylated in OSCC. Of these, MGMT hypermethylation was associated with pN0 status (p = 0.02) and with lower levels of MGMT protein as detected by immunohistochemistry (p = 0.02). DAPK1 methylation was associated with pN+ status (p = 0.008) but did not associate with DAPK1 protein expression. In conclusion, out of 28 reported metastasis-related methylation markers, two genes, DAPK1 and MGMT (7%), showed a predictive value for the pN status for nodal metastasis in a clinical group of OSCC. Therefore, DNA methylation markers are capable of contributing to diagnosis and treatment selection in OSCC.

Unfortunately, the validation of the 28 gene panel selected from literature did not yield DNA methylation biomarkers with sufficient predictive value for N-status in OSCC. Therefore, in **chapter 3** we performed a genome wide methylation assessment using OSCC patients with and without lymph node metastases. Sequences that were differentially methylated between pN0 and pN+ OSCC were selected based on the likelihood of differential methylation and validated using an independent OSCC cohort. Additionally, OSCC from the publicly available database The Cancer Genome Atlas (TCGA) were used for additional validation. MethylCap-Seq analysis revealed 268 differentially methylated markers. WISP1 was the highest-ranking annotated gene that showed hypomethylation in the N+ group. Expression of WISP1 using immunohistochemistry was analyzed on a large OSCC cohort (n=204). Bisulfite pyrosequencing confirmed significant hypomethylation within the WISP1 promoter region in N+ OSCC (p = 0.03) and showed an association between WISP1 hypomethylation and high WISP1 expression (p = 0.01). Both these results were confirmed using 148 OSCC retrieved from the TCGA database. In a large OSCC cohort high WISP1 expression was associated with lymph node metastasis (p = 0.05), decreased disease-specific survival (p = 0.022) and decreased regional disease-free survival (p = 0.027). These data suggested that

WISP1 expression is regulated by DNA methylation and that WISP1 hypomethylation contributes to LN metastasis in OSCC. WISP1 protein and WISP1 DNA methylation levels are potential biomarkers for identifying OSCC patients who require neck dissection treatment.

While WISP1 methylation status is predictive for N-status in our OSCC cohort, the clinical application of WISP1 methylation assessment is limited because WISP1 is hypomethylated in pN+ OSCC. Hypermethylation is much easier to measure than hypomethylation due to the background of normal methylation. In addition, possible treatment options with currently available demethylation drugs, should preferably be used for markers that are hypermethylated in OSCC with lymph node metastasis. Therefore, in **Chapter 4** we re-analyzed the MethylCap-Seq genome-wide methylation data combined with a reported gene panel predictive for LN-status in OSCC to focus on the selection of hypermethylated biomarkers predictive for pN+ OSCC. A two-step statistical selection of differentially methylated sequences revealed 14 genes with increased methylation status and mRNA down-regulation in pN+ OOSCC. RAB25, a known tumor suppressor gene, was the highest-ranking gene in the discovery set. In the validation sets, both RAB25 mRNA (P = 0.015) and protein levels (P = 0.012) were lower in pN+ OOSCC. RAB25 mRNA levels were negatively correlated with RAB25 methylation levels (P < 0.001) but RAB25 protein expression was not. Our data revealed that promoter methylation is a mechanism resulting in down-regulation of RAB25 expression in pN+ OOSCC and decreased expression is associated with lymph node metastasis. RAB25 methylation detection might contribute to lymph node metastasis diagnosis and serve as a potential new therapeutic target in OOSCC.

While the revised algorithm and data analysis did lead to the identification of a tumor-suppressor gene down-regulated in pN+ OSCC by hypermethylation, the ideal biomarker would be significantly hypermethylated in pN+ OSCC compared to pN0 OSCC. This would allow for a cheap and fast clinical detection method to define patients with LN metastases and to select those patients for whom potential treatment options would be available for this particular DNA methylation biomarker. Because RAB25 was not significantly hypermethylated in pN+ OSCC, the statistical selection to identify additional methylation markers was further expanded in **Chapter 5**. A multistep selection algorithm was performed using our OSCC-specific Methylome database, a gene expression signature predictive for pN+ OSCC and The Cancer Genome Atlas (TCGA) data to identify epigenetically down-regulated genes predictive for LN metastasis in OSCC. The gene with the most supportive evidence was characterized by immunohistochemistry and methylation-specific PCR using a cohort of OSSC and HNSCC cell lines. From a list of 26 previously identified markers, S100A9 was identified as the most promising biomarker for LN metastases. TCGA data showed that S100A9 methylation was negatively correlated with S100A9 expression and significantly associated with the presence of LN metastasis. In an independent OSCC cohort reduced S100A9 expression was significantly correlated with LN metastasis as well as with decreased patient survival. In HNSCC cell lines, treatment with demethylating drugs resulted in significant demethylation of the promoter and concomitant upregulation of S100A9 expression. This study showed that epigenetic down-regulation of S100A9 contributes to LN metastasis in OSCC providing a new tumor biomarker and a potential therapeutic target for the detection and treatment of OSCC patients with LN metastases.

In addition to the "clinical negative neck" dilemma, OSCC disease progress is complicated by a high frequency of both local recurrence and/or second primary tumor of 20-30%. This relatively high rate is partly caused by residual tumor cells of the first primary tumor and the presence of precancerous epithelium that has not clinically manifested but does have many genetic aberrations caused by the dietary risk factors of OSCC such as tobacco smoking and alcohol consumption. Since OSCC cells are shed into the oral cavity, the detection of tumor-specific DNA methylation markers in saliva could be a tool for the early detection of local recurrences of OSCC to increase the chance of early detection as well as making the follow-up easier for OSCC patients. In **chapter 6** we used the genome-wide data of the six pN0 and six pN+ OSCC acquired by MethylCap-Seq combined to identify genes hypermethylated in OSCC but not in normal oral epithelial cells and validate new methylation markers to detect OSCC cells in saliva. Potential OSCC-specific hypermethylation markers were validated on saliva from ten OSCC patients and five younger and five age-matched healthy controls using quantitative methylation specific PCR (QMSP). These new methylation markers were compared to markers reported to be methylated in saliva in the literature (EDNRB, HOXA9, NID2 and TIMP3). Using our OSCC methylome, seven genomic locations representing six genes (C11orf85, CMTM2, FERMT3, KCNA5, SIPA1 and TBX4) were identified that were significantly hypermethylated in tissues of OSCC compared to DNA from controls. QMSP analysis showed significant hypermethylation of KCNA5 in saliva of OSCC patients compared to saliva of age-matched controls ($p < 0.003$). Moreover, when combining QMSP results of KCNA5 with TIMP3, a 100% accuracy in detecting saliva from OSCC patients compared to non-cancer controls was observed. This study identified several new OSCC-specific methylation markers with a high sensitivity and high negative predictive value for the detection of OSCC. Two methylation (KCNA5 and TIMP3) markers might be useful for early detection of OSCC local regional recurrence in saliva cells. A larger prospective study should be done to confirm the clinical relevance of these two markers.

NEDERLANDSE SAMENVATTING

Hoofd- en hals tumoren is de verzamelnaam voor bepaalde gezwellen in het hoofd-halsgebied. Deze tumoren kunnen onder andere ontstaan uit de neus, neusbijholte, neuskeelholte, mondholte, mondkeelholte, keelholte, strottenhoofd, speekselklier en aan de lippen. De meeste van deze hoofd- en hals tumoren ontstaan in de bovenste adem- en voedingsweg en dan met name in de mondholte, de neus- keelholte en het strottenhoofd. Deze tumoren komen vooral voort uit de bekleding van deze organen, wat ook wel het epitheel wordt genoemd. Tumoren die uit het epitheel ontgroeien, zijn vrijwel altijd van het subtype plaveiselcelcarcinoom (HNSCC genoemd), wat betekent dat ze voortkomen uit een bepaald type cel waar de opperhuid uit bestaat: de keratinocyten.

De oorzaak van dit type tumoren komt voornamelijk door blootstelling aan tabak en alcohol. Regelmatige blootstelling van het epitheel aan deze risicofactoren zorgt voor veranderingen in het erfelijke materiaal van deze cellen: DNA-schade genoemd.

De grootste subgroep van hoofd-hals plaveiselceltumoren zijn die tumoren die ontstaan in de mondholte (OSCC genoemd). In 2012 zijn er naar schatting wereldwijd 300.400 nieuwe gevallen van OSCC gediagnosticeerd. Daarnaast zijn er in hetzelfde jaar 145.400 OSCC-patiënten overleden aan de gevolgen van deze tumoren. Bovendien is de incidentie van OSCC na 2012 nog verder toegenomen ondanks verbetering van de diagnose en therapie.

OSCC zaaien altijd eerst uit naar de lymfklieren in de hals. Dit komt doordat er veel lymfebanen beginnen in de mondholte waar losgeraakte tumorcellen makkelijk in terecht kunnen komen. Deze uitzaaiingen (metastasen genoemd) verlagen de overlevingskansen van de patiënten aanzienlijk. De overlevingskans van deze patiënten halveert indien er lymfkliermetastasen aanwezig zijn. Het is daarom van groot belang dat deze metastasen zo snel en accuraat mogelijk worden ontdekt tijdens het in kaart brengen van de ziekte voorafgaande aan de behandeling. Voor de detectie van deze metastasen worden diagnostische middelen zoals beeldvorming technieken waaronder Echografie, MRI en CT-scans gebruikt. Met deze technieken wordt echter slechts bij 60 tot 70 procent van de OSCC -patiënten met lymfekliermetastasen, deze uitzaaiingen ook daadwerkelijk gevonden. Dat er bij de overige 30 tot 40 procent van de OSCC-patiënten metastasen worden gemist komt onder andere doordat de metastasen soms te klein zijn om met deze beeldvorming technieken te detecteren, deze worden occulte metastasen genoemd. Indien er een risico is dat er metastasen aanwezig kunnen zijn maar dat ze niet detecteerbaar zijn, worden de lymfklieren in de hals preventief behandeld door middel van chirurgie of bestraling. Een deel van de OSCC-patiënten zal onnodig worden behandeld en daardoor bijwerkingen van deze procedures ervaren. Daarnaast zullen er patiënten zijn, waar het risico op occulte metastasen als klein wordt ingeschat. Vervolgens zullen deze patiënten geen preventieve behandeling zullen ontvangen, terwijl zij wel een uitzaaiing in de hals kunnen ontwikkelen. Het is daarom van groot belang dat de detectie van metastasen in de lymfklieren wordt verbeterd. OSCC-patiënten met lymfkliermetastasen worden als pN+ geclassificeerd terwijl OSCC zonder lymfkliermetastasen als pN0 worden geclassificeerd.

Als aanvulling op de beeldvormingstechnieken voor de detectie van lymfkliermetastasen in de hals wordt

onderzoek gedaan naar biologische en moleculaire eigenschappen van de primaire OSCC-tumoren die daarin een voorspellende waarde kunnen hebben. Onder andere de grootte van een tumor, de mate waarin een tumor het omliggend gezonde weefsel binnen groeit maar ook de aanwezigheid van bepaalde genetische of biologische veranderingen kunnen gebruikt worden om het klinisch gedrag van de tumor op te helderen en te voorspellen. Dit soort tumoreigenschappen worden biomarkers genoemd. Het gebruik van dergelijke tumor biomarkers bij patiënten met OSCC, zoals de infiltratiediepte van de tumor hebben geleid tot een verbetering van het detectiepercentage van OSCC-lymfkliermetastasen.

Een relatief recentelijk ontdekte vorm van tumor biomarkers zijn epigenetische biomarkers. In de essentie bestaat het DNA uit vier verschillende bouwblokken, oftewel nucleotiden. Een lange aaneenschakeling van deze moleculen dient als een soort blauwdruk voor de opbouw van cellen, weefsels, organen en uiteindelijk het gehele menselijk lichaam. Een specifiek stuk van het DNA dat codeert voor een functioneel molecuul, een eiwit, heet een gen. In het verleden werd bij het analyseren van het DNA alleen gekeken naar veranderingen, toevoegingen of verwijderingen van de nucleotiden in het DNA. Door deze wijzigingen in genen verandert de structuur van het functionele eiwit dat vaak samengaat met een andere functie van het betreffende eiwit. Door de jaren heen is echter gebleken dat de manier waarop het DNA wordt verpakt in een cel ook invloed heeft op de manier waarop DNA zijn functie als blauwdruk kan uitoefenen. De lengte van het menselijke DNA is in totaal 2 meter lang terwijl bijvoorbeeld een huidcel ongeveer 17 keer kleiner is dan een zoutkorrel (300 micrometer). Om ervoor te zorgen dat al dit erfelijke materiaal in een minuscule cel past maar ook afgelezen kan worden, is een complex systeem van archiveren nodig. De mate waarin DNA verpakt wordt, kan gevolgen hebben voor welke informatie er beschikbaar is. Dit kan betekenen dat bepaalde essentiële blauwdrukken niet afgelezen kunnen worden wat zou kan leiden tot ontregeling van het normaal functioneren van een cel. De wijze van "verpakken" van DNA door epigenetische veranderingen kan daardoor ook bijdragen tot het ontstaan van kanker.

Een van de meest voorkomende epigenetische verpakkingsvormen is DNA methylatie. In het algemeen is dit de toevoeging van een methyl-groep aan de nucleotide cytosine. De methylatie van cytosine veroorzaakt in het algemeen strakkere vouwing van het DNA waardoor genen die veel DNA methylatie bevatten slecht afgelezen kunnen worden en er dus geen eiwit gevormd wordt. Tijdens het ontstaan van kanker neemt deze uitschakeling van genen door verhoogde DNA methylatie toe, dit wordt DNA hypermethylatie genoemd. Daarnaast neemt de DNA methylatie op plekken in het DNA waar zich geen genen bevinden, juist af waardoor het DNA instabieler kan worden, dit wordt DNA hypomethylatie genoemd. Het gebruik van DNA methylatie als biomarker in de kliniek voor het voorspellen van de aanwezigheid van lymfkliermetastasen in de hals, vergroot de kans op de juiste diagnoses van de OSCC-patiënt. Daarnaast zijn DNA methylatie-markers uitermate geschikt om te dienen als klinische test. Omdat DNA methylatie vroeg optreedt bij het ontstaan van tumoren en omdat het testen goedkoop en betrouwbaar is te meten. Daarnaast geeft DNA methylatie metingen informatie over zowel de toestand van het DNA als het biologisch gedrag van de tumor.

Het doel van dit proefschrift is het vinden, testen en valideren van DNA methylatie biomarkers waarmee we in de primaire tumor van OSCC-patiënten metastasen in de lymfklieren in de hals kunnen detecteren.

In **hoofdstuk 2** hebben wij 28 bekende genen geselecteerd uit de literatuur die beschreven zijn als mogelijke DNA methylatiemarkers geassocieerd met de aanwezigheid van lymfkliermetastasen. Vervolgens hebben wij deze 28 biomarkers getest op patiënten uit een groot patiëntenbestand dat bestaat uit alle patiënten met OSCC die tussen 1997 tot 2008 zijn behandeld in het UMCG. Van deze patiënten is alle informatie over het ziekteverloop bekend en is er erfelijk materiaal van de primaire OSCC beschikbaar. Hiermee kan een associatie onderzocht worden tussen de OSCC-progressie en de mate van DNA methylatie. In totaal zijn 70 OSCC-patiënten getest doormiddel van de techniek Methylatie-specifieke PCR. Vijf van de 28 biomarkers (OCLN, CDKN2A, MGMT, MLH1 en DAPK1) vertoonden regelmatig een toename van DNA methylatie. Hypermethylatie van het gen MGMT was statistisch significant geassocieerd met pN0 OSCC en met de afname van expressie van het MGMT-eiwit. Daarnaast bleek de hypermethylatie van DAPK1 geassocieerd met OSCC met lymfkliermetastasen maar niet met verlaging van de DAPK1 eiwit expressie. Van de 28 geselecteerde biomarkers bleken er dus slechts twee geassocieerd met de aanwezigheid van lymfkliermetastasen in OSCC.

Hoewel de analyse van de methylatie status van MGTM en DAPK1 kunnen bijdragen aan de diagnostiek voor de detectie van OSCC-lymfkliermetastasen, is de voorspellende waarde van deze twee genen onvoldoende om alle occulte lymfkliermetastasen bij OSCC-patiënten te detecteren. Op zoek naar nieuwe en betere methylatiemarkers hebben we in **hoofdstuk 3** gekozen voor een techniek die in staat is om de methylatie status van het gehele menselijk DNA in één keer te analyseren. Hiervoor hebben we de techniek MethylCap-Seq toegepast op DNA van patiënten met OSCC met (pN+) en zonder (pN0) lymfkliermetastasen in de hals. Daarnaast hebben we gebruik gemaakt van DNA methylatie data van OSCC-patiënten afkomstig uit de publiekelijk toegankelijke database, The Cancer Genome Atlas (TCGA). TCGA-data van in totaal 148 OSCC-patiënten is gebruikt als eerste validatiestap van de MethylCap-Seq resultaten. Met de MethylCap-Seq hebben we doormiddel van statistische analyse 286 genen geïdentificeerd waarvan de DNA methylatie status significant verschillend was tussen pN0 OSCC en pN+ OSCC. Hiervan bleek dat het verlies van methylatie in pN+ OSCC ten opzichte van pN0 OSCC in het WISP1 gen het grootste verschil tussen pN0 OSCC en pN+ OSCC maakte. Doormiddel van kwantificatie van aankleuring van het WISP1 eiwit in tumorweefsel is vastgesteld dat de expressie van het WISP1 eiwit significant toegenomen is in pN+ OSCC-patiënten. En er bleek een verband tussen de hogere expressie van het WISP1 eiwit en de afname van methylatie in het WISP1 gen te zijn. De hogere expressie van het WISP1 eiwit bleek de overlevingskans van OSCC-patiënten significant negatief te beïnvloeden. Deze studie toonde aan dat zowel de methylatie status van het WISP1 gen als de expressie van het WISP1 eiwit gebruikt kunnen worden in de kliniek voor het voorspellen van de aanwezigheid van lymfkliermetastasen.

Hoewel theoretisch WISP1 goed kan dienen als tumor biomarker om de aanwezigheid of afwezigheid van de lymfkliermetastasen in OSCC-patiënten, is deze biomarker technisch gezien niet ideaal. Verlies van methylatie is lastiger en minder betrouwbaar te detecteren dan toename van methylatie. En hoewel te hoge niveaus van het WISP1 theoretisch goed behandelbaar zou zijn, bestaan er geen breed beschikbaar specifieke WISP1 medicijnen. Daarnaast zijn is er wel medicatie beschikbaar die hypermethylatie van genen kan behandelen. Daarom hebben we in **hoofdstuk 4** onze MethylCap-Seq en TCGA-data opnieuw geanalyseerd en deze analyse meer gericht op het selecteren van hypermethylatie in pN+ OSCC. Dit

leidde tot de identificatie van het gen RAB25. Validatie van deze biomarker leidde tot de conclusies dat RAB25 eiwit expressie significant is afgenomen in pN+ OSCC en dat er een correlatie is tussen RAB25 gen hypermethylatie en afname in expressie van het RAB25 eiwit. Daarnaast bleek dat de afname van expressie van het RAB25 eiwit een significante correlatie vertoonden met OSCC met lymfkliermetastasen. Echter, in onze patiëntengroep was het RAB25 gen niet meer significant gehypermethyleerd in pN+ OSCC ten opzichte van pN0 OSCC. Concluderend kunnen we vaststellen dat de expressie van RAB25 gebruikt kan worden om de aanwezigheid van lymfkliermetastasen in OSCC-patiënten te voorspellen.

Uiteindelijk doel van dit proefschrift is het identificeren van tumor biomarkers waarmee we de pN0 en pN+ OSCC kunnen onderscheiden op basis van DNA methylatie status, en daarmee het ziekteverloop en metastasering kunnen voorspellen. Hiervoor is het noodzakelijk dat de methylatie status van de gekozen biomarker significant verschillend is tussen de pN0 OSCC en pN+ OSCC-patiënten. RAB25 voldeed niet aan dit criterium. Daarom is in **hoofdstuk 5** het gebruikte algoritme verder uitgebreid door aanvullende analyses van de publieke TCGA-data toe te voegen voor de selectie van de methylatiemarkers met de beste voorspelling op aanwezigheid van lymfkliermetastasen. Dit leidde tot de selectie van het S100A9 gen. Validatie toonde aan dat S100A9 significant gehypermethyleerd is in pN+ OSCC en dat de expressie van het S100A9 eiwit significant lager is in pN+ OSCC, dat er een associatie is tussen S100A9 hypermethylatie en een lagere S100A9 eiwit expressie en dat een lagere S100A9 eiwit expressie de overlevingskansen van OSCC-patiënten beperkt. Daarnaast hebben we vastgesteld in een kankercel model dat bij behandeling met demethylerende middelen de hypermethylatie van het S100A9 gen verminderd kan worden en dat dit leidt tot toename van de expressie van het S100A9 eiwit. Dit suggereert dat hypermethylatie van S100A9 belangrijk is voor de regulatie van de expressie en daarmee het ontstaan van lymfkliermetastasen. Onafhankelijke studies zijn nodig om onze bevindingen te bevestigen, maar S1009 is dus een veelbelovende tumor biomarker voor toepassing in de kliniek om te bepalen welke OSCC-patiënten wel en welke patiënten geen lymfkliermetastasen hebben en daarom anders behandeld moeten worden. Onze bevindingen laten ook zien dat het biologisch proces waarin S100A9 expressie een rol speelt, zou kunnen dienen als target voor aanvullende behandelingen met specifieke medicatie.

Naast lymfkliermetastasen wordt de overlevingskans van OSCC ook geplaagd door een relatief hoge frequentie van het terugkeren van de eerste primaire tumor na behandeling, het lokaal recidief genoemd, of het ontstaan van een nieuwe tweede primaire OSCC-tumor. De hoge kans op deze recidieven en tweede primaire tumoren is gerelateerd aan de aard van de risicofactoren voor OSCC zoals alcohol- en tabak consumptie. Door veelvuldige blootstelling van het epitheel aan deze factoren ontstaat er een opstapeling van genetische afwijkingen in de epitheelcellen die de mondholte bekleden. Terwijl dit niet altijd direct leidt tot een tumor maakt dit het epitheel wel extra gevoelig voor het ontwikkelen van nieuwe OSCC of het terugkeren van oude OSCC. Daarom worden er bij een behandelde OSCC-patiënt regelmatig controles uitgevoerd om dergelijke tumoren tijdig te detecteren. Het is reeds gebleken dat het gemethyleerd DNA afkomstig van de OSCC-cellen uitgescheiden wordt in het speeksel. Dat maakt het mogelijk om de aanwezigheid van OSCC-cellen in de mondholte te detecteren door te zoeken naar gemethyleerd DNA specifiek voor OSCC in het speeksel. In **hoofdstuk 6** is daarom de MethylCap-Seq data opnieuw geanalyseerd met een ander algoritme om genen te identificeren die juist

gehypermethyleerd zijn in alle 12 OSCC-tumoren maar niet in gezonde cellen. Dit leidde tot de selectie van 6 mogelijke methylatiemarkers: C11orf85, CMTM2, FERMT3, KCNA5, SIPA1 en TBX4. Ter aanvullende controle van deze nieuwe biomarkers, zijn er vier DNA methylatiemarkers geselecteerd uit de literatuur: EDNRB, HOXA9, NID2 en TIMP3. Vervolgens is de voorspellende waarde van de methylatie status van deze genen voor de aanwezigheid van OSCC-cellen geëvalueerd op het speeksel van 10 patiënten met OSCC en van 10 gezonde individuen. Hieruit bleek dat twee markers (KCNA5 en TIMP3) geschikt zijn voor de detectie van OSCC-tumoren op DNA verkregen uit speeksel. Deze biomarkers zouden gebruikt kunnen worden om tijdens een controlebezoek eenvoudig in speeksel te testen of de OSCC als recidief weer gaat groeien. Onze studie is heel klein en de waarde voor de diagnostiek dient in grotere onafhankelijke studies bevestigd te worden. Omdat in de toekomst het speeksel door patiënten thuis verzameld en verzonden kan worden naar het lab, zou dit het controle proces sneller, goedkoper en gemakkelijker kunnen maken voor zowel de patiënten als de behandelaars.

DANKWOORD

Het is alweer even geleden dat ik ben begonnen met mijn promotie-traject. Het voelt als de dag van gisteren dat ik kwam voor mijn sollicitatie op de afdeling pathologie bij Prof. Dr. Philip Kluin en Dr. Ed Schuuring. Ik kwam gekleed in overhemd met nette broek en dat heeft mij blijkbaar geholpen om mezelf te onderscheiden van de andere biologen die ongetwijfeld met geitenwollen sokken in sandalen kwamen. Vier jaar onderzoek vlogen voorbij. Helaas verliep niet alles volgens plan, maar uiteindelijk is het toch gelukt. En ik ben niet alleen enorm blij en opgelucht dat dit proefschrift er eindelijk is, maar ik ben ook enorm trots op het eindresultaat. Mijn publicaties, de hoofdstukken en zowel de introductie als de algemene discussie geven mijn veel voldoening. Sterker nog, ik heb enorm genoten van de laatste loodjes. En daarom ben ik dan ook erg gelukkig dat ik deze periode zo heb kunnen afsluiten. Hoewel ik uiteraard mijn bloed, zweet en tranen in dit boekwerk heb gestort, had ik dit niet alleen kunnen doen. Ik wil daarom uit de grond van mijn hart een hele hoop mensen bedanken. Niet alleen voor hun hulp in het tot stand komen van dit proefschrift maar ook voor hun gezelligheid en steun tijdens deze reis vol wetenschap en persoonlijke groei.

Ik wil beginnen met mijn promotoren bedanken met als eerste Prof. Dr. E. Schuuring, oftewel Ed. Voor deze kans, het vertrouwen, geduld en de begeleiding gedurende de jaren die wij samen hebben gewerkt. Ik heb enorm veel van je geleerd. In mijn huidige werk pluk ik nog dagelijks de vruchten van alles wat je mij hebt bijgebracht. Met de name de drang om zoveel mogelijk de beste controles mee te nemen, steekt nog vaak de kop op. Toen ik op zoek ging naar een promotieplaats was het heel duidelijk voor mij dat ik translationeel onderzoek wilde doen, onderzoek dat dicht bij de kliniek staat en dat ernaar streeft om iets te betekenen voor de patiënt. Die kans kreeg ik bij uitstek bij jou op de afdeling pathologie en later ook bij de afdeling Mond-, Kaak- en aangezichtschirurgie. De samenwerkingen met Biobix van de Universiteit Gent, het CTMM-consortium en de verschillende klinische afdelingen van het UMCG hebben mij laten zien hoe ambitieus en breed onderzoek kan zijn. Veel grootser dan alleen werken in een kelder in het Biologisch centrum. Daarnaast heb ik mij altijd gesteund gevoeld in het najagen van mijn persoonlijke ontwikkelen door middel van bioinformatica cursussen en congressen. Ik heb enorme bewondering voor jouw werklust, jouw intelligentie en jouw beheersing van zoveel verschillende vakgebieden. En in het bijzonder hoe je je kennis weet toe te passen in de pathologie en moleculaire diagnostiek. En dat je dan ook nog tijd weet te vinden om zo'n familieman te zijn! Bedankt voor alles!

Als tweede wil ik uiteraard Prof. J.LN. Roodenburg bedanken. Hoewel ik niet direct ben begonnen bij de Mond-, Kaak- en aangezichtschirurgie, voelt het toch als of ik er altijd heb gewerkt. En dat heb ik aan u te danken. Ik heb mij altijd als bioloog thuis gevoeld bij de MKA. Ik ben dan ook enorm dankbaar voor uw vriendelijkheid en betrokkenheid. Ik heb mij geen moment de vreemde eend in de bijt gevoeld. Sterker nog, ik heb altijd het gevoel gehad dat de MKA erg enthousiast was over mijn onderzoek en dat ik zo echt wat kon bijdragen aan de kliniek. Dit heeft u ook versterkt door mij te betrekken bij uw tandheelkunde colleges en deelname aan de Nederlandse Vereniging van MKA-congressen. Ik heb enorme bewondering voor hoe u niet alleen in de kliniek heel veel heeft bijgedragen tot verbetering van de patiëntenzorg maar ook voor hoe u ook nog eens essentieel bent geweest voor de ontwikkeling van

de geneeskunde binnen de landmacht.

Daarnaast wil ik graag MKA-afdelingshoofd Prof. Dr. F.K.L. Spijkervet bedanken voor de unieke kans om als bioloog onderzoek te doen bij de MKA. Ik heb enorme bewondering voor hoe de MKA-afdeling zo professioneel en menselijk te gelijk kan zijn. Ik kijk enorm op naar alle kaakchirurgen en alle andere collega's van de MKA.

Uiteraard wil ik ook mijn leescomissie bedanken: Prof. Dr. M. Van Engeland, Prof. Dr. H. Hollema en Prof. Dr. A.J.W.P. Rosenberg voor hun tijd, moeite en expertise bij het beoordelen van dit proefschrift.

Ook ben ik veel dank verschuldigd aan alle wetenschappelijke samenwerkingen die ik tijdens mijn promotie aan heb mogen gaan. Te beginnen met onze samenwerking met Biobix van de Universiteit van Gent. Met in het bijzonder Prof. Dr. W. van Crieckinge en Prof. Dr. Tim de Meyer voor ondersteuning in de bioinformatica. Maar ook Dr. S. Denil en alle andere collega's van Biobix. Een hele hoop hele slimme mensen in het UMCG: Prof. Dr. B.F.A.M. Van der Laan van de KNO voor al uw kennis; Prof.Dr. H.J.M. Groen van de afdeling Longziekten voor alle translationele feedback; Prof. Dr. P.M. Kluin, voor het mogelijk maken van mijn promotieplaats en mijn begeleiding toen Ed nog geen hoogleraar was. Dr; G.B.A Wisman van de afdeling gynaecologie, voor alle feedback op mijn onderzoek, experimenten, teksten en alle expertise als het aankomt op DNA methylatie; Dr. M.J.H. Witjes van de MKA voor al het enthousiasme, alle ambitie, en al het klinische meedenken; Dr. T. Tomar, voor alle steun, vriendschap en dat we samen bioinformatica hebben mogen leren. Maar ook bij het UMCU: Dr. Stefan Willemsen, Dr. R. Noorlag en Dr. K. Koole, en alle andere collega's van de afdelingen Pathologie en MKA van het UMC Utrecht bedankt voor onze fijne samenwerking tijdens de FGF-studie. Ook wil ik het CTMM-consortium bedanken voor het mogelijk maken van mijn onderzoek en de vele inspirerende overleggen en meetings.

Behalve mijn promotoren, professoren en coauteurs ben ik aan nog een hele hoop andere collega's dank verschuldigd. Laat ik beginnen met Lorian en Mirjam. De DNA-diva's, de PCR-prinsessen, de Methylatie Masters! Waar Ed het brein is van de operatie, zijn jullie het hart en ziel van de groep. Zorgzaam en gezellig maar vooral enorm bekwaam in jullie werk. Altijd meedenkend en behulpzaam. Zonder jullie had ik nooit zoveel resultaten kunnen boeken. Want wat hebben jullie een werk verricht. Wat een hoop Immuunhistochemie, Grénman Assays en celkweek hebben jullie verricht sinds mij start bij de pathologie. Bedankt voor jullie hulp maar ook voor jullie vriendschap en de vele leuke herinneringen aan koffiepauzes, lab uitjes maar ook pop- en showbizz quizzen.

Mijn collega promovendi met wie ik tegelijk op de afdeling heb gewerkt: LJ, Koos, Emiel en Leonie. Lieuwe, wat heb ik enorme bewondering voor jouw professionaliteit, je kennis en werkhouding. Jij was de eerste kaakchirurg waar ik mee in aanraking kwam en ik was gelijk onder de indruk hoe je na je studie geneeskunde gewoon tegelijk promoveerde en tandheelkunde studeerde. Op magische wijze wist je ondanks de drukte geweldig onderzoek te doen en hoe je dat je helemaal eigen maakte. Koos, als opvolger van Lieuwe had je grote schoenen te vullen en dat is gewoon gelukt! Ook jij wist een even druk programma te baas te zijn en geweldig onderzoek te doen. Ik heb onze tijd op de kamer samen met Emiel altijd enorm leuk gevonden. Wat hebben we een lol gehad samen! Dat geldt ook voor Leonie.

Bedankt voor alle herinneringen, steun en ik wens jullie enorm veel succes in jullie verdere carrières en persoonlijke leven!!

En laat ik vooral de andere collega's niet vergeten: Wierd, Deborah, Gertrud, Sharon, Chris, Nancy, Bert, Rong, Frank, Aniek, Robin, Jasper, Brenda, Arja en alle andere collega's van het UMCG.

Daarnaast ben ik dank verschuldigd aan een aantal mensen die mij enorm veel geleerd hebben tijdens mijn bachelor en master en die hebben bijgedragen in mijn ontwikkeling als wetenschapper: Dr. Anghel Ostroveanu en Dr. Ingrid Nijholt. Tijdens mijn bachelor-stage en eerste masterstage hebben jullie de basis gelegd die mij uiteindelijk tot voltooiing van deze promotie hebben gebracht. En tijdens mijn tweede master-stage Prof. Dr. P. Luiten en Prof. Dr. Matsuyama. Tijdens mijn deze stage hebben jullie mij niet alleen enorm veel geleerd maar ook geholpen mijn droom om ooit in Japan te wonen te doen uitkomen en tot een groot succes te maken. Ook heb ik in die periode mijn eerste twee wetenschappelijke publicaties voltooid wat mij enorm heeft geholpen in mijn verdere carrière. Domo Arigato Onegaishimasu!

Mijn huidige collega's bij MRC die mij de laatste jaren ook enorm gesteund hebben in het afronden van mijn proefschrift.

Gelukkig heb ik naast mijn werk ook nog een hele hoop vrienden die ik mag bedanken. Naast dit wetenschappelijke boekwerk ben ik ook enorm trots dat ik al deze fijne mensen mijn vrienden mag noemen. Ik voel me enorm rijk dat ik zoveel regels en namen mag typen om al deze mensen te bedanken:

Het "magic clubje": Ron, Tom, Jeroen en Jeroen. Al zo'n 18 jaar zijn we vrienden! Bizar dat we elkaar al zo lang kennen en dat we nog steeds zulke nerds zijn. Heerlijk! Ik vind het echt ongelofelijk dat we nog steeds zo'n warme band hebben na al die jaren. Hoewel we een tijdje wat uitgewaaierd zijn, vind ik het fantastisch dat we elkaar nog steeds zo veel zien en spreken. Met name onze chill weekendjes zijn altijd een genot! Ik hoop dan ook dat we dit nog heel erg lang mogen blijven doen! Bedankt voor jullie vriendschap, alle herinneringen, tough love, truth bombs, nerdgasms en dat jullie mij hebben vergezeld tijdens al die jaren.

Het "bankje": Simone, Marten, Elize, Jasmijn, Johannes, Anemoon, Wicher, Jeroen, Els, Gert, Dominique, Willem, Melina, Mark en Nadine. Ook jullie ken ik eigenlijk bizar lang! Het is mij een eer om met jullie samen op te groeien en jullie allemaal te zien uitbloeien tot de mensen die jullie nu zijn. Niks dan bewondering en dankbaarheid!

En laat ik mijn andere middelbare schoolvrienden Dieuwke, Suzy en Jochem vooral ook niet vergeten. Onze etentjes in Groningen zijn nog steeds hele warme herinneringen voor mij en ik hoop dat we elkaar ook nog lang blijven zien.

Splintergroepering Plankton, The Family Guys, de Dipsauzers, Jaarclub Bowser: Jonathan, Toon, Jesse, Bas, Sietse, Egbert, Maarten, Joey en Edin. Bij gebrek aan een Albertus jaarclub, moest ik het maar met jullie doen. Maar dat bleek gelukkig geen straf. Bedankt voor de vele afleidingen van mijn werk en mooie herinneringen aan mijn studenten tijd! Wat tof dat we na al die jaren nog steeds contact hebben en ik ben dan ook erg vereerd dat jullie ongetwijfeld weer met grote getalen zullen komen opdagen voor mijn promotie. Fantastisch hoe het jullie allemaal een beetje gelukt is om door de jaren heen iets volwassener

te worden. Ik twijfel er geen moment aan dat jullie weer voorbeeldig zullen gedragen tijdens mijn promotiefeestje! Hoewel jullie, op Sietse na, niet veel bij hebben gedragen aan de wetenschappelijke kant van mijn studie en promotietraject hebben jullie wel enorm veel afleiding verzorgd!

De vrijdagmiddag PhD borrel gang: Amarins, Gert, Suzy, Mario, Emi, Jonathan, Guliz, James, Inge, Stefan, Barbara, Leonie, Jennifer, Melanie, Shamiso en Marnix voor de heerlijke Murphy's en goeie gesprekken in de Ierse Pub O'Ceallaigh op elke vrijdagmiddag na een week promoveren en labwerk.

De nieuwe vrienden die ik de laatste jaren heb mogen ontmoeten in Utrecht: Rob, Miriam, David, Chase, Kayla, Jeffrey, Valerie, Daniel, Jennifer, Mathijs en Edze. Bedankt dat jullie me zo thuis hebben doen voelen in Utrecht nadat ik na 10 jaar uit mijn dierbare Groningen moest vertrekken.

Maarten en Agnieszka. Ik kijk uit naar ons welverdiende diner bij de Fat Duck om dan eindelijk te kunnen vieren dat we alle drie gepromoveerd zijn!

En ook Willem Jan en Melissa: De Fokkinks! Dank voor jullie steun en vriendschap!

Mijn biologen en artsen clubje: Brenda, Sander, Lieke, Simon, Wendel, Willem, Jasper, Marieke, Jannes, Myrthe, Mendelt, Dieuwke, Martin, Lilly, Lenn, Bennet en Maria. Voor de borrelavonden, dansavonden, huisfeestjes, diepgaande gesprekken en vooral mooie herinneringen. Ik ben ook enorm blij dat ik nu eindelijk bij het clubje van gepromoveerde mag horen binnen deze club. In het bijzonder Brenda en Wendel, mijn huisgenootjes aan het Hanzeplein tijdens mijn promotie. Bedankt voor de rust en het thuisgevoel dat jullie mij gaven om bij te komen van de PhD stress. Het was altijd heerlijk thuiskomen met thee, wijn en Pointless.

Mijn Dungeons & Dragons groep van tijdens mijn promotie: Wouter, Els, Bennet en Mike. Voor de hilarische afleiding en avonturen tijdens mijn promotie!

Alle andere fijne biologen en levenswetenschappers van de GBC, Melior Vita en GLV Idun zoals Charlotte, Jelle, Iris, Anneleen, Robert, Daan, Freek, Hans, Frank, Jeroen, Steffen, Mike, Els, Bennet en alle vele anderen. Te veel om op te noemen!

En ik wil graag iets langer stil staan bij mijn paranimfen. Gert, ook wij kennen elkaar enorm lang en ik heb het idee dat we over de jaren alleen maar closer zijn geworden. Je bent een constante betrouwbare factor die altijd bereid is te helpen en voor gezelligheid. Het is mij een eer dat ik jouw paranimf heb mogen zijn en ik ben erg verheugd dat jij nu mijn paranimf bent. Ik ben erg bij dat je nu in ook in Utrecht woont met Dominique en dat we elkaar nu weer regelmatiger zien. Met name de bordspellen avonden bij brouwerij de kromme haring zijn mij altijd een genoegen. Onze reizen samen naar Japan en Zuid-Korea zijn herinneringen die ik de rest van mijn leven zal blijven koesteren. En ik twijfel er niet aan dat we in de toekomst alleen maar meer mooie herinneringen zullen maken!

En Toon. Jij was er bij vanaf bijna mijn allereerste dag in Groningen. Ik was net uit huis en nog verlegen en onzeker maar dankzij jouw vriendschap heb ik mijn vleugels uitkunnen slaan en uit kunnen groeien tot ik wie nu ben. Mijn studentenleven is bijna synoniem met onze vriendschap en ik ben daarom enorm

dankbaar dat jij mij ook tijdens mijn verdediging zal bij staan. Je bent nog steeds mijn grote steun en toeverlaat. Met jou als paranimf komt mijn promotie helemaal goed!

En tenslotte mijn familie. Met als eerste Vincent en Janneke. Vincent, ik heb altijd opgekeken naar mijn grote broer. Hoe hard je werkt, je levenslust en hoe jij altijd met iedereen vrienden kan maken, bewonder ik enorm. Ik ben dan ook blij te merken dat ik naarmate ik ouder word, ik voor mijn gevoel steeds meer op je ga lijken. Daarnaast ben ik je ook enorm dankbaar dat ik je mij niet alleen vroeg hebt afgeleerd een wijsneus te zijn maar ook voor dat ik door jou al vroeg in aanraking kwam met de muziek van de Heideroosjes, The Prodigy en The Offspring wat mij uiteindelijk tot de muziekliefhebber heeft gemaakt die ik nu ben. Het maakt mij ook enorm gelukkig om je nu met Janneke en Lucca te zien. Want Janneke, je hebt niet alleen overduidelijk Vincent enorm gelukkig gemaakt, maar de gehele Clausen clan! Vincent, dank je dat je Janneke en Lucca in mijn leven hebt gebracht. Al het geluk van de wereld!

En de belangrijkste heb ik voor het laatst bewaard: mijn ouders. Lieve pa en ma, ik weet niet hoe en met welke woorden ik eer kan doen aan mijn dankbaarheid naar jullie toe. Van het begin af aan heb ik van jullie geleerd om mijn best te doen op school en goed te leren. Zonder deze basis had ik nooit zo ver kunnen komen. Niet alleen zitten jullie tijdens mijn succes super trots vooraan, ook als het even minder gaat, staan jullie als eerste voor mij klaar. De afgelopen jaren heb ik jullie vast een hoop zorgen opgeleverd, ik ben dus ook enorm opgelucht naar jullie toe dat deze periode nu echt is afgerond. Het is zo fijn om te weten dat jullie altijd voor me klaar staan. Ik hou van jullie. Het is daarnaast fantastisch om na al die jaren jullie nog steeds zo gelukkig hand-in-hand te zien lopen en samen te blijven groeien. Ik ambieer net zoals jullie oud te worden!

LIST OF PUBLICATIONS

Ischemia-induced neural stem/progenitor cells in the pia mater following cortical infarction

Nakagomi, Takayuki; Molnár, Zoltán; Nakano-Doi, Akiko; Taguchi, Akihiko; Saino, Orie; Kubo, Shuji; **Clausen, Martijn JAM**; Yoshikawa, Hiroo; Nakagomi, Nami; Matsuyama, Tomohiro.

Stem cells and development 20 (12), 2037-2051, 2011

Ischemia-induced neural stem/progenitor cells express pyramidal cell markers

Clausen, Martijn JAM; Nakagomi, Takayuki; Nakano-Doi, Akiko; Saino, Orie; Takata, Masashi; Taguchi, Akihiko; Luiten, Paul; Matsuyama, Tomohiro;

Neuroreport 22 (16), 789-794, 2011

Identification of methylation markers for the prediction of nodal metastasis in oral and oropharyngeal squamous cell carcinoma

Melchers, Lieuwe J; **Clausen, Martijn JAM**; Mastik, Mirjam F; Slagter-Menkema, Lorian; van der Wal, Jacquiline E; Wisman, G Bea A; Roodenburg, Jan LN; Schuuring, Ed.

Epigenetics 10 (9), 850-860, 2015

Head and neck squamous cell carcinomas do not express EGFRvIII

Melchers, Lieuwe J; **Clausen, Martijn JAM**; Mastik, Mirjam F; Slagter-Menkema, Lorian; Langendijk, Johannes A; van der Laan, Bernard FAM; van der Wal, Jacqueline E; van der Vegt, Bert; Roodenburg, Jan LN; Schuuring, Ed.

International Journal of Radiation Oncology* Biology* Physics 90 (2), 454-462, 2014

FGFR family members protein expression as prognostic markers in oral cavity and oropharyngeal squamous cell carcinoma

Clausen, Martijn JAM; Koole, Koos; van Es, Robert JJ; van Kempen, Pauline MW; Melchers, Lieuwe J; Koole, Ron; Langendijk, Johannes A; van Diest, Paul J; Roodenburg, Jan LN; Schuuring, Ed.

Molecular diagnosis & therapy 20 (4), 363-374, 2016

Identification and validation of WISP1 as an epigenetic regulator of metastasis in oral squamous cell carcinoma

Clausen, Martijn JAM; Melchers, Lieuwe J; Mastik, Mirjam F; Slagter Menkema, Lorian; Groen, Harry JM; van der Laan, Bernard FAM; Van Criekinge, Wim; De Meyer, Tim; Denil, Simon; Wisman, G Bea A; Roodenburg, Jan LN; Willems Stefan M; Schuuring, Ed.

Genes, Chromosomes and Cancer 55 (1), 45-59, 2016

RAB25 expression is epigenetically downregulated in oral and oropharyngeal squamous cell carcinoma with lymph node metastasis

Clausen, Martijn JAM; Melchers, Lieuwe J; Mastik, Mirjam F; Slagter Menkema, Lorian; Groen, Harry JM; van der Laan, Bernard FAM; Van Criekinge, Wim; De Meyer, Tim; Denil, Simon; Wisman, G Bea A; Roodenburg, Jan LN; Schuuring, Ed.

Epigenetics 11 (9), 653-663, 2016

DNA methylation of KCNA5 and TIMP3 detects very accurate the presence of OSCC in DNA from saliva

Clausen, Martijn JAM; Boeve, Koos; Melchers, Lieuwe J; Mastik, Mirjam F; Slagter Menkema, Lorian; Groen, Harry JM; van der Laan, Bernard FAM; Van Criekinge, Wim; De Meyer, Tim; Denil, Simon; Wisman, G Bea A; Witjes, Max JH; Roodenburg, Jan LN; Schuuring, Ed.
Submitted 2019

Epigenetic regulation of S100A9 expression is related to lymph node metastasis and disease specific survival in patients with oral squamous cell carcinoma

Clausen, Martijn JAM; Melchers, Lieuwe J; Boeve, Koos; Mastik, Mirjam F; Slagter Menkema, Lorian; Groen, Harry JM; van der Laan, Bernard FAM; Van Criekinge, Wim; De Meyer, Tim; Denil, Simon; Wisman, G Bea A; Witjes, Max JH; Roodenburg, Jan LN; Schuuring, Ed.
Submitted 2019

The frequency of SMN gene variants lacking exon 7 and 8 is highly population dependent

Vijzelaar, Raymon; Snetselaar, Reinier; **Clausen, Martijn JAM**; Mason, Amanda G.; Rinsma, Marrit; Zegers, Marinka; Molleman, Naomi; Boschloo, Renske; Yilmaz, Rizkat; Kuilboer, Romy; Lens, Sylvia; Sulchan, Syamiroh; Schouten, Jan.
PLoS One. 14(7):e0220211, 2019

CURRICULUM VITAE

Martijn Jacobus Antonius Maria Clausen (Steenwijk, The Netherlands, 10th of March 1986) received his gymnasium diploma in 2004 from the R.S.G. Steenwijk.

The same year he started the bachelor Life Science and Technology at the Rijksuniversiteit Groningen. He received his degree in 2007 with a specialization in Human Physiology and Pharmacology and finishing his Bachelor Thesis "Age-dependent reduction of A-kinase anchoring proteins in human and mouse brain" at the department of Molecular Neurobiology. After completion of his Bachelor degree he spend a year working and travelling to China, Malawi and Tanzania. In 2008 he started the Master degree Biomedical Sciences at the Rijksuniversiteit Groningen. He received his Master degree in 2011 after completing the thesis "Ischemia-induced neural stem/progenitor cells express pyramidal cell markers" at the Rijksuniversiteit Groningen and the Institute for Advanced Medical Sciences, Hyogo College of Medicine, Nishinomiya, Hyogo, Japan.

After his masters he started his PhD at the department of Pathology and the department of Oral and Maxillofacial surgery at the Rijksuniversiteit Groningen under supervision of Prof. Dr. E. Schuuring and Prof. Dr. J.L.N. Roodenburg. His PhD was part of the CTMM consortium which helps to drive the development of molecular diagnostics and molecular imaging technologies, focusing on translational aspects of molecular medicine. His research focused on the discovery of genetic markers that can help predict the presence of lymph node metastases in the neck of patients with Oral Squamous Cell Carcinomas. His methods combined genome-wide discovery assays, experimental validation as well as bioinformatic validation using large publicly available datasets.

Currently he is employed at MRC Holland as a bioinformatician and product developer focusing on imprinting diseases.