Epigenetic and genetic influence on radiation response

by

Rita Halvorsen

Master i Biomedicine

Department of Health Sciences

Thesis submitted for the Master degree, 60 ECTS, Department of Genetics, Institute for Cancer Research, OUS Radiumhospitalet and Oslo University College

20 mai 2011

oslo university college
faculty of health sciences

Oslo
University Hospital
Radiumhospitalet

Contents

Abbreviations

AD	Allelic discrimination
cDNA	complementary DNA
DFS	Disease free survival
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dsDNA	Double stranded DNA
EGFR	Epidermal growth factor receptor
ER	Estrogen receptor
EtBr	Ethidium bromide
gDNA	genomic DNA
GPCR	G protein-coupled receptor
Gy	Gray
HER2	Human epidermal growth factor receptor 2
IDC	Infiltrating ductal carcinoma
ILC	Invasive lobular carcinoma
IR	Ionizing radiation
LD	Linkage disequilibrium
LOH	Loss of heterozygosity
MAF	Minor allele frequency
MGB	Minor groove binder
NBCG	Norwegian breast cancer group
NHEJ	Non-homologous end joining
PAM	Prediction analysis of Microarrays
PCR	Polymerase chain reaction
PI(4, 5) P2	Phosphatidylinositole-4, 5 biphosphonate
PI3K	Phosphatidylinositol 3-kinase
RT	Radiation therapy
RTK	Receptor tyrosine kinase

SAH	S-adenosylhomocysteine
SAM	S-adenosyl-L-methionine
SNP	Single nucleotide polymorphism
ssDNA	Single stranded DNA
TAE	Tris-acetate-EDTA buffer
UV	Ultra violet
V	Voltage
WT	Wild type

Abstract

Clinical experience has demonstrated that some tumors are more resistant towards ionizing radiation than others. There is strong evidence that both genetic and epigenetic characteristics of the tumor and surrounding tissue have impact on the response and outcome of treatment. As of today, there are no clinical tests that can identify patients with high risk for development of severe late adverse effects after radiotherapy, or those who are likely to respond poorly to the treatment. These tests would have provided the opportunity to optimize the radiotherapy, but in order to develop such tests we need better knowledge about the molecular mechanisms behind radiation response. We have studied 19 pairs of tumors sampled before and after radiation treatment for molecular markers of radiation response. The low number of cases and events increases the probability of false positive findings in microarray analysis. The results from this study should therefore be validated in independent datasets. The overall aim of this study was to investigate whole genome epigenetic effects for their impact on response to radiotherapy as well as selected genetic variations in *TP53, MDM2, MDM4* and *PIK3CA*. Mutations in *TP53* and SNP's in *MDM2* and *MDM4* were selected because of the pivotal role of the p53 pathway in stress response. Two exons in *PIK3CA* were sequenced because mutations in these regions are reported to contribute to radiation resistance. DNA methylation was investigated since this epigenetic alteration is thought to have great impact on tumorigenesis and response to cancer treatment, because of the ability to regulate gene expression. Based on the methylation level of 84 genes we were able to separate the irradiated tumor samples from the non-irradiated tumor samples with an overall error rate of 12,8%. Some of these genes, such as the inflammatory cytokine *IL1A* and *H2AFY* involved in DNA-repair were significantly differentially methylated in the irradiated compared to the non-irradiated tissue. For some genes such as *IL1A*, *PPGB* and *H2AFY* the alteration in methylation levels showed evidence for a dose dependency. Based on their clinical evaluation, samples were then divided into two major groups, good and poor responders. With an overall error rate of zero, samples from the two response groups were separated based on the methylation levels in 342 genes. The gene *ESRRB* was highly methylated in those with good response, while the gene *EGR2*, which is reported to be proapoptotic, was highly methylated in those reported with poor response. Overall, genes involved in the pathways connected to the immune response were found differentially methylated in tumor tissue exposed to irradiation compared with non-irradiated tumor tissue, as well as between the good and reduced responders. A tendency towards lower response to treatment was seen in those tumors harboring *TP53* mutations together with the unfavorable variant SNP's in *MDM2* or *MDM4*, or in those tumors carrying alterations *PIK3CA*. These results suggest that both genetic and epigenetic alterations have impact on the response to radiotherapy, and that methylation levels in both genes and pathways, are changed as a result of irradiation.

Sammendrag

Klinisk erfaring har vist at noen tumorer er mer resistent til strålebehandling enn andre. Mye tyder på at både genetiske og epigenetiske faktorer påvirker respons og utfall av behandling. Per dags dato foreligger det ingen test for å kunne identifisere pasienter som enten har høy risiko for å utvikle bivirkninger, eller som ikke responderer på behandling. En slik type test ville ha forbedret mulighetene for å optimalisere strålebehandlingen. For å kunne utvikle en slik test må mer av de grunnleggende biologiske mekanismene bak stråleresponsen undersøkes. Vi har undersøkt 19 tumor par samlet inn fra brystkreftpasienter før og etter bestråling, for å identifisere molekylære markører for strålerespons. På grunn av få pasienter i materialet er sjansen for tilfeldig eller ikke signifikante funn i analysene høy. Funnene bør derfor bekreftes i et uavhengig prøvemateriale.

Hovedmålet med denne oppgaven var å undersøke hvilken påvirkning helgenom epigentiske effekter har på stråleterapiresponsen, så vel som betydningen av genetiske variasjon i *TP53, MDM2, MDM4 og PIK3CA*.

Mutasjoner i *TP53* og SNP'er i *MDM2* og *MDM4* ble valgt på grunn av den avgjørende rollen p53 reaksjonsveien spiller i stress respons. To exoner i genet *PIK3CA* var valgt ut fordi mutasjoner i dette området er rapportert til å bidra til stråleresistens. DNA metylering ble undersøkt fordi man tror at denne epigenetiske endringen har stor betydning på utvikling av tumor og respons til behandling, på grunn av dens evne til å regulere genuttrykket.

På bakgrunn av metyleringsnivået til 84 gener og en feil rate på 12,8%, kunne de strålebehandlede tumorprøvene separeres fra de som ikke hadde vært utsatt for stråling. Gener som *IL1A,* et proinflammatorisk cytokine og *H2AFY,* involvert i DNA reparering, var signifikant ulik metylert i ubestrålt tumorvev, sammenlignet med bestrålt tumorvev. Endringene i metyleringen for genene *IL1A* , *H2AFY* og *PPGB* så ut til å være avhengig av dose. Ved hjelp av 342 gener og en feil rate på null, kunne tumorprøvene bli separert i forhold til rapportert respons til strålebehandlingen. Genet *ESRRB* var høyt metylert i de med rapportert god response, mens det proapototiske genet *EGR2* var høyt metylert i de med dårlig respons til strålebehandlingen. Gener involvert i ulike immunrespons signalveier ble funnet ulikt metylert i strålebehandlet tumorvev sammenlignet med ubestrålt tumorvev, samt mellom de gruppene som responderte forskjellig til strålebehandlingen. Det ble funnet en tendens til dårligere respons i de som hadde *TP53* mutert i tillegg til de ugunstige SNP ene i *MDM2* eller *MDM4*, eller som hadde en mutasjon i *PIK3CA*. Gener involvert i signalveier forbundet med immunrespons ble funnet ulik methylert i strålebehandlet tumorvev sammenlignet med ikke strålebehandlet tumorvev, i tillegg til i de ulike responsgruppene. Disse resultatene viser at både genetiske og epigenetiske endringer har innvirkning på responsen til stråleterapi, og at ioniserende stråling påvirker metyleringen i både gener og deres signalveier.

1 Introduction

Cancer is one of the most common human diseases, and in 2008 there were 26121 new cancer incidences recorded in Norway. The number of cases is expected to continue increasing towards 2020. Prostate cancer is the most frequent cancer in men, while breast cancer is the most commonly diagnosed cancer type among women. Environment, heritage and lifestyle is thought of as the main risk factors for cancer, and as much as 1/3 of the cases is suggested to be caused by lifestyle choices (1).

1.1 Cancer as a genetic and epigenetic disease

Cancer is an evolutionary process where the cells are not following the basic rules. Normal maintenance of the cell includes cell division and growth, differentiation and death. If one or more of these normal processes become disrupted by mutations, a cell can achieve advantages like more frequent cell division. Cells with selective benefits may be founder of a mutant clone, allowing further mutations to accumulate and hence make the clone more malignant (2). There are different hypothesis for explaining the tumor evolution. The clonal evolution model is based on the evolvement of tumor and accumulation of mutations, where the tumor cells acquire favorable genetic and epigenetic traits. In that way, the tumor cells are able to achieve stem cell features such as self-renewing division. The cancer stem cell hypothesis is based on the cancer stem cells opportunities to self-renewing, in addition to accumulate genetic mutations, and develop drug resistances and drive the tumor progression further (3).

The tumorigenesis is a multistep process where normal cells turn into premalignant cells that after multiple genetic alterations become invasive cancer cells. For a normal cell to become a cancer cell it has to acquire specific capabilities common for all types of human cancer, often referred to as the *hallmarks of cancer.* These biological acquired features include self-sufficiency in growth signals, insensitivity to growth inhibition (antigrowth signals), evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis. Recently, genomic instability and inflammation have also been suggested as hallmarks of cancer (4).

It is also clear that epigenetic alterations are important factors, adding more knowledge to the regulatory circuitry, without necessarily upsetting the established hallmarks (4). Epigenetic factors such as DNA methylation (described in chapter1.8.1), are important for tumorigenesis, because the methylation pattern are closely related to gene expression, and changes in this pattern may lead to favorable conditions for cancer cells (2). Typically, a cancer cell will acquire a methylation pattern with silencing of tumor suppressor genes and activation of oncogenes, leading to malignant transformation (5) .

1.2 Breast Cancer

Annually, 1 million women worldwide are diagnosed with breast cancer, and the incidence rates are increasing. In 2008, 2774 new breast cancer cases were registered in Norway, and the occurrence of the disease has doubled over the last 30 years. In addition to the increased frequency, the breast cancer survival rate has also increased markedly, and was in the period 2004-2008 on 87.8%, when looking at all stages combined (1). Today, breast cancer patients

live longer with their disease, which challenges the treatment in terms of how much side effects are acceptable with respect to quality of life after treatment.

1.3 Breast biology and breast cancer development

The human breast is a complex organ consisting of different types of cells and tissue (Figure 1), making breast cancer a highly heterogeneous disease, both molecularly and at the clinical level (3). The development of a normal mammary gland is hormone dependent during and after puberty. The mammary gland express both estrogen receptor α[1] (ERα) and progesterone receptor(PR), which can be activated by the hormones estrogen and progesterone, leading to intrinsic signaling and transcription of target genes crucial for development and function of the mammary gland (7). ERα signaling is required for elongation and branching of the mammary ducts. Progesterone is important for stimulating growth of lobuloalveolar structure (7). Several studies indicates that ERα influences tumor progression (6). Another important receptor in epithelial breast cells is HER2 (also termed ERBb2), a member of the ERBb family and known to play a role in several pathways involved in proliferation and cell survival pathways such as MAPK and AKT (8).

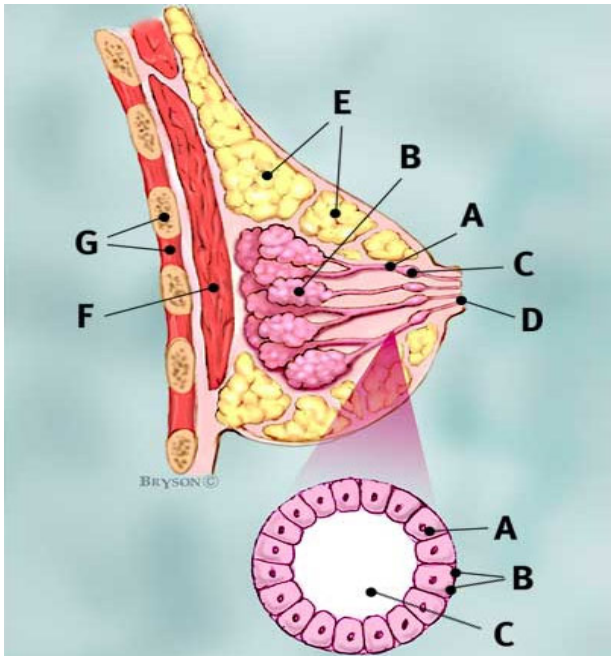

Figure 1: The major components of the human breast are the ducts (A), the lobules (B), the dilated section of duct to hold milk (C), the nipple (D) and fat (E),Pectoralis major muscle (F) and chest wall with ribs (G). Below: A normal duct with cell layer surrounded by the basement membrane (B) and the lumen (C), which is the center of the ducts (9)

Breast cancer usually begins in the cells of the lobules (B) or the ducts (A) (Figure 1, above). The lobules are milk-producing glands, and the ducts are the passages that drain milk from the lobules to the nipple. Rarely, breast cancer begins in the fatty or fibrous tissue, which is the stromal tissue(9). The development of breast cancer cells consists of multiple steps driven by both genetic and epigenetic alterations as well as changes in the tumor microenvironment (3).

In the last decade, an increasing interest in breast cancer subgroups and gene expression profiling as a new additional tool for clinical decisions, has emerged. On the basis of gene expression profiling, breast cancer can be divided into five distinct molecular subgroups: basal-like, luminal A, luminal B, Her2+/-, and normal breast-like (10).

[1] Estrogen receptor is present in two main forms ERα and ERβ transcribed from different genes (ESR1 and ESR2), different expressed and with different chromosomal location (6).

These groups differ in expression of a subset of genes, where Luminal A and B have high expression of estrogen receptor (ER+) and genes associated with luminal epithelial cells. Among the Her2+, high expression of genes characteristic for ERBB2 amplicon is seen in addition to be ER negative, whereas the group normal breast-like expressed genes associated with normal breast tissue (10). The basal like group is ER negative (10) Progesterone (PR) negative and HER2 negative (often named as triple negative)(3), in addition to show low expression of luminal epithelial genes. These molecular subgroups differ in response to treatment and clinical outcome (3).

1.4 Breast Cancer Treatment

When looking at all breast cancers combined the treatment modalities used include surgery, radiation treatment, chemotherapy, antihormone treatment and immune therapy. The latter is a targeted therapy and used in addition to conventional medicine. The type of treatment is selected on the basis of characteristics such as 1) tumor size (T), lymph node involvement (N) and metastasis status (M) (combined into the TNM stage), 2) histopathological scoring of estrogen and progesterone status, proliferation status (Ki67) and Her2 (as well as the woman's age and general condition) (11). Surgery in combination with radiation therapy are in almost all cases chosen (12). In order to improve the quality of treatment for breast cancer patients, The Norwegian Cancer Group (NBCG) has developed a web page with treatment recommendations for breast cancer patients in Norway. These advices are based on available knowledge and documented research (11).

In this thesis the focus will be directed at radiation therapy as a treatment modality.

1.4.1 Radiation therapy

Radiotherapy is important both as adjuvant treatment with curative intent and as palliative treatment, and has many benefits as a treatment modality. The use of radiation therapy is increasing in line with a tremendous development in treatment planning, with the use of specialized radiation therapy equipment, more accurate dose planning and powerful calculation algorithms. Optimal fractionations and precision of the dose to the target volume are important factors for minimizing the side effect as well as increasing the probability for tumor control (13).

1.4.1.1 Recommendations for radiation therapy from the Norwegian Breast Cancer Group (NBCG)

The recommendations for radiation therapy from NBCG are prepared on the basis of the factors mentioned above, in addition to choice of surgery. Different decisions is made depending on if the entire breast gland is removed (mastectomy) or just the lump (breast conserving surgery), and if the edges of the tumor are free of cancer cells after surgery (free resection edges). Adjuvant treatment like chemotherapy also influences the recommendations. Areas that should be considered for radiation therapy are the breast gland/chest wall and regional lymph nodes (11).

Radiation to the lymph nodes is recommended if 9 or less nodes are affected (Figure 2). If so, the axillary and supraclavicular nodes are always included in the radiation field because these localizations are most prone to develop a recidive (11). If 10 or more nodes are affected, they will be removed and there is no longer indication for radiating this area. The parasternal nodes are included if affected, but the importance is debated, because a relapse

in this area is not common. For patients below the age of 40 (some hospitals use 50), a boost with electrons of 2 Gray (Gy)[2] x 8 fractions is recommended (11).

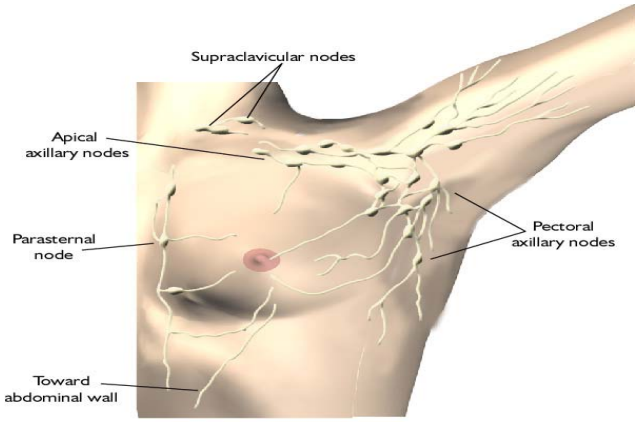

Figure 2 An overview of the lymph nodes surrounding the breast. The axillary and supraclaviculary nodes should always be included if 9 or less nodes are affected (15).

1.4.2 Radiation therapy and side effects

Despite improved recommendations, advancements in treatment planning and quality of equipments during the recent years, some patients still develop severe side effects and injury of normal tissue. Complications after radiation treatment can occur acutely during therapy or after a long latent period (16). Organs at risk defined by the Norwegian Breast Cancer Group (NBCG) are heart, lungs, medulla, contra lateral breast and Plexus brachialis[3] (11). Many breast cancer patients develop skin complications, and some of the most common cutaneous side effects are dermatitis and fibrosis (18). These complications are dependent on treatment related factors such as fractionation pattern, total dose, irradiated volume, anatomy, fractionation rate but also on individual factors like age and type of skin (19).

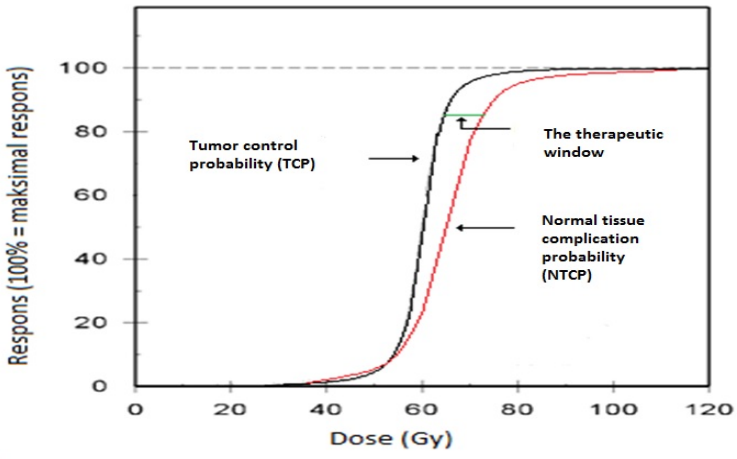

Figure 3 Response to tumor tissue and normal tissue is a function of dose. Response indicates the probability for tumor control or damage of normal tissue. The green line shows the therapeutic window. Modified from (14).

Radiation is a two-sided coin where the given dose is a balance between chance of healing and the tolerance of normal tissue surrounding the tumor. The differences in radiation dose which give the same response in normal tissue as in tumor tissue is called *the therapeutic window* (Figure 3). This is the area where the chances for

[2] Gray (Gy) is the most commonly used unit in describing radiation treatment referring to the absorbed radiation dose (14).
[3] Plexus brachialis is a bundle of nerve fibers originating from the spinal cord and continues through the neck and axillary region and into the arm (17).

complication free tumor control are highest. Increased distance between the two curves (larger *therapeutic window*) increases the chances of a radiocurable tumor (14).

It's also clear that patients receiving the same treatment get unequal reactions in the surrounding normal tissue, varying in strength and grade even after adjusting for the above mentioned factors, and this is thought to be caused by genetic and epigenetic variations between individuals (20).

1.4.3 Ionizing radiation

Photon beams are created by accelerating electrons to a very high energy and then stop the electrons by using a heavy metal. The stop energy generates primarily heat, but a part of this energy is converted to photon beams (21). Ionizing radiation (IR) deposits its energy randomly in the cell. This may cause damages to all the molecules in the cell, such as proteins, mRNA and water. Most of these components exists in multiple copies and undergo a rapid turnover with limited damage caused by the irradiation. However, the main target in cells is the DNA, which only has two copies and is essential for all cellular functions (14). A more indirect effect of the radiotherapy is caused by generation of reactive oxygen species and by-products, which may lead to damage on macromolecules and deregulation of transcription factors and kinases (22). There is also evidence for cell death occurring in unradiated cells caused by irradiated neighboring cells, termed the bystander effect. This can be explained by molecules secreted from irradiated cells, transferred into cells in close contact via gap junction, and thereby exchange molecules (23).

1.5 The cellular response to radiation

In order to survive, the cell injured by ionizing radiation must repair the damage. The cellular response to the damages caused by radiation can be divided into four categories:

Repair - DNA repair is very important for the cell after radiation. Repair mechanisms can take care of a sub-lethal damage if the cells get a few hours break before the next fractionation. Cancer cells often have mutations in DNA repair genes, which make them more radiosensitive than normal cells (21).

Repopulation - The radiated tissue can compensate for radiation damage by increasing the repopulation process. When treatment time is increased, a higher dose is required to control a tumor. Several studies have shown that gaps in treatment time often results in loss of local tumor control due to repopulation. Likewise an accelerated regime can improve the outcome (14).

Redistribution - After a radiation dose is given, the radiation sensitive cells die, and other cells are blocked in the pre-mitotic phase of the cell cycle. The cells are then redistributed in the cell cycle. Loss of cell cycle checkpoints affects tumor proliferation after radiation (21) (see section about cell cycle below).

Reoxygenation – The cells' response to ionizing radiation is strongly correlated to the cells access to oxygen (14). Tumor cells are often hypoxic because of poor blood supply, which can lead to more radiation resistant cells. By use of fractionated radiation, hypoxic cells can be reoxygenated, making them more sensitive to ionizing radiation (21).

1.6 The cell cycle and radiation

When a cell reproduces, it duplicates its content and then divides into two cells. Between the duplicate phase and dividing phase (S phase and M phase), the cell is growing and doubling its content of proteins and organelles. In these phases (G1 and G2), also called the gap phases, the cell has time to monitor the internal and external environment before entering the next phase (2). If the external conditions are unfavorable, the cell can enter a specialized phase known as G_0 (G zero). The cells can remain in the G_0 indefinitely, or they can re-enter the cell cycle when conditions permit (2).

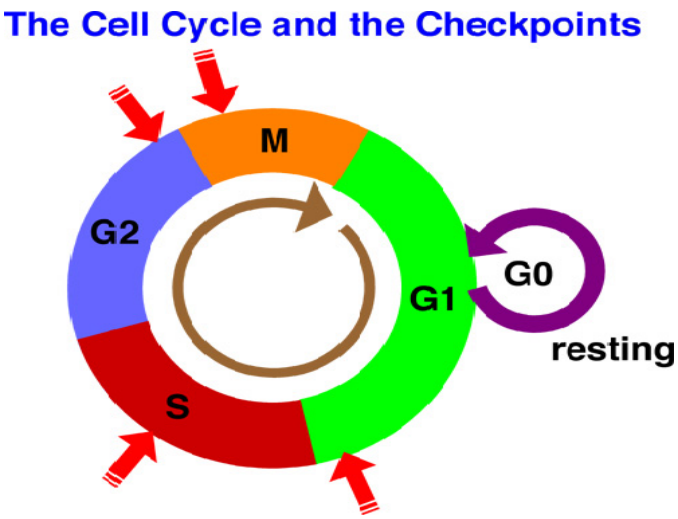

Figure 4 The different phases of the cell cycle with. In G1 phase the cell grows and prepares for DNA synthesis. In S phase DNA replication takes place. The cell will grow and prepare for mitosis in G2 phase and in M phase the cell divides into two cells during mitosis. Checkpoints are marked with red arrows (24).

To ensure accuracy and reliability of the cell-cycle progression, the cell has a control system performed by different checkpoints (2). Four checkpoints are activated by radiation at different phases in the cell cycle. The checkpoint between the G1 and S phase is important for entering S-phase. In this phase the cell is sensitive to growth factors, and unfavorable conditions will be captured by the sensors ATM and p53 (14). Cells that are irradiated in G1 phase will be delayed for entering S phase. The next checkpoint is in S phase, activated indirectly from ATM or ATR. Irradiation of cells in S phase leads to a dose- dependent reduction in the rate of DNA synthesis and increased replication time (14). G2 checkpoint is also ATM dependent, and irradiation in G2 phase can lead to blocking of cell cycle progression at the end of G2. The late G2 checkpoint (or M checkpoint in Figure 4) is activated by ATR, and marks a second delay before entry to mitosis. This delay applies to cells that have been irradiated in G1 or S phases, but still continued through G1- and S- phase checkpoints not adequately repaired(14). Most of the tumor cells have disabled one or more of the checkpoints because of genetic alterations during tumorigenesis (2). The radiation sensitivity is varying during the cell cycle and the most resistant cells are thought to be in S phase, while the most sensitive cells are in late G2 phase when the final checkpoint is passed (14).

1.7 Signaling pathways involved in radiation response

Several pathways are known to influence the radiation response such as non homologous end joining (NHEJ)(25), homologous recombination(HR)(25), PI3-K/Akt signaling pathway (25), Ras signaling pathway (25) and p53 pathway

(26). In this project the p53 pathway and Akt pathway are emphasized due to their central role both in radiation response and breast cancer tumorigenesis.

1.7.1 P53 pathway

The protein p53 transcribed from its gene *TP53,* is involved in a network of pathways comprising genomic stability, apoptosis and cell-cycle control (2). *The* most serious defect occurring to the cell is DNA damage, because it threatens the organism. It's essential for the cell to detect such injury before the cell continues through the cell cycle. If some irregularities are revealed, the cell can be sent to cell-cycle arrest for repair or programmed cell death (apoptosis) if the damage is to severe. The protein p53 has a pivotal role in regulating these actions (2).

1.7.1.1 Activation of p53 and downstream signaling

The different stress factors activating p53 includes genomic damage, hypoxia, nutritional starvation, oncogene activation and mitochondrial biogenesis stress. The first stress signal found to activate p53 was DNA damage(26). Damage caused by radiation can be captured by several elements in the p53 pathway, such as the protein kinase ATM. The protein phosphorylates various target proteins, but the major target is p53, which in turn can induce expression of numerous genes involved in cell cycle arrest, prevention of angiogenesis, DNA repair and apoptosis or senescence (27). High levels of p53 stimulate the transcription of p21, a protein binding to G1/S-Cdk and S-Cdk complexes, and block the progression through cell cycle by arresting the cell in G1 (2). If the DNA damage is to severe, p53 induces apoptosis by stimulating expression of pro-apoptotic genes, such as Bax, Puma (BBC3) and Noxa (PMAIP1). Another function of p53 is to bind and inactivate the anti-apoptotic Bcl2 protein, thereby promoting apoptosis (2).

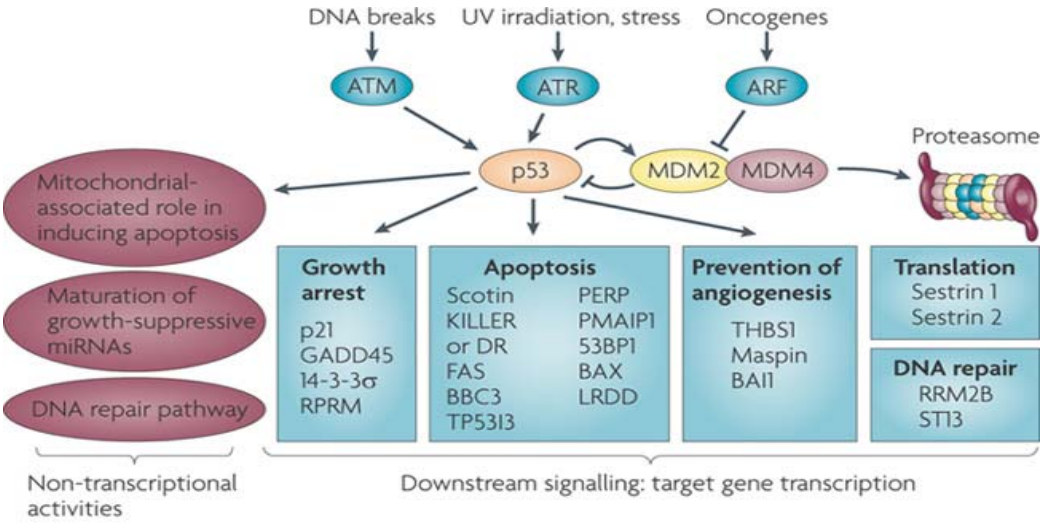

Figure 5 p53 is activated by several mechanisms and is regulated by Mdm2 and Mdm4. The protein p53 is involved in many downstream activities, like growth arrest, apoptosis, prevention of angiogenesis and translation of central genes in cell regulation, which underpins the p53 role as guardian of the genome. During these actions the protein can protect the cell from genomic instability and tumorigenesis (27).

1.7.1.2 Regulation of p53

Several proteins are known to modulate p53's activity. In a normal cell condition, the p53 level is low, and is regulated by the Murine double minute 2 (Mdm2) and the related protein Mdm4 which both interacts with p53 in non-stressed conditions (28). These interactions lead to ubiquitylation and degradation of p53 by the proteasome. In stressed conditions, Mdm2 and Mdm4 are phosphorylated by proteins such as Arf, and their functions are inhibited, which in turn leads to an increased level of p53 (27). Mdm2 is thought of as the key negative regulator of p53 and as such important to achieve well regulated p53 response (29).

1.7.1.3 Dysregulation of p53, MDM2 and MDM4

TP53 is a tumor suppressor gene that has a long time been regarded as "guardian of the genome", and several studies conclude that the p53 protein plays a central role in preventing cancer in humans (27). The loss of p53 or elements in the p53 pathway is a dangerous condition for the cell. A disrupted p53 pathway allows the cell to escape apoptosis, and thereby allowing further cancer-promoting mutations to accumulate in the cell. These DNA lesions will then be propagated to daughter cells, during cell division (2). Accumulation of mutations can also lead to resistance to anticancer drug and irradiation (2).

P53 can be inactivated in different ways. Inactivation of the *ARF* gene, amplification of Mdm2 or Mdm4, or miRNA regulation and repression of p53 expression, are some of the mechanisms that will lead to mitigated p53 function (26). Methylation is another mechanism reported to influence the p53 stability and activity (28). *TP53* is the most frequently mutated gene in all human cancers (27), and up to 50% of all human cancer contains mutation in both alleles of the gene (26). In breast cancer, the gene is found altered in 20 – 40% of all breast cancer cases, and have been reported to play a predictor role for the therapy response (30).

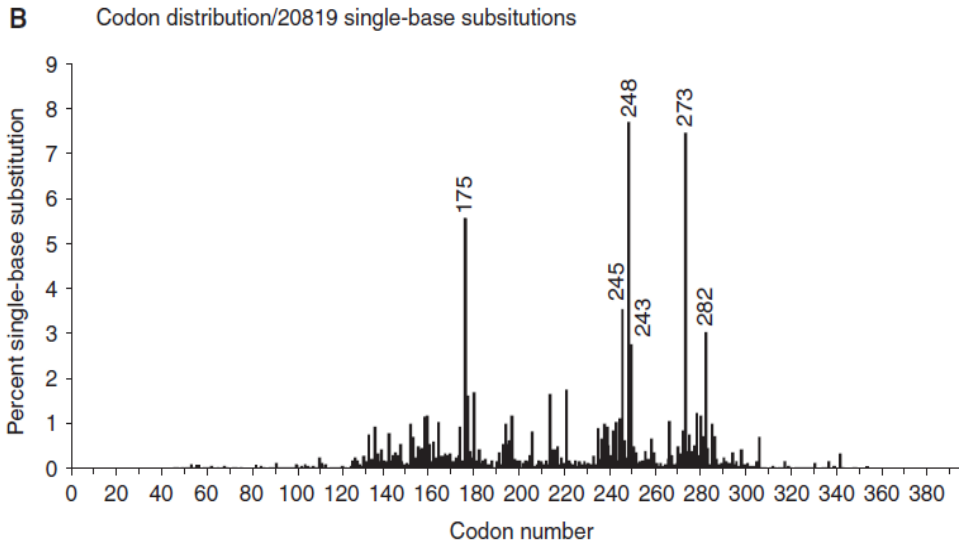

Figure 6 Distribution of the most frequent mutations in the coding sequence of *TP53* gene, and their position (31).

The p53 activity can be lost by several mechanisms, like *TP53* gene missense mutation or deletion of the gene *TP53* (26). Studies that examined the whole coding sequence, report that 86% of the mutations cluster between codon 125 and 300. Most of these mutations are missense mutations and located mainly in the DNA binding domain (31).

Mdm2 is shown to be overexpressed in a subset of tumors and associated with tumor progression and lack of response to therapy (32). Overexpression of Mdm2 leads to low levels of p53, suggesting that overexpression of Mdm2 can substitute for lacking a *TP53* mutation (32). Breast cancer patients with both overexpressed Mdm2 and mutant *TP53* have worse prognosis than those with a single defect (33). Mdm2 have also been implicated in other pathways of importance for tumor development independently of p53. Thus, a tumor will benefit from overexpression of Mdm2 even if p53 is mutated (33).

The structurally homologue protein Mdm4 is also found amplified or overexpressed in many human cancers and tumors without mutations in p53 or Mdm2 amplifications often show Mdm4 overexpression as an alternative molecular mechanism in tumorigenesis (34). Similar to Mdm2, Mdm4 regulates p53 activity, and is suggested to play an important role in tumorigenesis in human breast and ovary cancers (35).

1.7.1.4 Polymorphisms in TP53, MDM2 and MDM4

Several researchers have suggested that a polymorphism in *TP53* (rs 1042522, exon 4 codon 72) may play an important role in many cancers (36). This polymorphism is a substitution from CGC to CCC which leads to an amino acid change from arginine (Arg) to proline (Pro). The change does not affect the ability to bind DNA, but may still alter the biochemical and biological properties of the protein (37). Langerød *et.al* reported in their investigation of 390 breast cancer cases that 28, 5% of those who were homozygous for the Arg 72 allele also carried a *TP53* mutation compared to those homozygous for the Pro 72 allele were only 3,8 % had a *TP53* mutation (37). These results were significant with a P= 0.004. This observation indicates increased risk for carrying *TP53* mutated tumor for those harboring the Arg72 allele, which may affect the prognosis (37).

Recently, a single nucleotide polymorphism (SNP 309, T>G) rs 2279744, was identified in the promoter region of *MDM2* (promoter P2, position -309), reported to influence the expression level of Mdm2 (29). Transcription of *MDM2* can arise from one of two promoters, P1 or P2, the latter located in the first intron. The increase expression level caused by the SNP309 can be explained by an increased affinity for the transcription factor Sp1 when harboring the minor allele in SNP309 located in P2 promoter (29). Another SNP, rs 117039649 (SNP285 G>C) 24 base pairs upstream of SNP309 has recently been reported to have a neutralization effect on SNP 309(38). When harboring the minor allele in SNP 285, the affinity to the transcription factor Sp1 is reduced, leading to reduced levels of *MDM2* transcripts (38).

Hsi-Fang Tu *et.al* demonstrated that the polymorphism *MDM2* SNP309 and *TP53* codon 72 can estimate the prognosis for those with oral squamous cell carcinoma (OSCS). The prognosis was poorest for those harboring Arg/Arg in codon 72 and G/G in SNP309 in addition to receiving adjuvant radiotherapy for advanced OSCS(39). Toyama *et.al* suggested that harboring Pro/Pro in p53 codon 72 may alter the sensitivity of the tumor to treatment with chemotherapy. They found a relationship between the *TP53* Pro/Pro genotype and poorer disease free survival[4] (DFS) in breast cancer patients receiving adjuvant chemotherapy, but did not find any significant association

[4] Disease free survival (DFS) is the time from diagnosis until eventually recurrence (17). In this research the median followed up was 61.7 months (36).

between *MDM2* SNP309 and survival. On the other hand, they did observe a tendency (non-significant) towards better DFS for patients receiving adjuvant tamoxifen and who carried the T/T genotype (36).

Regarding Mdm4, Atwal and colleagues have reported a haplotype of *MDM4* (denoted the neutral haplotype) to be associated with an increased risk of cancer and early onset of tumorigenesis. The neutral haplotype can be separated from the non-neutral haplotype (not observed with the mentioned factors), by genotyping the polymorphism rs 1563828 in intron 10 in *MDM4* (35). Because of the regulating role of Mdm2 and Mdm4 in the p53 pathway, harboring these minor SNP's leads to lower p53 response which may result in a higher mutation rate, poorer DNA repair process, reduced cell cycle arrest, apoptosis and senescence (35). All these factors can lead to tumor formation and may also affect the response to treatment.

1.7.1.5 P53/ Mdm2/Mdm4 and response to radiation

Ionizing radiation is one of the stress factors known to increase p53 expression. In a study of MCF-7 breast cell lines, exposure with 8 Gy of IR lead to a 3 fold increased level of p53 mRNA and a 6-fold induction in the p53 protein (40). Westphal *et.al* revealed the importance of p53 for response to radiation with the study of p53 null[5] mice. They showed that p53 null mice survived after having been exposed to 10Gy γ-irradiation while all the wild type (WT) mice died within 1-2 weeks. Approximately, half of the p53 heterozygous mice died, indicating that this effect was dose dependent (41).

Mendrysa *et.al* found similar results in their study of the p53 pathway and response to ionizing radiation. They used transgenic mouse with reduced levels of Mdm2 and discovered that this protein was important in radiosensitivity. All the transgenic mice died 12-22 days after being exposed to 10 Gy irradiation, while 50% of the wild type mice were still alive 40 days after irradiation (32).

Wang *et.al* demonstrated in their mouse models the importance of Mdm4 down regulation for an effective p53 mediated radiation response and tumor suppression. They created a mouse model 3SA (mice with ATM and Chk2 mutations leading to Mdm4 phosphorylation and subsequent mitigated p53 function) and exposed them to 10 Gy whole body radiation. The models showed increased resistance to death induced by ionizing radiation, compared with WT mice. All the WT mice died within 18 days, while 60% of the mice 3SA lived at least 100 days after irradiation. To achieve adequate p53 activation, phosphorylation of Mdm4 targeted by the damage kinases ATM and Chk2, is required. From a radiation point of view, a reactivation of Mdm4 may have some benefit in cancer therapy by preventing damage in normal tissue when exposed to radiotherapy (42).

These examples illustrate how pivotal the p53 pathways and its elements are to radiosensitivity.

1.7.1.6 TP53 and regulation of Estrogen receptor

Several studies have shown a relationship between ER positive breast tumors and expression of WTp53, while tumors carrying *TP53* mutations are suggested to lead to loss of ER expression (43). Both *TP53* mutations and estrogen negative tumors are associated with tumor development and progression, in addition to decreased disease

[5] Null mice are transgenic mice with a gene knock out, leading to mice not expressing that particularly gene.

free survival (43). Angeloni *et.al* showed in their study that ERα expression is regulated by p53 in the human cancer cell line MCF-7 (43). This is supported by Shirley *et.al in* their study of MCF-7 cells. They suggested that p53 regulates ER expression and to confirm the hypothesis they exposed cells for radiation or doxorubicin (both known to stabilize p53) which led to increased p53 and ER levels. They repeated the experiment with cells targeted with depletion of *TP53*, which resulted in down regulation of ER expression, which confirmed their assumption (40). They also reported a correlation between increased level of Mdm2 expression and ERα positive tumor cells, confirmed by Phelps and colleagues when studying 6 breast cancer cell lines (44).

1.7.2 The Akt-pathway

The surface receptors tyrosine kinase (RTK) and G protein-coupled receptor (GPCR) are activated by binding an extracellular signal protein, such as growth factor. This activation leads to phosphorylation of phosphatidylinositol 3-kinase (PI3K) in interaction with P85 and P110 subunits. This complex activates phosphatidylinositole-4, 5 biphosphonate (PI (4, 5) P2), which converts to PI (3, 4, 5) P3. PTEN can dephosphorylate PIP3 back to PIP2, which limits Akt activation. PIP3 recruits the proteins Akt and PDK1, which leads to activation of Akt via its Ph domain. Once activated, Akt controls the cell cycle and cell survival (45). Akt stimulates cells to survive and grow, mostly by targeting of proteins regulating these functions (2). The PI3-K/Akt pathway can be abnormally activated by several types of alterations, such as mutations, gene amplifications and promoter hypermethylation (45). A dysregulated PI3-K pathway is known to be involved in tumorigenesis (46).

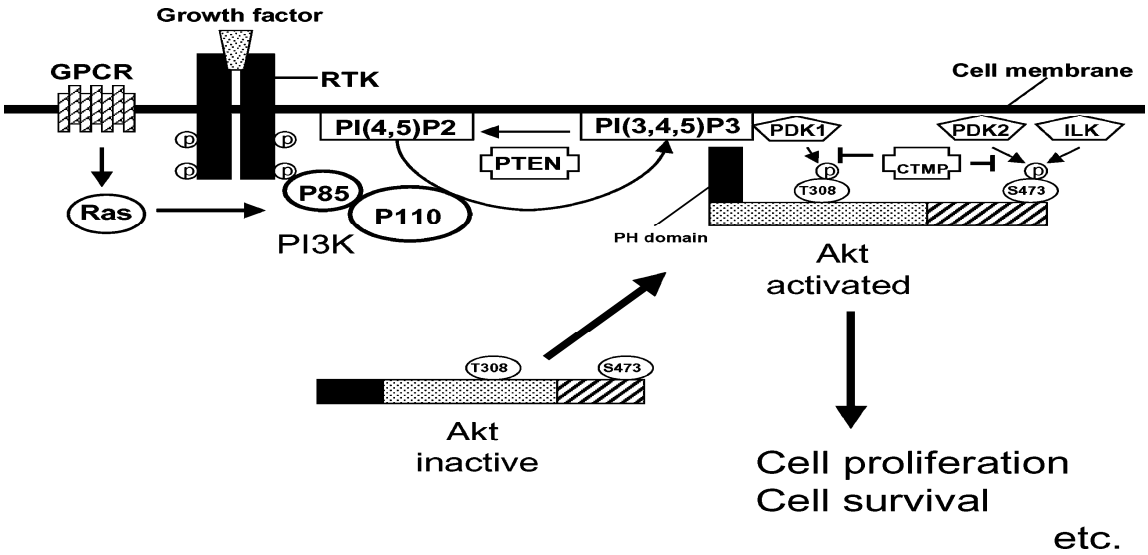

Figure 7 A model for the regulation of the PI3K-Akt signaling pathway (45)

The *PIK3CA* gene encodes for the p110α catalytic subunit, and can be described as an oncogene, because of its ability to increase the PI3K signal, which leads to stimulation of downstream Akt signaling (47). H1047R, E545K and E452K are the three most frequent mutational spots of the *PIK3CA* gene, and are located in exon 9 and exon 20 (48). Mutations in these regions lead to increased kinase activity (47). It is proposed that *PIK3CA* is one of the two most common mutated genes identified in human cancer, the other being *KRAS* (46). In tumors with these mutations, inhibition of the p110 α is pointed out as an attractive therapeutic strategy (46).

PIK3CA gene amplification, activating mutation in *PIK3CA* or inactivation of the tumor suppressor gene *PTEN* are mechanisms detected in ovarian cancers that lead to increased Akt phosphorylation and oncogenic transformation (49). A strong and independent association between the level of pAkt and treatment outcome in Non-small Cell Lung Cancer has been established (25). The PI3/Akt pathway has been demonstrated to play an important role for the development of radiation resistance (25). Many investigators have proposed this pathway as a promising target for increasing radio sensitivity in tumors that have activated the PI3-k/AKT cascade (25). The PI3-K/AKT pathway is essential in three of the main mechanisms to radiation resistances mentioned in chapter 1.5, which are intrinsic radio sensitivity (DNA repair), tumor cell proliferation (repopulation) and hypoxia (25). Ionizing radiation causes not only multiple DNA double and single strand breaks as mentioned earlier, but the radiation also activates pathways leading to repair mechanism, such as homologous recombination or non-homologous end joining (NHEJ). Epidermal Growth Factor Receptor (EGFR) and PI3-K/AKT signaling are involved in the DNA repair by NHEJ, which constitutes the repair of the majority of DNA double strand breaks caused by irradiation (25).

The decreased radiation responsiveness has been associated with increased expression of pAkt in several types of cancer, including head and neck squamous carcinoma (50) , lung carcinoma (25) and breast cancer (51). A significant association between the level of pAkt expression and radiation resistance in cervical cancer has also been revealed (52). AKT has been suggested as a prognostic marker, and the PI3K/AKT-signaling pathway is pointed out as an targeted approach to improve the outcome of radiotherapy (52).

1.8 Epigenetics

The term *epigenetics* designates the heritable modifications in the phenotype, which cannot be explained by changes in the primary DNA sequence (53), and is known to play an important role in normal development as well as in disease initiation and progression (5). The epigenetic machinery includes histone modification, changes in remodeling complexes and DNA methylation (53). It is also debated whether miRNA should be included as an epigenetic feature. The epigenetic status is dynamic and is influenced by factors such as age and environmental stimuli (53). Feinberg describes the term epigenetic disease as: “disruption of phenotypic plasticity – the ability of cells to change their behavior in response to internal or external environmental cues” (5). Defects in the epigenome are known to be involved in abnormal development, which may cause diseases such as Beckwith-Wiedemann syndrome, Prader-Willi syndrome and Angelman syndrome (5).

Chromatin contains 4 core histone proteins, which DNA is wrapped around. Each core histone has an N-terminal amino acid tail, which extends out from the histone core. The histone tails are modified at many sites with different modifications such as acetylation, methylation, phosphorylation and ubiquitylation. The histone modifications are dynamic and have consequences for the chromatin structure (54). The set of modification determines whether DNA is accessible for transcription as in euchromatin, or inaccessible as in heterochromatin where the chromatin is highly condensed (54). In this way, the histone proteins are also involved in regulation of gene expression (55).

DNA methylation is an important epigenetic mechanism because of its influence on many cellular processes and regulation of gene transcription. DNA methylation has a crucial role in maintaining genomic stability by silencing

repeat elements and endogenous transposons (53). To ensure proper differentiation, some tissue-specific genes are known to be controlled by DNA methylation (55). A subset of genes is known to be imprinted by DNA methylation during development and growth of the embryo, and for postnatal function. These genes are located in clusters and becomes imprinted or expressed only from the maternal or paternal allele in order to control normal development (53). In order to achieve dosage compensation, one of the two female X chromosomes in mammals becomes transcriptional inactivated early in development. A random silencing of one of the X chromosomes is required, such that both male and female embryos only have one X chromosome active (53). Aberrant methylation pattern have been detected in several diseases such as Prader-Willi and Angelman syndrome, but cancer is perhaps the most studied disease in terms of change in methylation pattern (53). This thesis will focus on DNA methylation and cancer.

1.8.1 DNA methylation

DNA methylation occurs when a methyl group is transferred from S-adenosyl-L-methionine (SAM) and attached with a covalent binding to 5' position at cytosine (figure 8), a modification mainly seen in combination with CpG dinucleotides. This action is performed by DNA methyltransferases (DNMT) and four DNMTs involved in the methylation process are so far detected, DNMT1, DNMT2, DNMT3, and DNMT3B. The remaining product S-adenosylhomocysteine (SAH) will in high concentrations inhibit the action of DNMT (53).

Figure 8 Modified from (56). The picture depicts how SAM acts as a methyl donor with DNMT as a catalyzator. High levels of SAH can inhibit DNMT.

1.8.1.1 DNA methylation and cancer

DNA methylation is thought to be a reversible modification, playing a crucial role in tumorigenesis. Changes in the methylation pattern may affect numerous of pathways related to development of cancer (57). Increased methylation (hypermethylation) is associated with gene silencing, and decreased methylation (hypomethylation) often leads to gene activation. CpG islands which are unmethylated in normal somatic cells can be methylated and result in inactivation of tumor suppressor genes, which in turn can lead to cancer. On the other hand, hypomethylation of oncogenes, which are associated with increased expression, can also cause development of cancer (5). The global loss of methylation found in cancer may be involved in initiating and developing cancer (53). Hypomethylation of repeat elements may cause genomic instability, another hallmark of cancer (58).

Gene silencing through methylation can affect several cellular processes like cell cycle checkpoint, apoptosis, signal transduction, cell adhesion and angiogenesis (57). DNA methylation can also affect the expression of genes involved

in maintaining the integrity of the genome through DNA repair, detoxification of reactive oxygen species and the induction of senescence, which often is associated with cancers showing mutations in the Ras signaling pathway (57).

Recently, the p53 pathway has been identified as important in relation to epigenetics (26). The gene *TP53* is often mutated in cancer, but in tumors with wild type p53, researchers have detected aberrant methylation patterns in central elements in the p53 pathway. The promoter methylation profile of *ARF* and *PTEN* revealed a significant hypermethylation level whereas *MDM2* was hypomethylated (59). ASPP2 (locus *TP53BP2*), a protein known to bind p53 and induce apoptosis, is reported to be regulated by methylation (60). Recent data have revealed a correlation between increased risk of relapse and demethylation of *TP53BP2*, in *TP53* wild type populations. In *TP53* mutated samples, the gene *TP53BP2*, did not influence on the survival rate (61).

Epigenetic therapies hold promising prospects for personalized treatment regimes. DNA methyltransferase (DNMT) inhibitors and histone deacetylase (HDAC) inhibitors are already in trial. Disadvantages for these class of drugs is the lack of nonselective agents, meaning that these drugs may activate as many genes as they silence (5). Histone acetyltransferase inhibitors is reported to be more selective and may be useful in cancer treatment (5). The methylation status of driver genes may be prognostic markers for breast cancer survival. Revealing the methylation pattern may also predict treatment response in addition to serve as useful biomarker for tumor diagnosis (62).

1.8.1.2 DNA methylation and subgroups in breast cancer

As mentioned in chapter 1.3, breast tumors can be divided into subgroups depending on their gene expression. Recently, researcher have also been able to subgroup the tumors on the basis of their methylation status. The clustering shared some of the same patterns as those performed by expression analysis (61). In another study of methylation profiling in breast cancer, three major clusters were defined, but both Luminal A and basal like were split into two clusters. The subgroups were strongly correlated with *TP53* mutation status and ER status, as well as survival. This may indicate that methylation profiling can add further information about the subgroups (63).

1.8.2 Radiation and epigenetic changes

There is increasing evidence that radio resistance is epigenetic of nature, although the mechanisms are not fully understood. Luzhna and colleagues were able to show a correlation between radiation responsiveness and global levels of DNA methylation in cultured cells (64). Studies have also demonstrated that irradiation induces epigenetic alterations, including DNA methylation, reported as a result from alterations in DNA methyltransferases (DNMT). Down regulation of DNMT may result in loss of global DNA methylation (65). There is also suggested that DNA methylation may have a function in regulating the radiation response. Elements involved in DNA damage pathways, such as ATM, are associated with radio-sensitivity. In that way, the methylation status of these genes may influence the sensitivity to radiation (65). A recent study performed by Aypar *et.al* discovered hypomethylation of repeat elements LINE-1 and Alu when exposed to radiation, compared with controls. Repeat elements are supposed to be inactivated by hypermethylation to prevent genomic instability (66).

Epigenetic drugs have huge potential in cancer treatment, especially in combination with other treatment modalities such as radiotherapy. It may also be an alternative for those resistant to convential therapy (67). For instance, DNMT inhibitors may sensitize the tumor for radiotherapy, leading to increased response and decreased toxicity. Several DNMT inhibitors are developed affecting different parts of cell maintenance and tasks. Still little is known about the effect on non-cancerous cells and conflicting evidence is reported (68).

2 Aim of the study

Today, the knowledge about how, and the extent to which epigenetic alterations influence radiation sensitivity and resistance are limited to data generated from cell lines and retrospective studies where the variable molecular characteristic in tumor tissue after radiation therapy are compared with patients with different radiation sensitivity and -resistance.

In this study, the aim is to increase our knowledge about which effect radiation therapy has on the methylation pattern in tumor tissue, and see if these changes are dose dependent and predictive for response to the treatment. For the samples collected before and after radiation the analysis of the alteration in methylation pattern may highlight the pathways induced in response to irradiation.

Also, the aim is to study if genetic variation and somatic mutations in genes indicated to play a role in the radiation response as well as breast cancer etiology, such as *TP53*, *MDM2, MDM4* and *PIK3CA,* can be shown to have an effect also in this material.

More specifically, we aimed to focus on the following topics:

- Is the methylation pattern in tumor tissue changed as a result of the radiation treatment and is the changes dose dependent?
- Can changes in methylation pattern be connected to differences in radiation response?
- Are changes in the p53 and Akt pathway predictive for radiation response?

3 Materials

Breast tissue biopsies taken from 22 patients treated with radiation were included in this study. The samples were collected both before radiation (22 patients) and after receiving 10 – 20 Gy (19 patients). All patients had inoperable breast cancer stage 3 – 4 or local relapse. The material was collected between 2002 and 2005, and stored mainly as biopsies in -80°C. Extracted DNA was available in a 5°C fridge. Extracted DNA from blood samples drawn from the same patients before radiation therapy was started, was stored in a -20°C freezer.

Clinical information on ER status and response to radiotherapy treatment were available.

All patients had given informed consent, and the project is approved by the Regional Ethical Committee REK Sør. All the samples were de-identified.

Tissue from healthy women who underwent breast tissue reduction was available from -20°C freezer storage, and used as reference material. These women had signed an informed consent.

4 Methods

In this part, the different practical laboratory methods are described. The methods used for DNA extraction and sequencing are performed using standard procedures.

4.1 DNA extraction by Maxwell®16

DNA was extracted from breast tumor biopsies as well as from normal tissue using a Maxwell®16 instrument. This is a fully automated instrument that extracts DNA effectively from blood, cells and tissue in 45 minutes. The purified DNA can be used directly in further analysis. This analysis was performed according to Maxwell® 16 DNA Purification Kits Technical Manual, Literature # TM284 (http://www.promega.com).

4.1.1 Procedure

The Maxwell®16 accepts 16 samples, each of them placed in a cartridge consisting of 7 chambers (Figure 9). The first chamber contains lysis buffer, important for tissue and cell destruction. The other chambers contain washing buffers. MagnesilR paramagnetic particles (PMP) essential for the DNA extraction are located in chamber 2.Tissue with a maximum weight of 50 µg was added to the first chamber, and a plunger was placed in the 7th chamber. Blue cuvettes were filled with 200 – 600 µl elution buffer (dependent of amount of tissue at hand and the required DNA concentration), and placed in a rack in front of the cartridge.

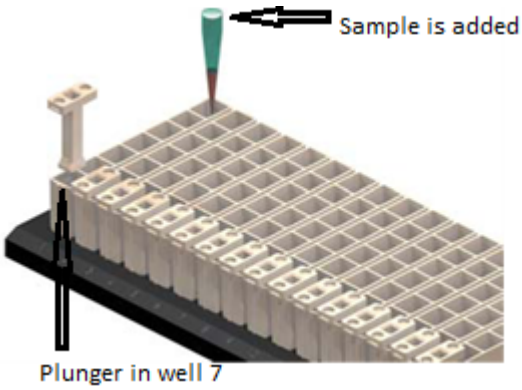

Figure 9 A cartridge consists of 7 chambers. The sample is added chamber 1, and a plunger is placed into chamber 7

The instrument started by picking up the plunger in chamber 7 and homogenizing the sample material in chamber 1. Secondly, the plunger picked up the magnetic particles in chamber 2, and transferred them to chamber 1 for DNA binding. The instrument performed a 5 step washing procedure, before the DNA was finally eluted in the blue cuvettes.

4.2 DNA quantification with absorbance

A NanoDrop® ND-1000 (Saveen Werner A/S) instrument is used for reading DNA-concentration and for evaluating the pureness in the DNA-solution. The instrument is a spectrophotometer measuring the absorbance in the 220 – 750 nm specters. The measurement does not distinguish between RNA, single stranded or double stranded DNA, or other compounds absorbing at 260nm. All the nucleic acids absorb in the same area. A 260/280 ratio is used to calculate the pureness of the DNA-solution. A ratio below 1.8 indicates that the sample contains protein, phenols or

other pollutions which absorb light at 280nm. Only 1µl of sample is required, which is placed directly onto the end of a fiber optic cable. The spectrophotometer is operated from a connected computer.

4.2.1 Procedure

The analysis was performed by loading 1µl of the sample onto the measurement surface and the sampling arm was lowered. The 260/280 ratio was checked for an indication of the pureness of the DNA.

4.3 DNA quantification with fluorescence

To make sure we had enough doubled stranded DNA (dsDNA) for our analysis, the DNA solution was also measured with the Quant-iT™dsDNA Assay Kit, Broad Range (Invitrogen) on Thermo Scientific NanoDrop 3300 Fluorospectrometer (Saveen Werner A/S). The fluorophore PicoGreen binds only to dsDNA and is used for quantifying dsDNA. When the PicoGreen is bound to DNA and exposed to Blue Led light at λ 525 +/- 20nm, the amount of emitted light detected correlates to the dsDNA concentration of the sample. The sample DNA concentrations are obtained using a 5 point standard curve in the range of 0 – 1000 ng following the supplier's handbook Quant-iT™dsDNA Broad-Range Assay Kit (http://www.invitrogen.com). The fluorospectrometer is controlled by a computer.

4.3.1 Procedure

- 1µl Quant-iT™dsDNA BR reagent (PicoGreen) was mixed with 199µl Quant-iT™dsDNA BR buffer and vortexed for 2-3 seconds used for working solution.
- 2µl of the sample was mixed with 198µl working solution, vortexed and spun down. The mix was incubated for 2 minutes at room temperature and protected from light using aluminum foil.
- The sample mix was measured by loading 2µl onto the NanoDrop 3300 Fluorospectrometer.

4.4 Sequencing of *MDM2*, *TP53* and *PIK3CA*

The sequencing analysis of the genes *MDM2, TP53 and PIK3CA* were performed on an Applied Biosystems 3730 DNA Analyzer according to the supplier's handbook Applied Biosystem 3730/3730X/DNA Analyzers Part 4331467 Rev.B (http://www.appliedbiosystems.com).

The promoter region of *MDM2* contains two SNP's of interest, SNP 309(rs 2279744) and SNP 285 (rs 117039649), where SNP 309 is the most studied one. A region covering both variants was sequenced. In the gene *TP53,* an interesting SNP in codon 72 is described in the literature, and therefore included in the study. The SNP analysis was performed using blood samples since the SNP's are germ line polymorphisms inherited from the parents. An exception was the healthy control samples and patient number 127, where only DNA from healthy breast tissue and tumor tissue was available respectively. Exon 9 and exon 20 in the gene *PIK3CA* and all exons and flanking introns of TP53 (except exon 1 which is noncoding) were sequenced and investigated for mutations. The mutations are somatic and therefore DNA isolated from tumor tissue was used in this analysis.

The sequencing procedure is divided into seven different steps, starting with a PCR reaction followed by gel electrophoresis to ensure that the PCR-reaction was successful, clean-up of the PCR product, the sequencing reaction itself, clean-up of the sequence reaction product using Sephadex® columns, the capillary electrophoresis in the ABI PRISM® 3730 instrument and finally aligning and analyzing the samples in SeqScape (v.2.5, Applied Biosystem).

4.4.1 The PCR reaction

The purpose of PCR amplification is to generate many copies of the DNA area of interest for the cyclic sequencing.

Table 1 shows the recipe of the PCR reagent mix required for analyzing 1 DNA sample. The DNA samples with high concentration were diluted with water to 10ng/µl. In samples with low concentration, 1,5µl DNA was used and the MQ-H_2O was reduced to 5.65µl.

PCR	*Mdm2* *TP53* *PIK3CA*
Primer F (6pmol/µl)	0,5µl
Primer R (6pmol/µl)	0,5µl
dNTP mixture, Takara BIO inc. (2,5mM)	0,75µl
10xbuffer (m/$MgCl_2$), Qiagen	1µl
Hot Star Taq Polymerase, Qiagen	0,1µl
MQ-H_2O	6,15µl
DNA (10ng/µl)	1µl

4.4.1.1 Procedure

- A reagent mix was made according to Table 1, vortexed and spun down.
- 9µl of the mix was transferred to each separate well in a 96-well plate.
- 1µl of the DNA samples were added to the wells containing mix.
- At least one of the wells was used for a no template control (NTC), *i.e.* only containing the reagent mix and no DNA template.
- The lid was put on and the wells spun down.
- For samples with low DNA concentration 8,5 µl reagent mix and 1,5 µl of DNA samples were used

The PCR primers used in the *TP53* analysis are modified by a 5' incorporation of an oligonucleotide sequence complementary to the universal primers (-21M13 and –M13) used in the sequencing reaction. The 3' end is complementary with the gene sequence. The p53 primers are designed with help from the program OLIGO Primer Analysis Software (National Bioscience, Plymouth, MN, USA). The PIK3ca primers are designed as described in the article by Samuels Y. *et.al* 2004, Science 304:554. The Mdm2 primers are designed as described by Knappskog *et.al* (38). Information on primer sequence and fragment lengths are given in Table 2.

Table 2 The PCR- primers used in the sequence analysis of the genes *MDM2, TP53 and PIK3CA*. Forward and reverse primers are linked to universal M13 sequences in the 5' position: -21M13: TGT AAA ACG ACG GCC AGT (forward), -M13 REV: CAG GAA ACA GCT ATG ACC (reverse).

Region of interests	Forward primer (5')	Reverse primer (3')	PCR fragment length (bp)
Tp53: exon 2,3	GGAGTGCTTGGGTTGTGGT	CGGCAAGGGGGACTGTA	586
Tp53: exon 4	GACTTCCTGAAAACAACG	CACACATTAAGTGGGTAAAC	593
Tp53: exon 5,6	TTT CTT TGC TGC CGT CTTC	TTG CAC ATC TCA TGG GGT TA	588
Tp53: exon 7	GAC CAT CCT GGC TAA CGG	CAC AGG TTA AGA GGT CCC AAA	595
Tp53: exon 8,9	TTT GGG ACC TCT TAA CCT GT	CAG GCA AAG TCA TAG AAC CAT	733
Tp53: exon 10	CAT GTT GCT TTT GTA CCG TC	GGC AAG AAT GTG GTT ATA GGA	396
Tp53: exon 11	AAG GGA AGA TTA CGA GACT	TA GCT GGT ATG TCC TAC TC	500
Mdm2: P 2	CGG GAG TTC AGG GTA AAG GT	AGC AAG TCG GTG CTT ACC TG	352
PIK3CA: exon 9	GAT TGG TTC TTT CCT GTC TCT G	CCA CAA ATA TCA ATT TAC AAC CAT TG	487
PIK3CA: Exon 20	TGG GGT AAA GGG AAT CAA AAG	CCT ATG CAA TCG GTC	525

A thermal cycler DNA Engine Biorad Tetrad 2 Peltier Thermal Cycler was used to run the PCR reaction temperature profile. For activating the Hot Star Taq polymerase (Qiagen), an initiation step was performed at 95°C for 15 minutes. The PCR program was divided into three steps with repeated cycles at different temperatures and durations.

Repeated 8 times: Denaturation at 94°C for 25 seconds
Annealing at 68°C for 20 seconds
Extension at 72°C for 40 seconds
} temperature adjusted -1°C pr repetition

Repeated 10 times: Denaturation at 94°C for 25 seconds
Annealing at 60°C for 20 seconds
Extension at 72°C for 40 seconds

Repeated 12 times: Denaturation at 94°C for 25 seconds
Annealing at 58°C for 20 seconds
Extension at 72°C for 40 seconds

Repeated 12 times: Denaturation at 94°C for 25 seconds
Annealing at 56°C for 20 seconds
Extension at 72°C for 40 seconds

After the last cycle the temperature remained 4°C forever.

4.4.2 Agarose gel electrophoresis

Agarose gel (1,5%) electrophoresis was used as a semi-quantitative/qualitative check to ensure that the PCR-reactions were successful. The DNA is negatively charged and will during the electrophoresis move towards the positive electrode. Gel loading buffer 0,1% bromphenol blue (recipe in appendix B) is added to the DNA, providing color and density to the DNA, and making it easier to load. The bands are visualized by using EtBr (VWR) or GelRed Nucleic Acid Stain (Biotium) into the agarose gel. Both EtBr and GelRed bind as an intercalating agent to the DNA, and will absorb invisible UV- light and transmit its energy as visible light. The visualization was performed in SynGene GeneSnap version 7.01 (SynGene). A DNA fragment size ladder GeneRuler™DNA Ladder Mix (Fermentas) is added in one of the wells. The no template control (NTC) is used to reveal contamination of the PCR reaction mix.

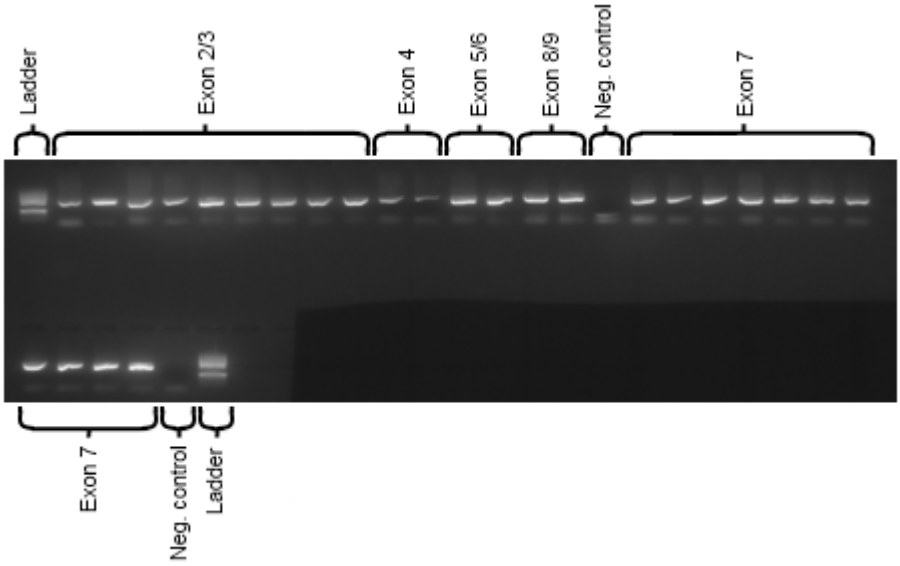

Figure 10 The gel picture visualizes PCR products from 5 different exons in *TP53* in addition to NTC and DNA fragment size ladder.

4.4.2.1 Procedure

- 5.25 gram BioRad Agarose, BioRad Laboratories was added 350ml 1xTAE buffer (recipe in appendix B) and heated in a microwave oven until fully dissolved.
- The solution was cooled down to approximately 60°C. 7 µl of GelRed was added[6] and the solution was left to polymerize in the electrophoresis chamber (about 20 minutes).
- 2µl PCR product was mixed with 2µl Gel loading buffer and all samples, the NTC and the DNA fragment size ladder was applied in separate lanes of the agarose gel
- The electrophoresis was run at 200V for 25 minutes.

4.4.3 Purifying of PCR product on epMotion 5075 VAC

The PCR products must be purified before the sequence reaction. Primer dimer, unincorporated dNTP and other pollution products can affect further analysis. This procedure was performed using the fully automated instrument epMotion 5075 VAC (Eppendorf Nordic).

[6] Ethidium Bromide (EtBr) was replaced by GelRed after the first batch of samples due to the less toxic properties of GelRed. For the first batch of samples, this step was performed in a LAF bench, since EtBr is carcinogenic.

4.4.3.1 Procedure

In the epMotion 5075 VAC the PCR products diluted in MQ-water are transferred to a MultiScreen®PCRμ96 filter plate, Millipore. By using vacuum, small polluting particles was drawn through the filter together with the water, while the DNA remained in the filter. Finally, the DNA was resolved in 30μl water and transferred to a clean 96 well plate.

4.4.4 The sequencing reaction

In the sequencing reaction, the purified PCR products were linearly amplified by cyclic sequencing. BigDye®Terminator v1.1 cycle sequencing, Applied Biosystem contains labeled dideoxy-ribonucleotides (ddNTP), which is a modified nucleotide where the OH group at the 3' positions of the ribose ring is replaced with an H atom. This prevents further elongation of the sequencing product and the chain will be terminated when ddNTP is inserted instead of a dNTP. The ddNTPs are tagged with different fluorescent dyes so that the emitted fluorescence will correspond to the base terminating the chain.

4.4.4.1 Procedure

Specifications of the Big Dye terminator sequencing reaction mix are given in Table 3.

- 7 μl sequencing reaction mix and 3μl PCR product were added to a PCR plate and quickly spun down.
- A thermal cycler Biorad Tetrad 2 was used to run the sequencing reaction temperature profile. An initialization step at 96°C for 2 minutes is required for activating the polymerase
- Denaturation at 96°C for 15 seconds
- Annealing at 50°C for 5 seconds
- Extension at 60 °C for 4 minutes

} 25 cycles

- Storage 4°C

All fragments were analyzed from both directions (forward and reverse) except p53 exon 7, where only the reverse primer was used due to poor performance with the forward primer. When sequencing Mdm2, the forward primer was not used because the SNP is located near the start of the reverse primer leading to insufficient area coverage.

Table 3 The Big Dye terminator sequencing reaction mix was made in two 1,5ml micro centrifuge tubes, one for the forward primer and another for the reverse primer. In this reaction universal primers are used as mentioned earlier. The mix should be kept on ice to avoid premature reactions.

Reaction mix pr. Sample	Volume
Big Dye Terminator reaction buffer v1.1	1,0 μl
Big Dye Terminator reaction mix v1.1	2,0 μl
Primer -21M13 or –M13 (0,8μM)	1,0 μl
Template: Purified PCR-product	3 μl
MQ-H_2O	3,0 μl
Total volume	10,0 μl

4.4.5 Purifying the sequencing products with Sephadex®

Sephadex®G-50, GE Healthcare is prepared by crosslinking dextran with epichlorohydrin. It is available as dry powder that is swollen in water before use. The swollen Sephadex® separates molecules according to their size and weight. High molecular weight substances will elute first, while small particles will be bound to the gel. The clean sequence products are collected in a 96 well plate.

4.4.5.1 Procedure

- Dry Sephadex® was loaded into a MultiScreen®HV filter plate (Millipore) with 96 wells
- 300µl MQ- water was added to each well and after 2 hours the mini-columns were swollen
- The plate was then centrifuged at 910 rpm for 5 minutes,
- 150µl MQ- water was added for pre-rinse and the plate centrifuged for 5 minutes.
- The samples were added to separate wells and the plate was centrifuged for 6 minutes at 910 rpm. The sequencing products passed through the Sephadex® and were collected in an underlying plate.

4.4.6 Applied Biosystems 3730 DNA Analyzer

The labeled sequencing products were separated according to fragment size and detected using the 48-capillary electrophoresis instrument Applied Biosystems 3730 DNA Analyzer. When the labeled DNA fragments in electrophoresis reach the detection window, a laser beam illuminates the fluorophore, causing an excitation. The emitted light is registered and separated by a spectrograph. The results were read and interpreted in the data collection software (Applied Biosystem), and the results visualized as an electropherogram.

4.4.6.1 Procedure

- The sample plate with the cleaned sequencing reaction products was sealed with a septa membrane and put into a black base plate and a corresponding retainer.
- The plate was inserted into the plate stacker of the Applied Biosystems 3730 DNA Analyzer. The results were automatically saved in the instrument operating software.

The results from the capillary electrophoresis were exported to SeqScape v2.5 (Applied Biosystem) for further analysis.

4.4.7 SeqScape v2.5

SeqScape v.2.5 is a software program used to align and analyze sequences from the Applied Biosystems 3730 DNA Analyzer. Every sequence was analyzed according to a chosen reference sequence, and the mismatching nucleotides and sequence areas were automatically highlighted. Forward and Reverse sequences are analyzed simultaneously. The four differently labeled nucleotides are visualized as colored graphs in the electropherogram, with their corresponding letter above. A heterozygote mutation is shown with two smaller peaks, (due to lower intensity caused by fewer incorporated ddNTP's) marked with different color and letter above. A homozygote mutation does not have a smaller peak, but another color on the graph and nucleotide letter.

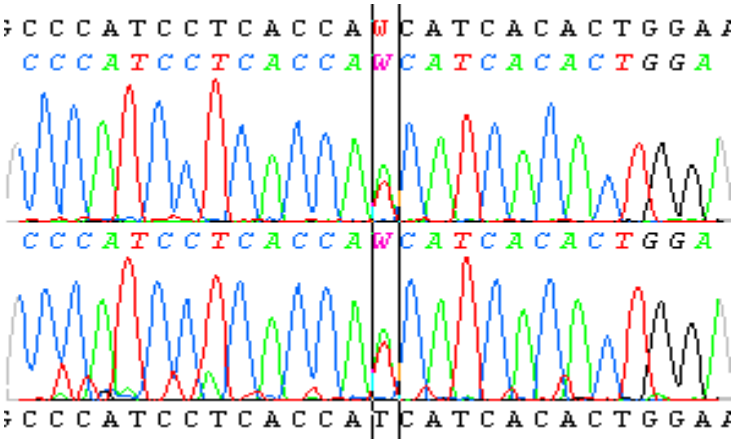

Figure 11 The electropherogram from TP53 exon 5 in sample 116 reveals a heterozygous mutation. This is shown with a red W and to similar peaks in green and red color. The sample was sequenced in two different runs, to confirm the result.

4.4.7.1 Procedure

- *Tp53, PIK3CA and MDM2* project templates and results from sequencing were imported to the computer
- The imported sequences were then analyzed and aligned according to the project template (*TP53* accession nr: NM_000546 (http://www.ncbi.nlm.nih.gov/nuccore/NM_000546), *PIK3CA* accession nr*:* NM_006218 (http://www.ncbi.nlm.nih.gov/nuccore/NM_006218), *MDM2* accession nr: NM_002392.3 (http://www.ncbi.nlm.nih.gov/nuccore/NM_002392.3)).
- All the sequences were manually and independently evaluated by two persons to assure quality control
- The sequence scoring from the two independent evaluations were compared. When a discrepancy occurred the results were discussed and potentially the fragment was re-sequenced if agreement could not be reached.

4.5 TaqMan genotyping

For detection of the SNP polymorphism in *MDM4* SNP-7 (rs1563828), a TaqMan®SNP Genotyping Assay, Applied Biosystem was used. The TaqMan® Genomic Assays are allelic discrimination assays exploiting the 5'-exonuclease activity of *Taq* polymerase. A PCR with one TaqMan® probe for each SNP allele is performed in a Real-Time PCR instrument like 7900HT Fast Real-Time PCR System and the fluorescence signals generated during the PCR reactions are determined by laser detection. TaqMan®probes are sequence specific oligonucleotides with a fluorescent reporter dye linked to its 5' end and a minor groove binding non-fluorescent quencher (MGBNFQ) in its 3' end. The most widely used reporter dyes are VIC® for allele 1 and FAM® for allele 2. The minor groove binding part of the quencher stabilizes the probe-template complex and allows the design of shorter probes, which make the allelic discrimination more robust.

The AmpliTaq Gold DNA polymerase extends the primers bound to the DNA template, and its 5'-exonuclease activity cleaves the probes that are hybridized perfectly to the target. The cleavage separates the reporter dye from the quencher, which results in a fluorescence signal. The signal indicates which allele is present in the sample.

Homozygosity for allele 1 is identified by a VIC-dye signal. Homozygosity for allele 2 is identified by a FAM-dye signal. If fluorescence signal from both VIC- and FAM-dye are detected, it indicates heterozygosity.

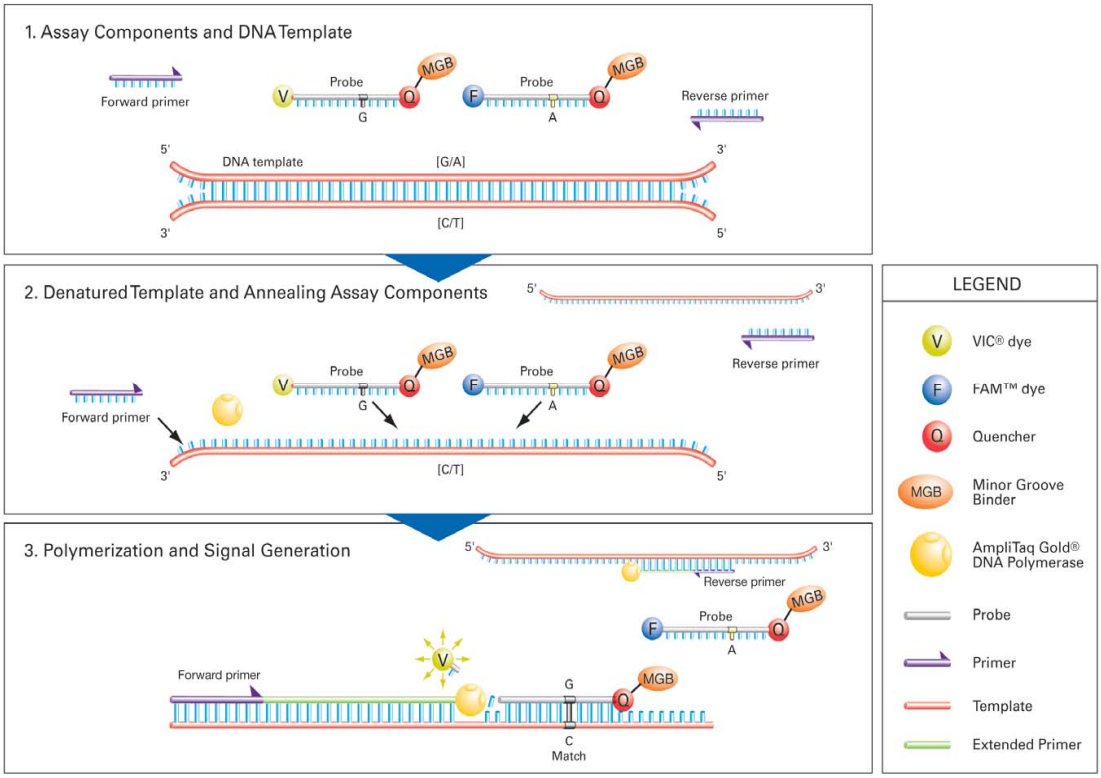

Figure 12 SNP detection by TaqMan® technology. The PCR primers and TaqMan®probes are annealed to the denatured DNA templates. The probes have a fluorescent reporter dye linked to its 5' end and a minor groove binding non-fluorescent quencher (MGBNFQ) linked to its 3' end. The fluorescence signal is generated during the polymerization reaction (source TaqMan®SNP Genotyping Assays Protocol).

4.5.1 Procedure

This analysis was performed in a 384-well plate. The TaqMan® SNP Genotyping assay consisted of primers and probes. In addition, TaqMan®Genotyping Master Mix was used.

TaqMan reaction mix for 45 samples:

- 109,7µl of the TaqMan®Genotyping Master Mix
- 2,8µl of the TaqMan® SNP Genotyping assay. The solution was vortexed and spun down.

The primer sequences used in this analyze is confidential. Probe 1 was tagged with Vic dye and probe 2 was tagged with FAM dye.

Following was mixed and added to each well:

- 2,5µl of reaction mix
- 2,5µl of the sample (10ng/µl)

An optical adhesive film MicroAmp™ (Applied Biosystem) was used as a cover, and the plate was quickly spun.

The PCR was initiated with a DNA denaturation/polymerase activation step at 95°C for 15 minutes, followed by 40 cycles consisting of following steps: 95°C for 10 minutes (denaturation), and 60°C for 1 minute (annealing and extension).

The fluorescence signals were analyzed by the software, and the results were shown by clustering of the alleles in a scatter plot depending of the signal strength from the probes detected. Each plotted point corresponds thereby to the fluorescence values from both dyes. Alleles with weak signal were not clustered but placed in bottom of the diagram and shown as not confirmed.

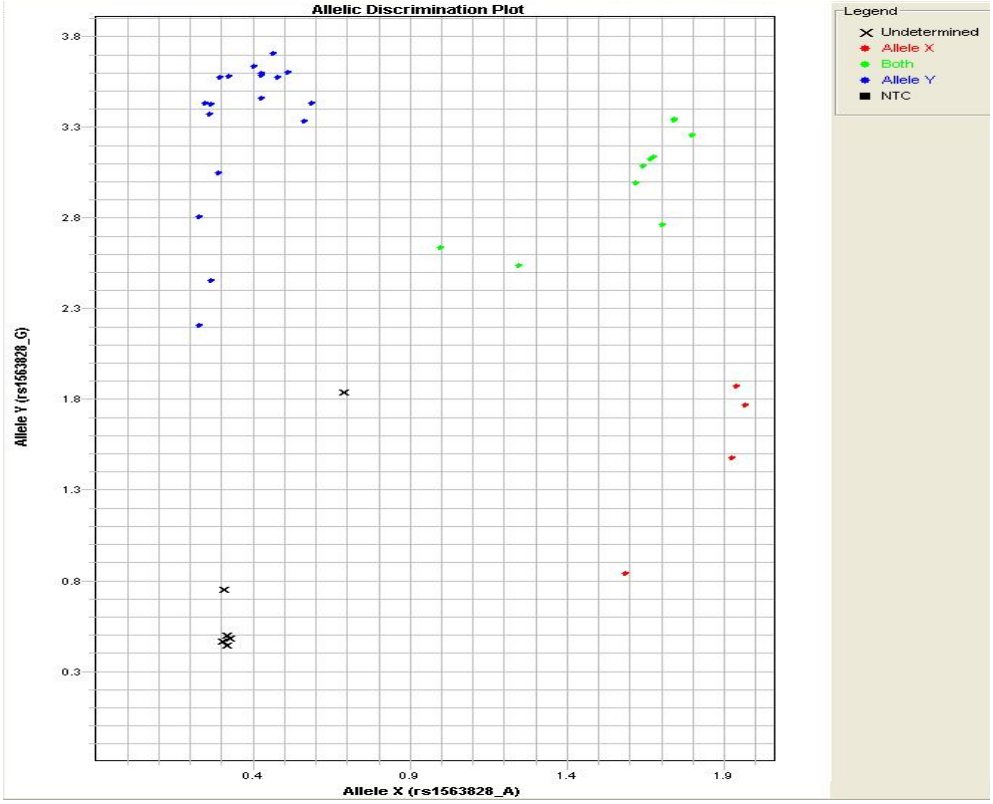

Figure 13 Allelic discrimination plot. The red spots are allele A/A, the blue spots indicate G/G and green spots indicate both alleles. The black crosses are samples not determined.

4.6 Methylation

Methylation of DNA occurs on cytosine residues, especially on CpG dinucleotides. For determination of the methylation pattern, it's necessary to distinguish between the methylated cytosine and the unmethylated cytosine. This is solved by treating the DNA with sodium bisulfate. In this step, unmethylated cytosine will convert to uracil, while methylated cytosines remains unchanged as shown in Table 4.

Table 4 Bisulphite treatment of DNA results in conversion of unmethylated cytosines to uracil, while methylated cytosines remain unchanged.

	Original sequence	After bisulphite treatment
Unmethylated DNA	N-C-G-N-C-G-N-C-G-N	N-U-G-N-U-G-N-U-G-N
Methylated DNA	N-**C**-G-N-**C**-G-N-**C**-G-N	N-**C**-G-N-**C**-G-N-**C**-G-N

The methylation analysis was performed using Illumina Infinium HumanMethylation27 BeadChip Kits. This is a high throughput genome-wide array covering 27,578 CpG cites, mainly located in promoter regions and in the gene transcription area (69). The methylation process is divided into the following steps: denaturation, fragmentation, precipitation, re-suspension, hybridization and staining of DNA.

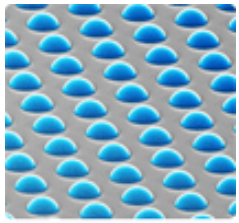

Figure 14 The BeadChips are covered with 3-micron silica beads assembled in micro-wells.

The BeadChips consist of 3-micron silica beads assembled in micro-wells. Each bead is covered with specific oligonucleotides which will bind sample DNA fragments with perfect match. The degree of methylation for each locus is analyzed using two different beads types, one with oligonucleotides complementary to the methylated sequence and one with oligonucleotides complementary to the unmethylated sequence. In a single base extension reaction, the oligonucleotide is extended with the fluorescently labeled base complementary to the sequence in the hybridized DNA strand (Figure 15). The signal from the labeled base is intensified using a multi-layer staining procedure prior to laser detection using the Illumina BeadArray Reader. The results from each bead on the array are analyzed in Illumina Genome Studio Methylation Module v1.8.The ratio between the signals from the two probes is calculated and gives information about the methylation beta value. This value is in the range between 0-1 and can be expressed with this formula:

$$\beta = \frac{Max(SignalB,0)}{Max(SignalA,0) + Max(SignalB,0) + 100}$$

Signal A corresponds to the signal in Red channel and signal B corresponds to the signal in the Green channel. The 100 is added to the denominator just to ensure that the Beta value does not become a negative value in the rare case that max(signalA,0) + max(signalB,0) is less than zero. Information about each bead type ID used, the expected intensity and color channel are listed in table 20 in the methylation protocol guide (Part#11322371 Rev.A). Illumina Infinium Methylation Assay system contains various controls for evaluating the quality of the procedure, both sample independent and, sample specific.

The sample independent controls are used for evaluating the different steps in the methylation process, and include staining controls which examine the staining efficiency, the extension controls which test the extension efficiency of the nucleotides, target removal controls which evaluate the stripping step and hybridization controls which monitor the overall performance of the assay.

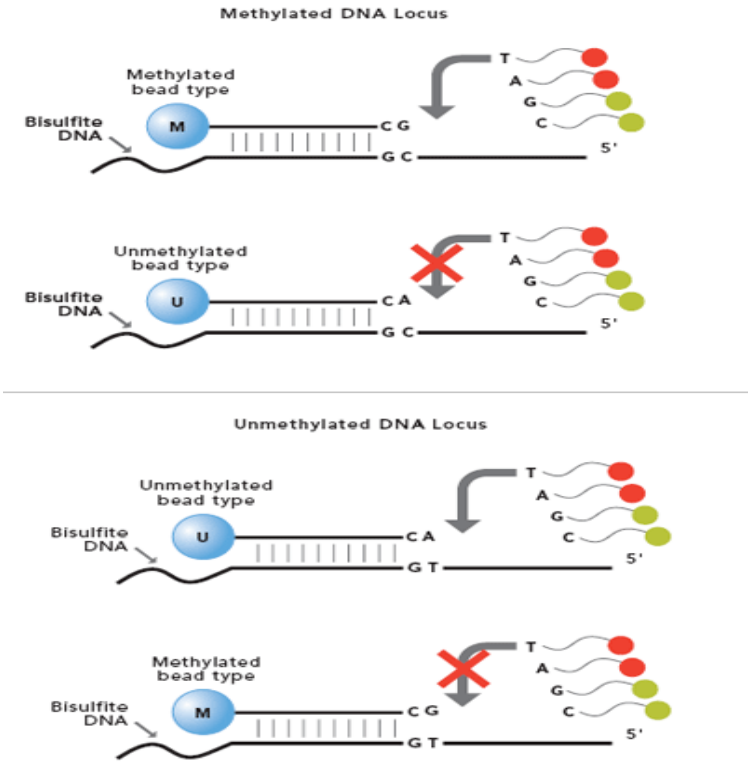

Figure 15 Two different bead types are used in the methylation array. Top figure: the M probe will fit the methylated CpG target site enabling single-base extension and detection. Extension will be inhibited when a single-base mismatch occurs. Bottom figure: shows the U probe, which will match with an unmethylated CpG site and M probe which will mismatch the unmethylated DNA locus.

The sample dependent controls are designed to evaluate performance across the samples. Bisulphite-conversion controls- test the efficiency of the bisulphite conversion of genomic DNA (gDNA). A successful conversion will let the converted probes match the converted sequences, get extended and vice versa. The specificity controls are designed for detecting perfect match giving high signal or mismatch leading to a low signal. Negative controls-consist of sequences that should not hybridize to the DNA template. Mean signal from these controls defines the system background and is also used for detecting cross-hybridization, non-specific extension and imaging system background. Negative controls should be monitored both in green and red channels. The non-polymorphic controls (NP) are designed to control the process from amplification to detection.

4.6.1 Bisulphite treatment

The bisulphite treatment and clean up procedure were performed using EpiTect® Bisulphite Kit 48 (Qiagen). This is the most critical step in the methylation analyzing process. The DNA is treated with high bisulfate salt concentrations at high temperature and low pH. The low pH is necessary for optimal cytosine conversion. It is important that the denaturation is performed at high temperature, since the conversion reagents only work on single-stranded DNA (the different steps are shown in Table 6). (Source: EpiTect® Bisulphite Handbook, QIAGEN 09/2009) These hard conditions usually lead to a high degree of DNA fragmentation and loss of DNA.

4.6.1.1 Procedure

All the tumor samples were diluted with RNase free water in their respective tubes, to an optimal amount of 500ng with a total volume of maximum 40µl. Some of the sample had low concentration of DNA and were not diluted.

Table 5 Reagents and volume needed per reaction.

Component	Volume per reaction
DNA solution (1-500ng)	Variable (max 40µl)
RNase-free water	Variable
Bisulphite mix	85µl
DNA protect buffer	15µl

- The bisulphite mix was dissolved by adding 800µl RNase-free water. For completely dissolving, 5 minutes with vortexing was necessary. One aliquot is sufficient for 8 conversions, and for 55 samples, 7 aliquots were needed
- Samples and reagents were transferred to a 96-well PCR plate
- The plate was closed with lids, vortexed easily and spun quickly
- The thermal cycling process was performed on BioRad TetRad 2, and the instrument was programmed with denaturation and incubation steps as described in Table 6.

Table 6 Thermal cycler conditions during the bisulphite conversion.

Step	Time	Temperature
Denaturation	5 min	95°C
Incubation	25 min	60°C
Denaturation	5 min	95°C
Incubation	85 min	60°C
Denaturation	5 min	95°C
Incubation	175 min	60°C
Hold	Indefinite	20°C

The denaturation steps are important for keeping the DNA single-stranded, and the incubation steps are necessary for the sulfonation and cytosine deamination.

4.6.2 Cleanup procedure

For purification of the DNA and removing desulfonation agents, a cleanup procedure is performed. During the washing process, the desulfonation agents are removed, while the single-stranded DNA is bound to the membrane in the column. In the last step, pure converted DNA can be eluted from the spin column by adding an elution buffer.

4.6.2.1 Procedure

This procedure was performed according to recommendations in EpiTect bisulphite handbook. Before starting the process, the buffers had to be prepared. The BW buffer (washing buffer for removal of desulfonation agents) was added 30ml ethanol (96-100%), the BD buffer (for desulfonation) was resolved in 27ml ethanol (96-100%) and the BL

buffer was added carrier RNA (310µg/µl). The BL buffer promotes binding of ssDNA to the EpiTect spin column, which is further enhanced by the carrier RNA.

- Each sample was transferred to a clean 1.5ml micro centrifuge tube, added 560µl of BL buffer, vortexed and centrifuged briefly
- The solutions were dispensed to the EpiTect spin columns with corresponding collection tube underneath, and centrifuged at maximum speed for 1 minute. The filtrate was discarded
- 500µl BW buffer was added to each spin column, centrifuged and the waste were discarded
- 500µl BD buffer were added, and the samples incubated for 15 minutes. The spin columns were then centrifuged, and the waste was discarded.
- 500µl BW buffer were added, centrifuged and the waste discarded. This step was performed twice. After the last centrifugation, new clean collection tubes were placed under the spin columns and centrifuged.
- To remove the rest of the ethanol, the tubes were placed in a heating block for 5 minutes at 56 °C for evaporating.
- Clean 1.5ml micro centrifuge tubes were placed under the spin column and 20 µl EB buffer were added each column. By centrifuging the samples at 15 000 x g for 1 minute, the DNA was eluted into the micro centrifuge tubes. This step was repeated to increase the yield of DNA in the eluate. The samples were stored at -20°C before used in the methylation analyze.

4.6.3 Quantification of bisulphite converted DNA with Real time PCR

The bisulphite analyze is a rough procedure that fragments and converts the DNA. In this process over 90% of the DNA can be lost (70). To make sure that the samples contain sufficient DNA for the methylation arrays, the amount of converted DNA was measured with Real Time PCR. The TaqMan probe principal is explained in chapter 4.5. Real Time PCR with specific primers and TaqMan probes is a sensitive method that can detect small amount of DNA. The method is also well suited to analyze fragmented DNA since the amplified sequence is short (50-150 base pairs). A threshold value is set automatically in the PCR reaction exponential phase, above the intensity of background noise. For absolutely quantification, a standard curve is required. A standard curve with 9 point diluted standards (EpiTect®Control DNA, methylated (100), Qiagen) in the range 0 – 5 ng/ µl) are set up together with the samples, and the software program (SDS v.2.3, Applied Biosystem) will calculate a standard curve based on the Ct values. The Ct value is defined as the number of cycles needed to reach the chosen threshold value and is inversely proportional with the amount of DNA in the samples. The analysis software automatically read out the concentration of bisulphate converted DNA from the standard curve.

4.6.3.1 Procedure

An oligomix containing 6µM of each primer and 2µM of the probe(in Table 7) was prepared according to description in MethyLight (71).

Table 7 Custom designed Primers and probe (Applied Biosystem) used in the Real-time PCR concentration analysis of bisulphite converted DNA.

Primer/probe	Sequence
ALU-C4M_F	GGT TAG GTA TAG TGG TTT ATA TTT GTA ATT T
ALU-C4M_R	ATT AAC TAA ACT AAT CTT AAA CTC CTA ACC T
ALU-C4M_P	FAM- CCT ACC TTA ACC TCC C - MGBNFQ

- A reagent mix with master mix, oligomix and MQ water was made as shown in Table 8.
- 9,5µl of the reagent mix was transferred to each well in a 384 well plate.
- 2 NTCs and each sample were added as triplets to the plate into their respective wells. The 9 standards were added in triplets.

Table 8 Reagents used for the PCR reaction mix for 1 sample

Reagent	µl/well
2xTaqMan®Universal PCR Master Mix, No AmpErase®Ung (Applied Biosystem)	5
Oligomix (6µM ALU-C4M_F, 6µM ALU-C4M_R, 2µM ALU-C4M_P)	3
MQ	1,5
Bisulphite converted DNA	0,5
Total volume	10

The real-time PCR program:

- Incubation on 95°C in 10 minutes for activation of the polymerase.

40 cycles with following steps:

- Denaturation at 95°C for 15 seconds
- Annealing and extension at 60 °C for 1 minute

The threshold line and base line were set automatically. The results from this analysis were evaluated and gave an indication of whether the samples could be used for further analysis.

4.6.4 Denaturation, amplification and fragmentation

In this step the bisulphite treated DNA is chemically denatured before amplified and fragmented using Illumina supplied reagents. Because of low concentration, all of the bisulphite treated DNA (30-40µl) was added instead of the recommended 4µl. This was compensated by using 30µl 0.1M NaOH instead of 4µl proposed by the methylation

protocol guide (Illumina, Part#11322371 Rev.A). All other steps were performed according to methylation protocol guide.

4.6.4.1 Procedure

A midi plate with 96 deep wells was used

- 20 µl Multi-Sample Amplification 1 Mix (MA1) was transferred into each well
- All the bisulphite treated DNA (30-40µl) were added the wells
- 30µl 0.1M NaOH were added to each well to denature the DNA
- 68µl of the buffer Multi-Sample Amplification 2 Mix (MA2) were added each well for neutralizing the DNA
- 75µl of the multi-sample amplification master mix (MSM) were dispensed into each well for DNA amplification
- The MIDI plate was incubated at 37°C for 22 hours for amplification.
- The DNA was then fragmented by the enzyme FMS, during a 1 hour incubation at 37°C.

4.6.5 Precipitation and resuspending

For precipitation of the DNA, 100% 2-propanol and PM1 were used.

- 100µl with PM1 was added to each well in the MSA2 plate and incubated at 37 °C for 5 minutes
- 300µl 100% 2-propanol was dispensed into each well and the samples incubated at 4°C for 30 minutes,
- The MSA2 plate were centrifuged at 3000xg and 4°C for 20 minutes, causing the DNA to precipitate to blue pellets attached to the bottom of the well.
- The supernatant was removed by inverting the plate onto an absorbent pad. The pellets were air dried in room temperature by leaving them uncovered in the inverted plate for 1 hour, showed in Figure 16. It is important that the alcohol is completely removed from the DNA, because of its tendency to interfere with further analyzes. The plate was sealed with a cap mat and stored at -20°C until the next day.

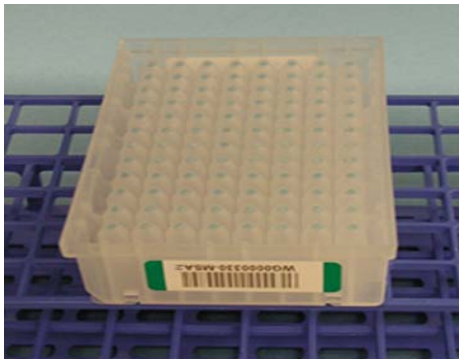

Figure 16 An uncovered MSA2 Plate Inverted for Air Drying. The blue pellets are shown in the bottom of the wells

The precipitated DNA was resuspended in 42µl Resuspension, Hybridization, and Wash Solution (RA1). The MSA2 plate was sealed and placed in an Illumina Hybridization Oven and incubated for 1 hour at 48°C. Then the plate was vortexed at 1800 rpm for 1 minute and centrifuged at 280g for 1 minute. The vortexing and centrifugation step was repeated once. The plate was put in a -80°C freezer, which is required when storing more than 24 hours.

4.6.6 Hybridization

The hybridizing of DNA to the BeadChips is performed in Hyb Chambers shown in Figure 17. The different Hyb Chamber components were put together according to the protocol and the Hyb Chamber reservoirs were filled with 200µl Humidifying Buffer (PB2). Then the Chamber was closed and locked until the BeadChips were loaded.

Figure 17 shows the BeadChip Hyb Chamber components. A: Hyb Chamber, B: Hyb Chamber Gasket, C: Hyb Chamber Inserts.

The BeadChips were placed into a Hyb Chamber insert, and 12 µl of the fragmented and resuspended DNA were loaded to the BeadChips according to a lab tracking form (Part# 11327244 Rev.B). When all the samples were loaded, the Hyb Chamber inserts were placed into the Hyb Chamber and the Hyb Chamber was placed into a 48°C Illumina Hybridization Oven. The incubation time is recommended to be at least 16 hours but no more than 24 hours. To achieve optimal distribution of the samples to the beads, a rocker function was used.

Figure 18 shows a BeadChip where 6 samples can be loaded with a multi-channel precision pipette into the left side of the BeadChips, and 6 samples into the right side according to a loading protocol. On the top of the BeadChip a unique Barcode is placed.

4.6.7 Washing

The washing process was performed by the core facility lab at Radiumhospitalet, Institute for Cancer Research. The washing process is performed to remove residues from the hybridization step. The cover seals must be carefully removed without touching the BeadChips, and then the BeadChips were immediately slided into the wash rack which was filled with washing buffer (WB1). The rack was moved up and down for 1 minute and then removed to next wash dish containing Hybridization Preparation Reagent (PB1). The rack was moved up and down for 1 minute. After the washing process, 4 black frames are placed in a Multi Sample BeadChip Alignment Fixture filled with PB1. Each BeadChip can then be placed into the respective black frame, according to the protocol. A plastic spacer was put onto the BeadChips and an Alignment Bar placed onto the Alignment Fixture. A dust-free glass back plate was then placed onto the BeadChips. Two Metal clamps were attached to each BeadChip, which were placed in a Flow-Through Chamber for single base extension.

4.6.8 Single-Base Extension and Staining

The single base extension and staining steps were performed in a liquid handling system, Tecan Freedom Evo. Unhybridized and non- specifically hybridized DNA were first removed by RA1 reagent. XStain BeadChip Solution 1 (XC1) and XStain BeadChip Solution 2 (XC2) are required for the extension reaction, and Two-Color Extension Master Mix (TEM) reagents are used for the single base extension reaction. In this process labeled nucleotides are extending the oligonucleotides on the beads. The hybridized DNA is removed with 95% formamide/1mM EDTA (recipe in appendix A). The XStain BeadChip Solution 3 (XC3) reagent neutralizes the single base extended oligonucleotides, which now can undergo a multi-layer staining process. After finished the staining, the BeadChips were washed in PB1 reagent, coated with XStain BeadChip Solution 4 (XC4) reagent and then dried.

4.6.8.1 Procedure

The liquid handling system was prepared according to the Infinium II Methylation Assay Automated Protocol (Part# 11322195, RevC). A chamber rack was inserted on Robot Bed and the reagents (RA1, XC1, XC2, TEM, 95% formamide/1mM EDTA and XC3) filled in respective reservoirs. When the process was completed, the BeadChips were removed from the Chamber Rack, and placed on the bench in room temperature. The Flow-Through Chambers were carefully disassembled and the BeadChips were placed in a clean staining rack. The staining rack was moved up and down 10 times in a dish chamber filled with PB1 to wash the BeadChips. The staining rack was soaked for 5 minutes and then transferred to a dish chamber containing XC4. The same procedure was performed in this step. The BeadChips were then removed from the rack and placed horizontally on a tube rack, which were placed in a dessicator for drying. The dessicator dries the BeadChips with use of vacuum. The clean and dry BeadChips are scanned in an Illumina BeadArray scanner.

4.6.9 Image BeadChip iScan System

The scanning process was also performed by the core facility lab at Radiumhospitalet according to the Updated Hyb, Wash, and Image Protocols for the Universal-12 BeadChip (Illumina, Part#11325425 Rev.A). The iScan Control software is automatically connected to the iScan Reader. The fluorophore incorporated in the single base extension reaction is excited with a laser, and the emitted light are detected and recorded in high-resolution images of the BeadChip sections. Data from these images are then analyzed by the corresponding iScan Control software.

4.6.9.1 Procedure

The four BeadChips were loaded into the iScan Reader tray and the barcodes were scanned. The scanning process took 8-10 minutes per BeadChip. The image is saved as a JPEG file or a TIFF file. After scanning, every bead was registered, meaning that location on the bead was matched to information in the bead map file. This file is unique for each BeadChip array. The intensity values were extracted and determined for every bead on the image. This information was stored in an intensity data file.

4.6.10 Data Preprocessing

This data preprocessing procedure was performed by Jörg Tost and his colleagues, a collaborating epigenetic group at Centre Nationale de Genotypage (CNG) in Paris. This process is a quality control where the dataset is cleaned, normalized and prepared for further analysis. Targets that contained a 'zero' value in one sample for methylated or unmethylated signals were removed (n=10). In intra-sample normalization, color bias and background levels were corrected for. An imbalance between the two color channels may occur because the labeling efficiency is different for the two dyes. Different scanning properties may also lead to different intensities measured in the two color channel. A smoothing of the quantile normalization was used. When using Infinium methylation arrays a large number of CpG sites are probably unmethylated. This leads to a mode of intensities measured by the methylated probes. The background level is due to this mode position, and must be corrected.

For in between sample normalization, quantile normalization is used. This was performed at probe level, meaning that the intensities of methylated and unmethylated probes are normalized, instead of normalizing the summarized methylation levels. In quantile normalization, all the beta-values are sorted according to intensity and then matched with similar values. In that way the smallest values are identical, the second values are identical, and so forth. This is independent of the gene identity, meaning that the smallest value for one array may represent a different gene than the smallest on another array.

4.7 Bioinformatics and statistics

Due to the high amounts of probes, different bioinformatic tools which could handle the huge dataset were used.

4.7.1 Prediction Analysis of Microarrays (PAM)

PAM analysis is a prediction analysis for microarrays for sample classification, using the nearest shrunken centroid method. A normal distribution of the data set is not required for this analysis. The dataset is divided into different classes or groups, such as phenotypes, by the user. According to these classes a significant gene list which separates the classes best can be identified using PAM. PAM performs a cross validation with a range of threshold values for separating the predefined classes, and a misclassification error is connected to each threshold value. The number of genes in the gene list can be changed by using different thresholds. A higher threshold provides a shorter gene list. This analysis can handle large set of data, and will come out with a significant gene list, depending of the classes you want to separate. This significant list of genes can be used in further bioinformatic tools (72).

PAM analysis uses the method nearest shrunken centroid, which computes a standardized centroid for each class. The centroid is the average expression/methylation for each gene in each class, divided by the with-in class standard deviation for that gene. A new sample will be compared with each of the class centroids, and put into the class whose centroid it is closest to in squared distance. The class centroids are modified by shrinking the centroid towards zero by using a threshold. This modification makes the method more accurate by reducing effect of “noisy” genes (72).

4.7.1.1 Procedure

The dataset was loaded into the program, and cross validation of the set of genes according to the predefined classes, were performed by the program. The threshold values with misclassification error rate, together with the associated number of genes, were illustrated by the use of a diagram. A threshold value according to the diagram was chosen, which next gave an indication of the overall error rate. A confusion table with the number of samples misclassified according to the predefined groups was depicted. Different plot showing the distribution of genes in the predefined groups were available, and a significant list of genes could be extracted and used for further analysis.

4.7.2 Hierarchical clustering in J-express

J-express is a bioinformatic and statistical data analysis program, owned by Molmine AS and developed by bioinformatics group at the Department of Informatics, University of Bergen, Norway. Different statistical analysis and methods are available in the software program, but only the hierarchical cluster method was used in the analysis included in this thesis.

The clustering segments objects into clusters, such that those within a cluster are more likely related to one another than the objects in different clusters. This can be performed supervised, where the groups of interest are predefined, or unsupervised, where no predefining of groups is provided. By using the supervised clustering, genes and groups of genes differently methylated between the predefined groups can be imaged. It should be noticed that the supervised clustering only is a visualization of the gene list imported to the program and not an independent

statistical test. The unsupervised clustering is performed for revealing underlying biological mechanisms differently regulated independent of predefined groups. There are options for single, complete and average linkage of the clusters. Average linkage is based on the average distance between two clusters. The results are visible in a dendrogram, where the samples are branched into subsets based on similarity in methylation pattern. The gene names imported to the program are listed up at the right side of the dendrogram (73). The statistic tool Spearman Rank Correlation was chosen in order do a comparison of the data from the two groups of interests. The data are then converted into rank ordering before the analysis is performed.

4.7.2.1 Procedure

The gene list provided in PAM analysis was loaded into J-express for visualization. Columns containing gene identifiers, rows with sample id's, and the methylation values were marked. First, a supervised hierarchical clustering for the different subset of data was chosen. As statistical tool, Spearman rank correlation and average linkage were chosen. The results were visualized by a dendrogram, showing both clustering of the correlated samples and genes.

4.7.3 Ingenuity Pathway Analysis

Data were further analyzed through the use of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com). This is a licensed software program with a lot of possibilities for pathway and gene analysis. The program provides information about the interactions with other genes, cellular phenotypes and disease processes in the experimental data. Biological models can be made, using search in the Ingenuity® Knowledge Base. The canonical pathways are curated from the scientific literature and provides with p-value, false discovery rate (FDR) and ratio. The p-value is calculated from Fisher's exact test, and measures if the association of the pathways related to your dataset can be explained by chance alone. The false discovery rate (FDR) indicates the proportion of falsely rejected null hypothesis (Type I error), i.e. false positives. Calculating the FDR is performed by using Benjamini-Hochberg multiple testing correcting p-values, and will give an indication of how many false positives you maximally can expect among the significant findings. The ratio is the number of molecules from your dataset that map to the canonical pathway, divided by the total number of molecules that map the canonical pathways displayed (74).

4.7.3.1 Procedure

The gene list provided from PAM analysis was imported into the program for identification of which pathways that were most significant to the imported data set. A core analysis of the data set was performed and a summary of the analysis was provided. The summary included top networks, top bio functions, top canonical pathways, top tox list and top tox functions in the imported gene list.

The top canonical pathways were then further examined for relevance to the project's purpose.

4.7.4 Chi-square and Fisher's exact test

Chi-square is used to consider a relationship between data in cross-tabulation on nominal level. If more than 1/5 of the cells have less than five observations, Chi-square is not recommended. Fisher's exact test is then preferred. Chi-square or Fisher's exact test (when appropriate) was calculated to get an indication if the variables in treatment response and development of side effects could be explained by carrying different polymorphisms. Fisher's exact test was also used to see if there was a relationship between ER status and abbreviations in p53/p53 pathway. These tests were performed using the software program IBM SPSS Statistics 19.

4.7.5 Student's t-test

Student t-test is a parametric test used for normal distributed data. For my data, two tailed with two sample equal variance were used for comparing average methylation between different groups. The most significant genes separating the groups of interests were analyzed with Student's t-test. This was performed in order to check if these genes are statistically different methylated in the different groups. The cut-off is set to 0.05, indicating that any p-value <0.05 is treated as significant. A low p value indicates that the difference in methylation cannot be explained by chance, but that the genes are actually differentially methylated between the groups (75).

4.7.6 Mann Whitney test

This test is performed in IBM SPSS Statistics 19. This is a non parametric test for comparison of two groups, and does not require normal distribution of the data. The test performs of sums of ranks and all the observations are therefore ranked. The cut-off is set to 0.05 (75).

5 RESULTS

This study was designed to investigate epigenetic alterations as a result of radiation exposure and to identify biomarkers related to response to radiation. Also, genetic variation and somatic mutations in genes reported to play a role in the radiation response were analyzed.

5.1 DNA extraction and concentration measurement

The concentration and purity of the DNA extracted from biopsies were measured both with absorbance on NanoDrop®ND-1000 and by fluorescence using the fluorochrome PicoGreen on NanoDrop®3300. The DNA concentration measurements varied between the two methods. In Figure 19 the result from the two analyses are shown.

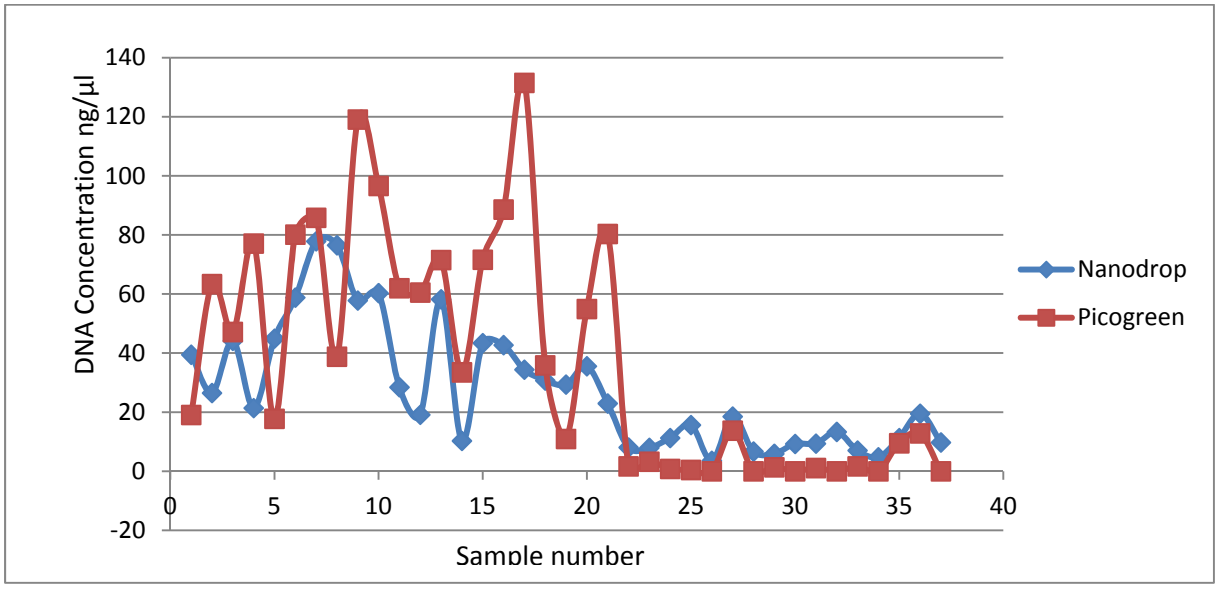

Figure 19 DNA concentrations for each sample measured with both fluorescence (NanoDrop®3300 with PicoGreen) and absorbance (NanoDrop®ND-1000).

In the low concentration areas, the fluorescence method tended to measure lower concentrations than the absorbance method while the opposite was found in high concentration areas. The largest difference between the DNA concentrations measured by the two methods was detected in high concentration areas.

5.2 Clinical data

Clinical data like response to radiation and estrogen receptor (ER) status were available for 18 of the 19 samples. The response to radiation was evaluated according to EORTC scoring system based on alterations in the tumor size, leading to four different categories of response (good response, partial response, no response and progression). Regarding ER status the samples were divided into ER negative and ER positive, where those with ≥10% of the cell nucleus stained were classified as ER positive. The distribution of the samples in the four response group and ER

groups are listed in Table 9. Data on other clinical factors such as Her2 and progesterone receptor status, tumor size and lymph node status was also available but were not used in the analysis included in this thesis.

Table 9 Clinical data such as ER status and response were available for the samples analyzed in this thesis. The different tumor samples are shown with recorded ER status and response to radiation. In addition, the frequencies of each ER group and response class as well as missing values (ND) are displayed in the bottom row.

Sample ID	ER	Response
101	Neg	Good
102	ND	Good
103	Pos	No Response
104	Neg	Good
105	Neg	Good
106	Pos	No Response
107	Pos	Good
108	ND	ND
109	Neg	Good
112	Pos	Partial
113	Pos	No Response
114	Pos	Good
117	Neg	Progression
118	Neg	ND
120	Pos	No Response
121	Pos	Partial
123	Neg	Good
124	Pos	Good
125	Pos	Good
127	Neg	Partial
Total	Pos: 50% Neg: 40% Missing(ND): 10%	Good: 50% Partial: 15% No Response: 20% Progression: 5% Missing(ND): 10%

5.3 Methylation analysis

A major part of this project was the investigation of methylation pattern in normal tissue, non-radiated tumor tissue and radiated tumor tissue to understand more of the molecular mechanisms underlying the radiation response. The methylation analysis can be divided into three main parts: 1) normal tissue versus tumor tissue, 2) tumor tissue before versus after irradiation and 3) methylation and response.

5.3.1 Average methylation in normal and tumor tissue

The average methylation level in normal tissue, non-radiated and radiated tumor tissue was 29,5%, 29,9% and 30% respectively with similar standard deviation in all three tissue types.

5.3.2 Unsupervised Clustering of normal and tumor samples

An unsupervised hierarchical clustering based on the raw beta values was performed by a collaborating epigenetic group at Centre Nationale de Genotypage (CNG) in Paris. In this cluster analysis the samples are clustered without any predefined parameters, by using Euclidian distance and the Ward agglomerative method. The cluster analysis separated normal samples from tumor samples, but not radiated tumor samples from non-radiated tumor samples (Figure 20). For eight of the samples the before and after samples clustered together (denoted "paired samples" for the remaining part of this thesis).

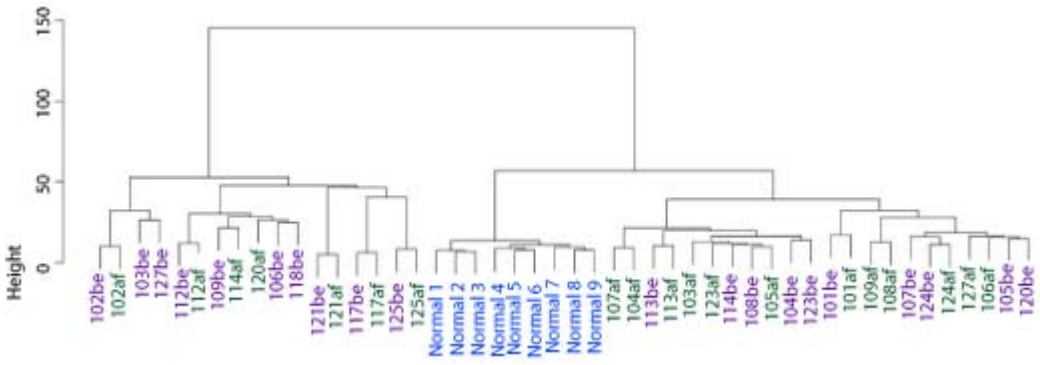

Figure 20 Cluster dendrogram provided from unsupervised hierarchical clustering, based on raw beta values. Samples clustering together share more similarities in methylation level across the whole dataset containing 27578 probes than those separated into two clusters. The normal samples are marked with blue color and are clustered in a separate sub-cluster. Non-radiated and radiated tumor samples are indicated with purple and green respectively.

5.3.3 Methylation pattern in tumor samples and normal samples

The normal samples (healthy control samples) and tumor samples taken before radiation (termed as before samples), were analyzed to uncover differences in methylation pattern between normal breast tissue and tumor tissue.

5.3.3.1 PAM analysis of differentially methylated genes in normal samples and tumor samples

The normalized methylation data were analyzed using prediction analysis of microarray (PAM) for detecting genes differentially methylated between normal tissue and non-radiated tumor tissue. The misclassification error is dependent on the number of genes included in the classifier at a given threshold (Figure 21). Here, a threshold higher than 5.8 resulted in 1 tumor sample misclassified as a normal sample. At lower thresholds the number of genes increases with an increasing misclassification error. A threshold of 5.8 was chosen as it separated best between the two groups. At this threshold 14 genes were included in the classifier (Table 19 in Appendix A). Despite a short gene list, the overall misclassification error was zero.

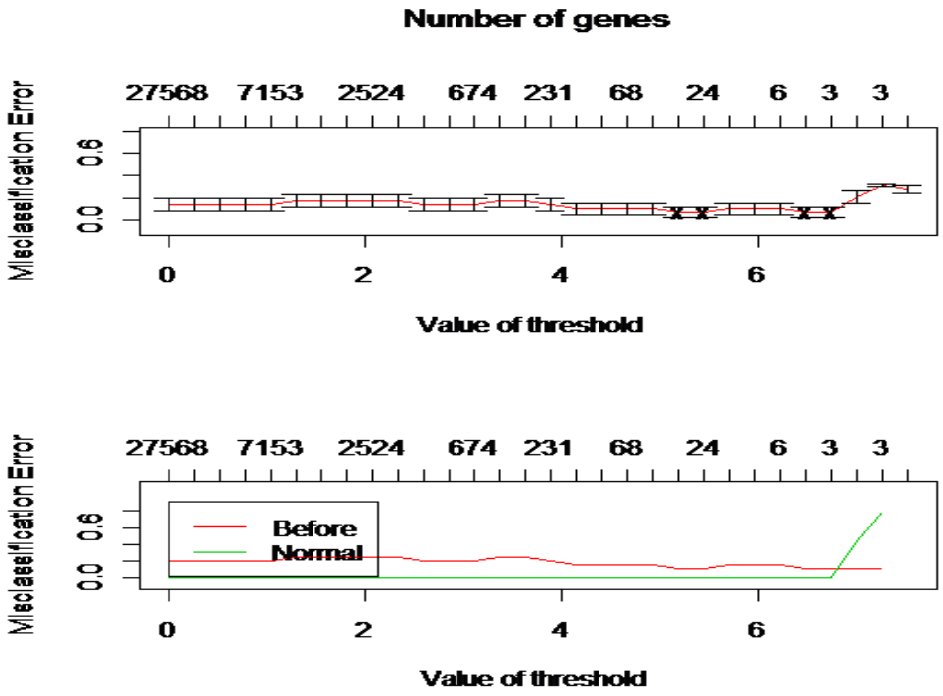

Figure 21 The misclassification error figure gives an indication of the number of genes included in the classifier used to score the samples into the groups of interest. A lower threshold results in more genes and a risk of over fitting. The y- axis gives the misclassification error and the x-axis the threshold and corresponding number of genes included. In the comparison of normal versus non-irradiated tumor tissue, a threshold of 5.8 was chosen, at this threshold 14 genes were included in the classifier.

5.3.3.2 Supervised hierarchical clustering of normal samples and non-radiated tumor samples

The list of 14 genes provided from PAM was imported into J-express for visualization. Using supervised hierarchical clustering illustrates further the potential of these genes to separate the samples according to the phenotype of interest. Spearman Rank Correlation and average linkage was used in order to show the degree of correlation between the methylation levels of the different genes selected from PAM. In the resulting cluster a clear separation between samples from normal tissue and tumor tissue can be seen with only one of the tumor samples (number 123) misclassified as a normal sample (Figure 22).

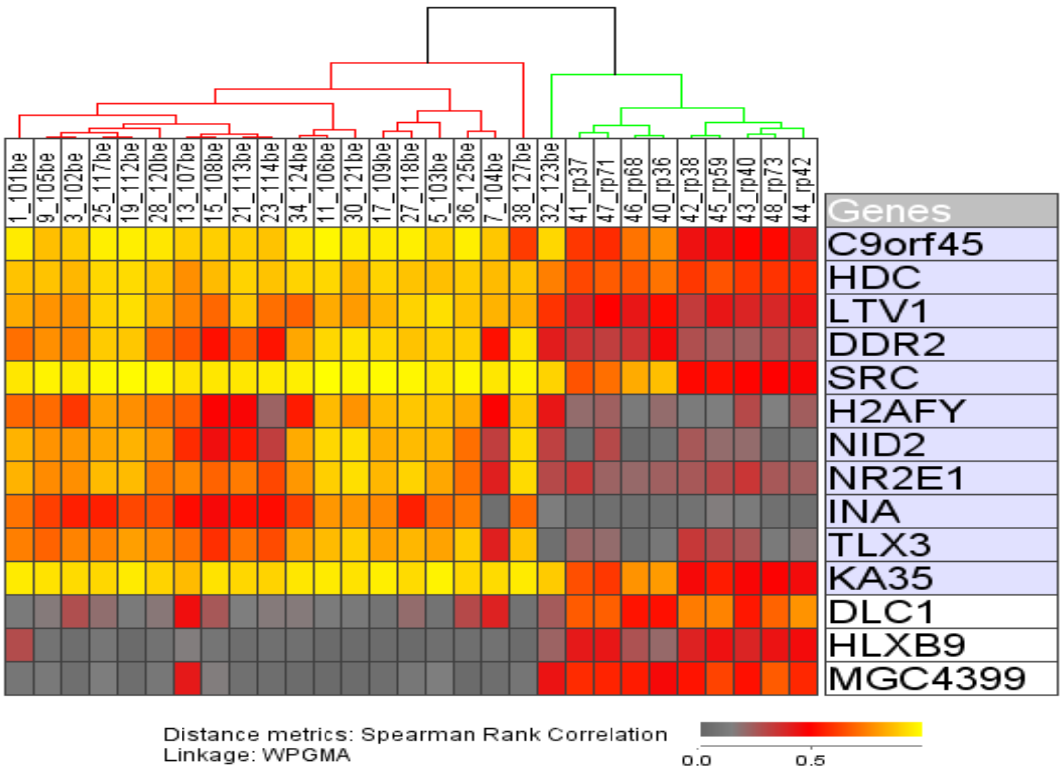

Figure 22 The dendrogram shows the genes names selected from PAM on the right side and the gene clusters at the left side. The normal (green color) and tumor samples (red color) are shown in branches on the top, except from sample number123, which is a tumor sample clustered with normal samples. The heat map mirror the beta value provided from the methylation analysis. The yellow spots indicate high degree of methylation and the grey spots low methylation levels. The normal and tumor samples are clearly separated with only one tumor sample misclassified as a normal sample.

The list of genes provided from the PAM analysis contained only 14 genes for the selected threshold which is not enough to perform a valid pathway analysis. All further analysis on this gene set is based on the genes as separate entities.

5.3.3.3 Statistical testing of genes differentially methylated in normal tissue versus non-radiated tumor tissue

All the 14 genes described above (Table 19 in appendix A) were analyzed using a t-test (two-sided with two sample equal variance) to determine to which extent they are differentially methylated between the two groups. All the results are listed in Table 10. Because of high standard deviation a F-test was performed to determine if the two sample groups had equal variance. The genes C9orf45, HLXB9, HDC and DLC revealed no significant variance between the normal versus tumor group in the t-test analysis.

Table 10 Genes identified through the PAM analysis listed together with information of average methylation level and standard deviation within each group, p-value from t-test and p-value from the Man-Whitney U-test. Genes with a F value lower than critical value corresponding to equal variance are marked with equal. Those with significant different variance between the two groups reached a F-value higher than critical value.

Genes	Average methylation norm/tumor	SD norm/tumor	T-test p-value	F-test critical value=2,477	Man-Whitney P-value
C9orf45	0,55/0,90	0,12/0,07	5,5E-10	Equal	
DDR2	0,32/0,75	0,07/0,16	1,5E-8	Different	0,022
H2AFY	0,18/0,67	0,05/0,18	2,4E-8	Different	<o,oo1
HDC	0,63/0,86	0,04/0,04	4,5E-13	Equal	
HLXB9	0,38/0,08	0,08/0,06	6,2E-11	Equal	
INA	0,07/0,57	0,04/0,19	2,4E-8	Different	<0.001
KA35	0,58/0,91	0,13/0,03	1,1E-11	Different	<0,001
NID2	0,13/0,70	0,10/0,20	2,2E-8	Different	0,001
NR2E1	0,26/0,75	0,05/0,16	1,7E-9	Different	<0.001
SRC	0,61/0,71	0,15/0,02	1,3E-10	Different	<0,001
TLX3	0,19/0,71	0,09/0,19	2,8E-8	Equal	
DLC	0,65/0,19	0,09/0,10	1,1E-11	Equal	
MGC4399	0,56/0,11	0,05/0,11	1,8E-11	Different	<0,001
LTV	0,42/0,80	0,05/0,09	2,5E-12	Different	<0,001

The non-parametric test Mann Whitney U-test was performed on the genes with high variance between the groups. All the genes were also significant differentially methylated between the two groups, though with a higher p-value (Table 10). The distribution of the beta-values for the genes *H2AFY* and *NID2* are shown as box plot in

Figure 23.

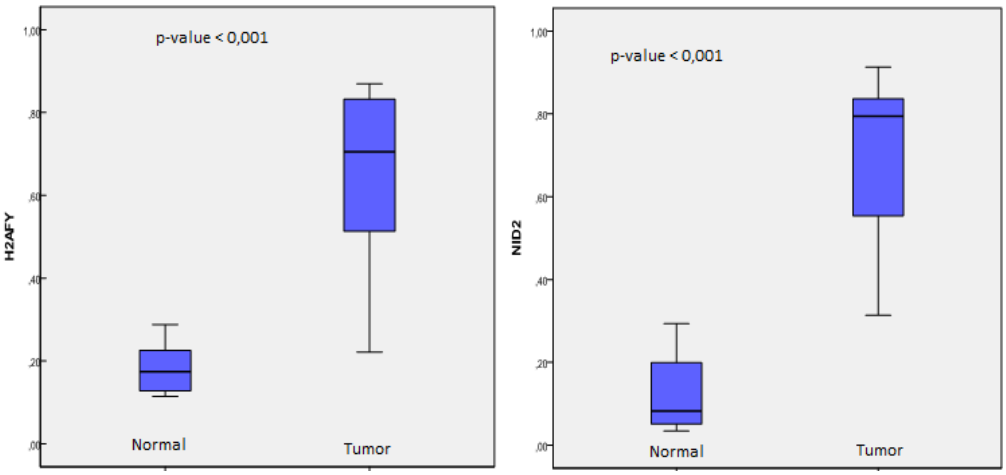

Figure 23 the genes H2AFY and NID2 were significantly differentially methylated (p <0.001) between normal samples and tumor samples. The box plot visualizes the median value in addition to upper and lower quantile, as well as maximum and minimum methylation value within each group.

5.3.4 Methylation pattern before and after radiation

The methylation patterns were compared between non-irradiated tumor samples and irradiated tumor samples (termed be and af) to find genes with a significant different methylation level between the two groups.

5.3.4.1 PAM analysis of tumor samples before and after radiation

The two groups were first analyzed in PAM in order to find genes whose methylation profile best separated the two tumor groups. Several thresholds values were tested, but the one with lowest error rate was chosen. A threshold from 0 to 0.8 seemed to give the lowest misclassification error, but testing resulted in an error rate on 0.377 (37 %) with a classifier build on more than 3000 genes (Figure 24). A threshold value of 1.7 was chosen, leading to a gene list with 140 genes (Table 20, appendix A) and an overall error rate on 0.278, indicating that 27.8% of the samples are misclassified.

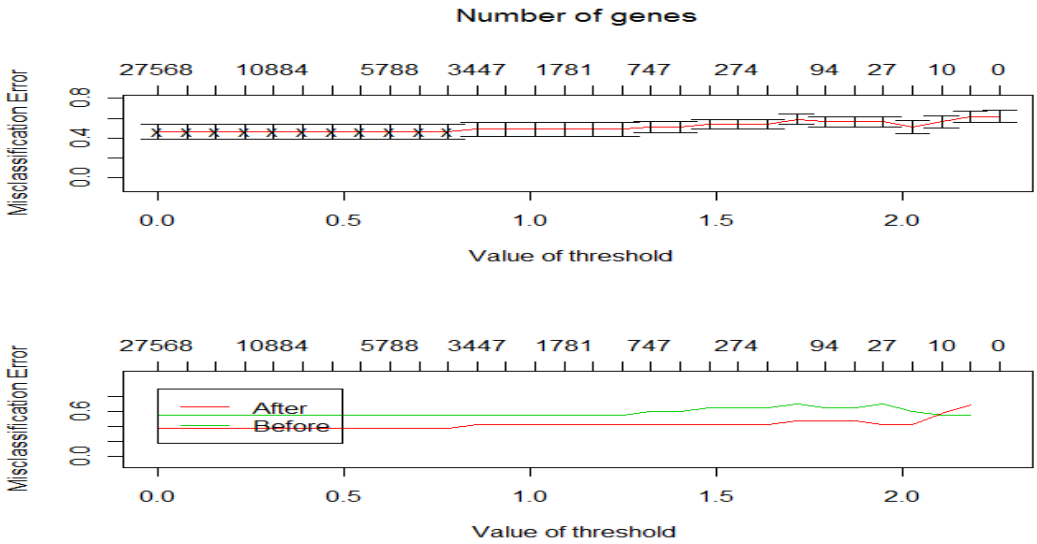

Figure 24 The misclassification error figure indicates that a threshold between 0 and 0.8 best separates the groups before and after. Several threshold values were tested but a value on 1.7 was the one which best separated the two groups, giving a list of 140 genes differentially methylated between the two groups.

5.3.4.2 *Supervised hierarchical clustering of before and after samples*

By using supervised hierarchical clustering with Spearman Rank Correlation and average linkage, the different samples were clustered into two main clusters based on the 140 genes identified from the PAM analysis (Figure 25).

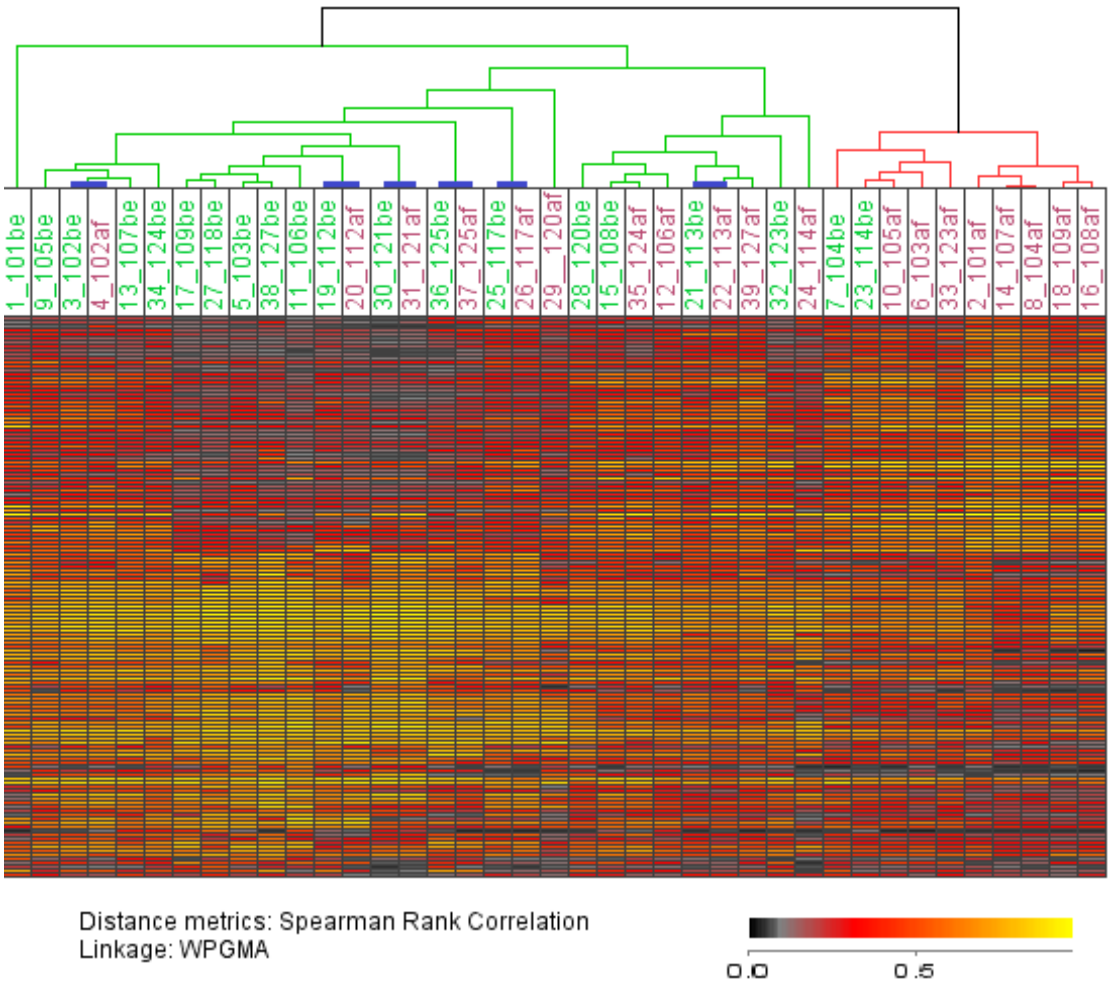

Figure 25 Before and after samples clustered using the 140 genes extracted from PAM. Eleven of the after samples (58%) are clustering together with the majority of the before samples and two of the before samples (10%) are clustered with the after samples. Six of the after samples clustered with its respective before sample, (marked with a blue line).

In the cluster dendrogram shown in Figure 25, most of the before samples are clustering together at the left side (green branching) of the dendrogram. 11 after samples are also clustered at the left side together with the before samples. The remaining 8 after samples are clustered at the right hand side (red branching) in the dendrogram, with two before samples also clustering with this group. Six samples are clustered together as a pair (sample number 102, 112, 121, 125, 117 and 113, marked with blue line) and further two samples are located nearby its respective pair sample (120 and 104). Paired samples clustering together give an indication of more similarities with its respective sample partner, than with the other samples within the corresponding class. Knowing that the second biopsy was drawn after five to twenty days after the first dose of radiation, the exact days and received dose before harvesting were investigated. This led to the discovery that for all the incidences where the before and after samples clustered together as pairs or within the same cluster, the after samples were collected after 5 to 11 days (which means after receiving only 10 to 16 Gy) into the radiation treatment, while samples that were taken out after 9 to 20 days (or

after 16 to 30 Gy), were better separated, shown in Figure 26. The exception being sample 120 which was collected after only 12 Gy but still gave a clear separation of the before and after sample.

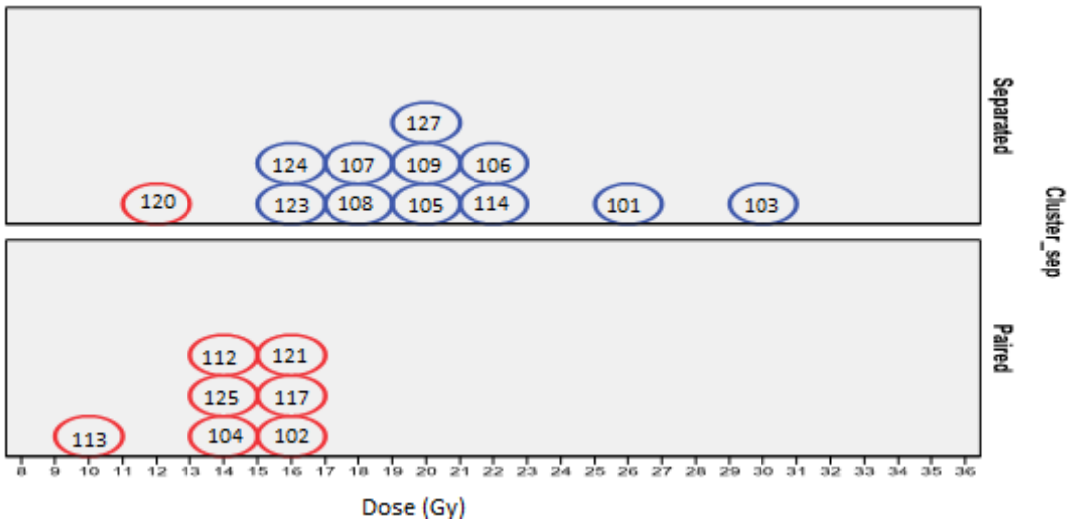

Figure 26 All the samples which were separated are shown in the upper panel and the samples clustering as pairs are displayed in the lower panel. The X-axis gives information about the given dose to each sample. The red circles are those after samples which were removed in the new be/af analysis.

In order to find genes stronger correlated to the radiation response, after samples which appeared to not separate from their respective before sample or received less than 16Gy, were removed.

5.3.4.3 PAM analysis of before/after tumor samples after removal of eight after samples

The reduced dataset, where eight after samples were removed, was loaded into PAM and analyzed. The misclassification error figure (Figure 27) indicates that thresholds within the areas 0.1 to 1, and 2.8 to 3 are best in order to separate the before and after samples. A threshold of 2.8 was chosen, leading to a gene list of 84 genes (Table 21, Appendix A). This threshold gave the lowest overall error rate (0,128) with one of eleven (9%) after samples and three of twenty (15%) before samples misclassified.

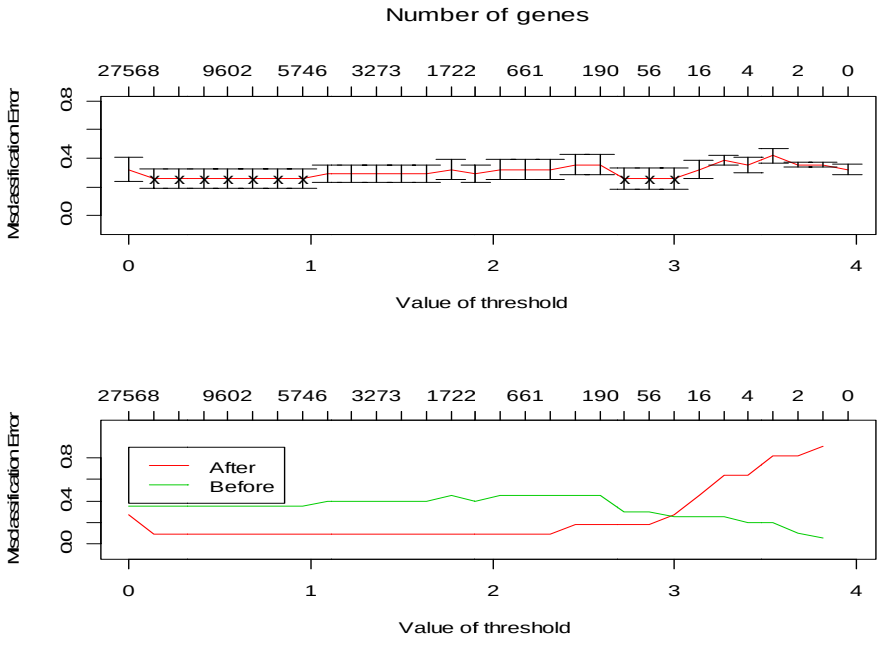

Figure 27 Threshold is chosen depending on misclassification error and number of genes. These graphs indicates that the lowest misclassification error rate can be found between threshold 0-1 giving a gene list of 5000-27568, or between 2.8 and 3 giving a gene list with 40-90 genes.

5.3.4.4 Supervised hierarchical clustering with the modified before and after gene list

The new gene list with 84 genes provided from PAM was imported into J-express, where a supervised hierarchical clustering was performed. The clusters in Figure 28 show a better separation of the two classes and reveal a more distinct difference in the methylation pattern between the two main clusters than compared with the clustering containing all the after samples.

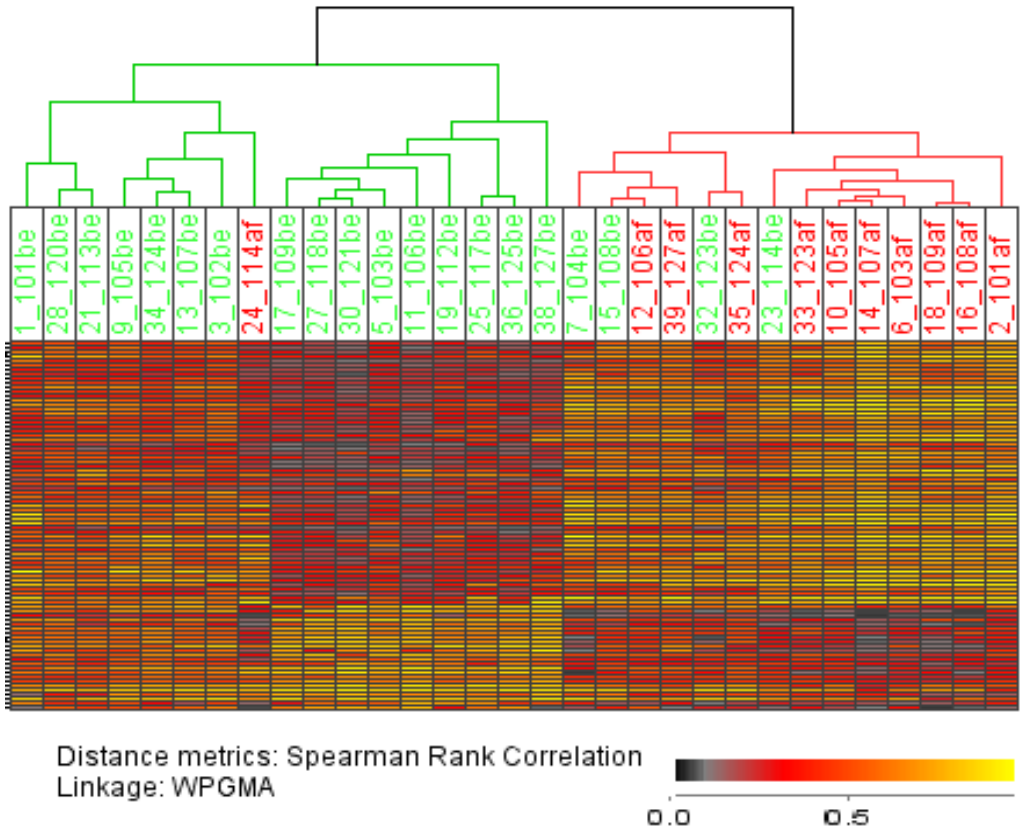

Figure 28 A new cluster analyses was performed after removing after samples not separated from its respective before sample in the previous analysis. This heatmap reveals a more distinct difference in methylation pattern between the two groups.

When looking at the dendrogram, the samples are divided into two main groups and four subgroups, which are further branched into smaller groups. The first main cluster, marked with green, contains mostly before samples and only one after sample. The second main cluster, marked red, contains mostly after samples, but also 4 before samples.

5.3.4.5 Ingenuity Pathways analysis in tumor samples before and after radiation

Both the original (no samples excluded) and the modified gene list (after removal of 8 after samples) extracted from the PAM analysis of the before- after samples, were further investigated in the Ingenuity Pathway Analysis software program to identify pathways overrepresented within the gene lists. The top canonical pathways for both lists are shown in Table 11. IL-10 signaling pathway, acute phase signaling and LXR/RXR pathway were found to be significantly overrepresented within both gene lists ($p<0.01$).

Table 11 The top five canonical pathways identified with in the gene lists extracted from the comparison of all before –after samples (A) and only a subset of the samples (B). For each canonical pathway the level of significance, FDR (false discovery rate) and ratio (genes included in the imported gene list / the total number of genes in the canonical pathway) is provided. The last column indicates the genes from the imported gene list that are involved in the given pathway

A. Top Canonical Pathways (no af samples excluded)	P-value	FDR	Ratio	Molecules involved in the pathway
NF-kB signaling	1,07E-03	0,08	6/172	FCER1G, **IL1B**, **IL1R2**, TANK, **TNFRS1B**
Acute Phase Response Signaling	1,92E-03	0,08	6/177	FGG, **IL1B**,**LBP**, PIK3CD, SERPINA3, **TNFRSF1B**
IL-10 Signaling	1,71E-03	0,08	4/72	CCR1, **IL1B**, **IL1R2**, **LBP**
Serotonin Receptor Signaling	1,92E-03	0,08	3/33	HTR2A, HTR3D, SLC18A1
LXR/RXR Activation	2,57E-03	0,08	4/83	**IL1B**, **IL1R2**, **LBP**, **TNFRSF1**B
B. Top Canonical Pathways (8 af samples excluded)	**P-value**	**FDR**	**Ratio**	**Molecules involved in the pathway**
Hepatic Cholestasis	1,75E-03	0,113	4/142	**IL1A**, **IL1R2**, **LBP**, SLCO1C1
IL-10 signaling	2,18E-03	0,113	3/72	**IL1A**, **IL1R2**, **LBP**
LXR/RXR activation	2,99E-03	0,113	3/83	**IL1A**, **IL1R2**, **LBP**
Acute Phase Response Signaling	3,68E-03	0,113	4/72	FGG, **IL1A**, **LBP**, RBP1
IL-6 signaling	5,43E-03	0,113	3/98	**IL1A**, **IL1R2**, **LBP**

The top network functions in the before after analysis (no after samples removed) were “Immunological Disease, Cellular Movement, Hematological System Development and Function” with a score of 90 involving genes such as *IL1B*, *IL1R2*, *LBP* and *BCAN*. In the modified before after analysis, “Amino Acid Metabolism, Cellular Assembly and Organization, Connective Tissue Development and Function” was the top associated network function with a score of 89. In this network, 45 molecules from the modified gene list were involved and included *BCAN*, *H2AFY, IL1A*, *IL1R2* and *LBP.*

5.3.4.6 Statistical testing of genes from the modified before/after gene list

The gene list extracted from PAM is sorted from the most differentially methylated to the least differentially methylated gene between the analyzed groups. *BCAN*, *H2AFY*, and *PPGB* are the three genes most differentially methylated between the before-after groups for the modified sample set. These three genes, in addition to three genes identified through the pathway analysis to be central for several of the identified canonical pathways (*LBP, IL1A, IL1R2*), were further investigated. For testing to which extent they are differentially methylated between the two groups, a t-test (two tailed with two sample equal variance) was performed. The genes *IL1A* (p=0,0002), *IL1R2* (not significant), *BCAN* (p=2,1E-5), *H2AFY* (p=2,1E-5) were found to have a higher methylation level before than after radiation. The genes *LBP* (p=0,0003) and *PPGB* (p=0,0001) were less methylated in the after samples, than in the samples collected before radiation. A F-test uncovered a significant different variance between the groups in the gene *IL1R2*. This gene was therefore tested with Man-Whitney U-test.

The beta value distribution in the two groups (before and after) is illustrated with box plot for the genes *H2AFY, LBP, BCA, PPGB, IL1A and IL1R2* in Figure 29. All the genes were significant differently methylated (P<0,05) between the two groups, using Man-Whitney U-test.

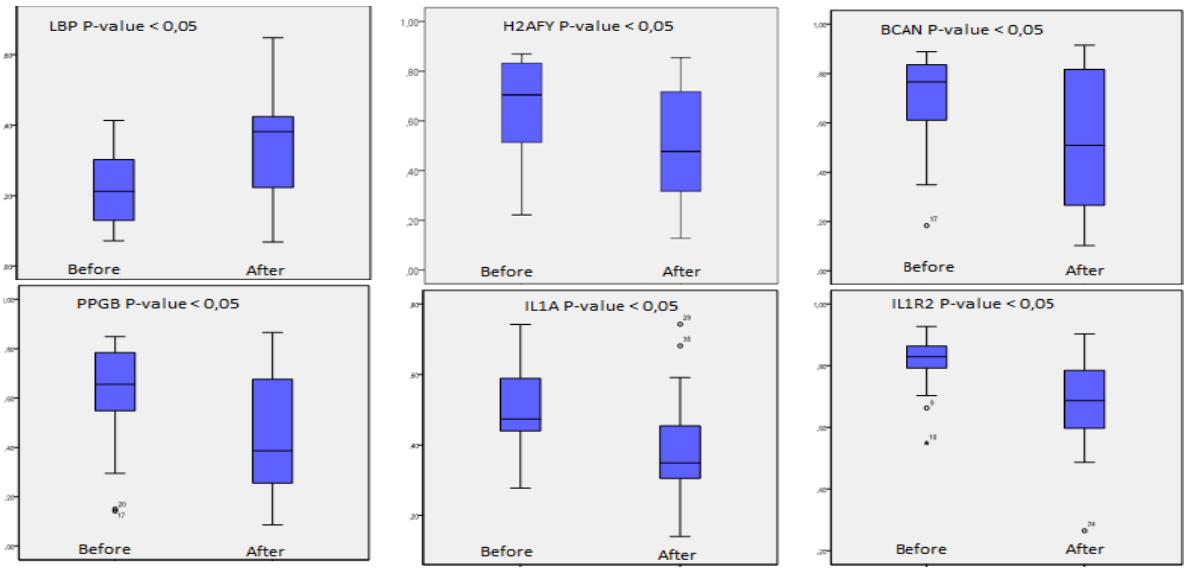

Figure 29 The median methylation level were found to be significantly different between the before and after group for the genes *LBP, H2AFY, BCAN, PPGB, IL1A and IL1R2* using a Mann Whitney U-test. The maximum and minimum value, upper and lower quantile, in addition to the median value for each group is visualized with the box plot. Outliers are shown with small black circles.

5.3.4.7 Dose dependent methylation changes in before and after samples

The genes *H2AFY, BCAN, LBP, PPGB and IL1A* revealed a significant change in methylation level in breast tumor tissue after exposure to radiation. In order to investigate if the changes in methylation level were significantly correlated to dose, a regression analysis were performed in SPSS (ver.19). The delta value between before and after samples were calculated and for the genes *H2AFY*, *PPGB* and *IL1A* the changes in methylation level were found to be significantly

correlated with dose (P=0,002, P=0,001 and P=0,001, respectively). The methylation level of the genes *BCAN* and *LBP* were not found significantly correlated to radiation dose.

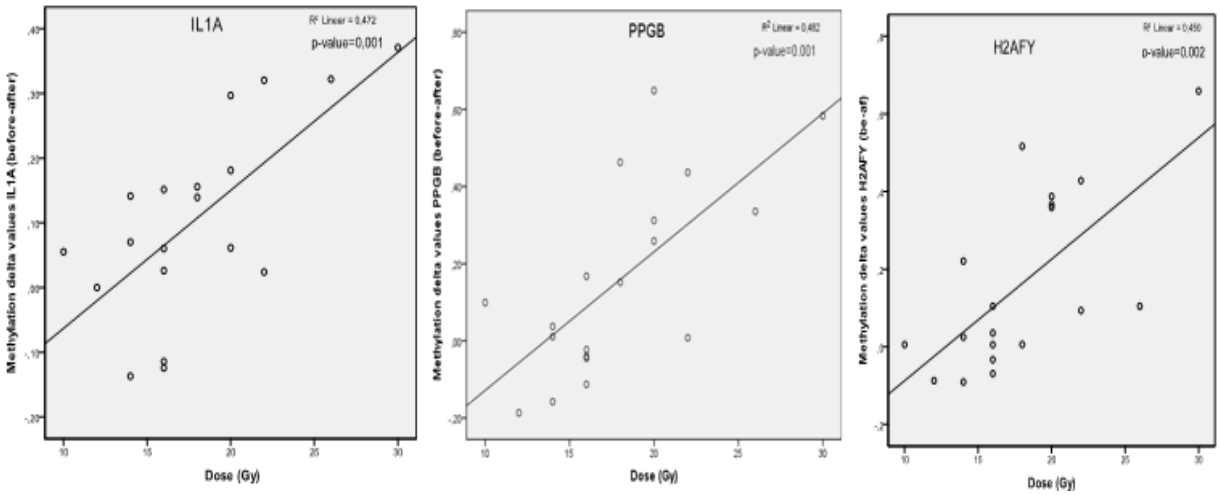

Figure 30 The delta methylation values between before and after samples were calculated and correlated with given dose. For the genes *IL1A*, *PPGB* and *H2AFY* a significant correlation with given dose was detected.

5.3.5 Methylation pattern and response

The non-radiated tumor samples were grouped into two classes. The good response group contained those reported with good response to treatment and the poor response group contained those with reduced response, meaning those reported with partial response, no response, and progression, measured as discussed in chapter 5.2.

5.3.5.1 PAM analysis in non-radiated tumor samples with different response to radiation

The methylation level in the predefined response groups, described above, was analyzed in PAM in order to find genes differentially methylated between the two groups. The misclassification error figure indicates threshold values from 2.2 to 2.9 to give the best separation between the two groups (Figure 31). These thresholds gave similar overall error rate, zero, meaning that the identified subset of genes for each threshold, separated the samples correctly according to the predefined phenotype. In order to obtain a longer gene list, a threshold value of 2.2 was chosen, resulting in a gene list with 342 genes differentially methylated between the two response classes. The gene list is shown in

Table 22 in appendix A.

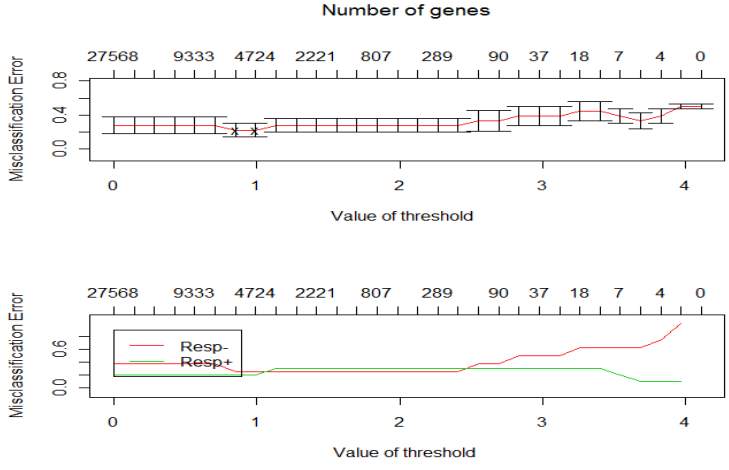

Figure 31 threshold values from 2.2 – 2.9 gave an overall error rate of zero for the classification of the before samples in to good and poor responders. A threshold value of 2.2, resulted in a gene list of 342 genes**.**

5.3.5.2 Supervised hierarchical clustering of non-radiated tumor samples according to response

The methylation data for the 342 genes extracted from PAM (

Table 22, appendix A) were used in J-express for supervised hierarchical clustering, (Figure 32). Only two samples, reported with no response were misclassified (sample 113 and 120).

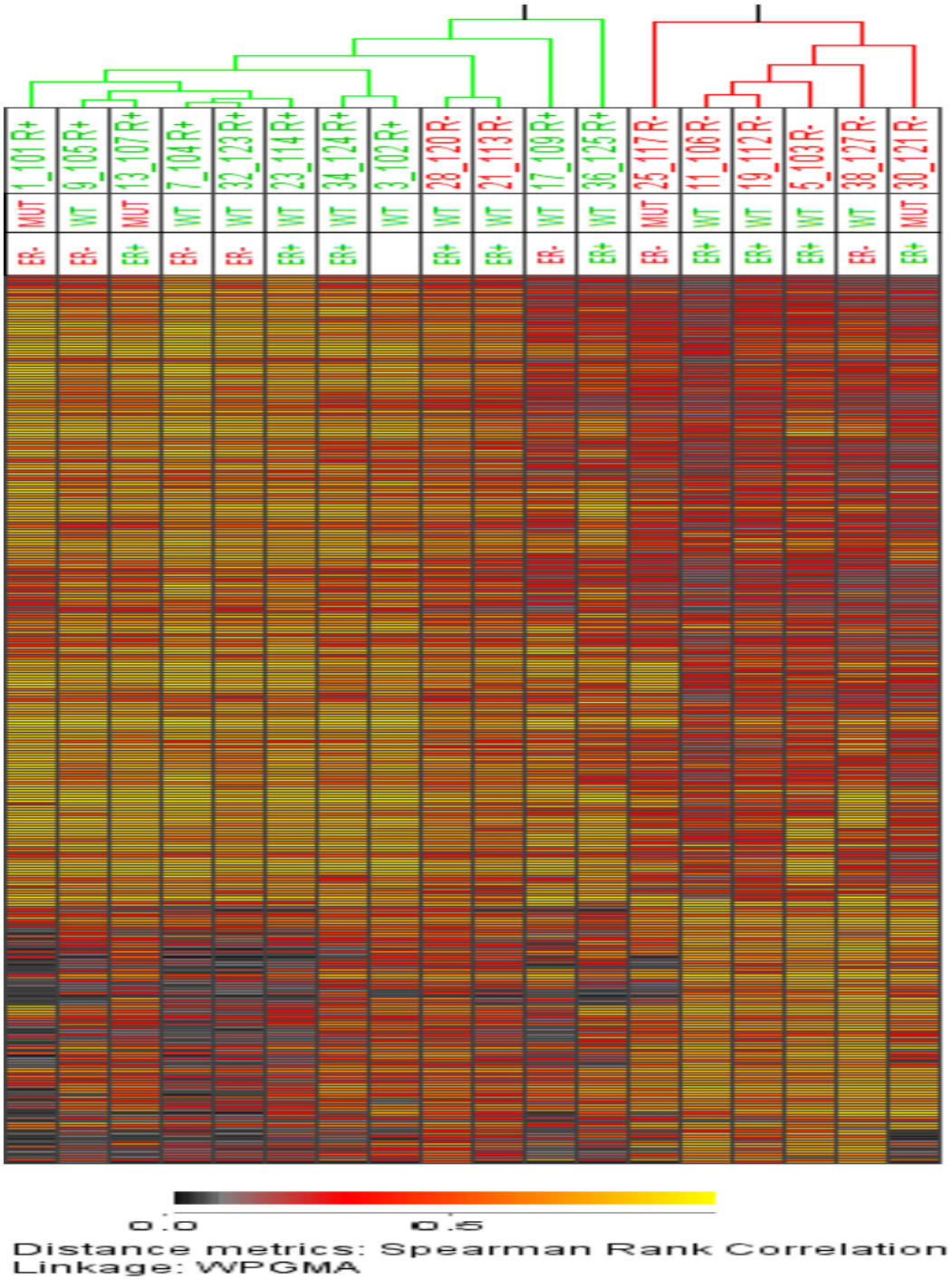

Figure 32 The before tumor samples were divided into two response groups, good (R+) and poor (R-), depending on the response to radiation treatment. All the samples in the cluster at the right side in the dendrogram are reported with poor response. The samples in the cluster at left side are primarily reported with good response. Sample number 120 and 113 were misclassified as good responders but are reported with poor response. ER status and p53 status for each sample are displayed in boxes under the sample number.

The dendrogram revealed a clear difference in the methylation pattern between the two response groups. Patients with good response are found in the left cluster, while those with a reduced response are primarily located in the right side cluster. Sample number 117 is the only sample reported with progression, and this sample clustered together with the poor response group, but in a sub cluster without other samples. ER status and p53 status are indicated in boxes below the sample numbers. No clear clustering of samples harboring p53 mutations or showing the same ER status is seen.

5.3.5.3 Ingenuity Pathway Analysis of genes differently methylated in the response groups

In order to find pathways and networks influenced by the genes differentially methylated according to response, the gene list (n=342) provided from PAM was imported into the Ingenuity Pathway Analysis software. The five top canonical pathways identified as overrepresented within the gene list are shown in Table 12. The p-value is significant for the five canonical pathways, but the FDR indicate a high false discovery rate (between 17.2 and 19.1 %). The associated genes from the imported gene list are shown at the right side in the table.

Table 12 The five top canonical pathways overrepresented within the 342 genes identified as differentially methylated between the two response groups: good and poor responders. For each canonical pathway the level of significance, FDR and ratio (genes included in the imported gene list / the total number of genes in the canonical pathway) is provided. The last column indicates the genes from the imported gene list that are involved in the given pathway. Genes in bold are found in two or more of the top canonical pathways.

Canonical Pathways	P-value	FDR	Ratio	Molecules involved in the pathway
G-protein Coupled Receptor Signaling	1,11E-03	0,172	18/517	**CNR1**, GPR1, **HRH2**, PTGFR, **CAMK2A**, **TAAR1**, GPR14, **ADCY4**, F2RL2, **FPR1**, BAI1, **GRM8**, ADRA1D, **PDE10A**, ADRA1B, MASIL, GPR83, GHRHR
cAMP mediated signaling	1,73E-03	0,172	10/212	**CNR1**, **HRH2**, **CAMK2A**, **TAAR1**, **ADCY4**, , **PDE10A**, **GRM8**, **FRP1**, PK1A, CNGA2
Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	4,64E-03	0,191	3/22	DEFB103A/DEFB103B, **CCL4**, **CCL2**
Chemokine signaling	4,71E-03	0,191	5/70	**CCL2**, **CCL4**, CCL13, CCL7, **CAMK2A**
Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza	4,81E-03	0,191	4/44	**CCL2**, **CCL4**, IFNA14, IFNA13

In the associated network analysis “Cell Morphology, Cellular Development and Nucleic Metabolism” involved the highest number of genes (n=96 genes) from the imported gene list, including genes such as *ADAMTS8*, *RASGPR3*,

RASSF6, *RUNX3* and *SFRP2*. Following this, the "Antigen Presentation, Cellular Movement, Hematological System Development and Function" network with totally 72 genes from the imported list (including genes like *CCL1*, *CCL2*, *CCL4*, *CCL8*, *CCL13*, *BAI1*, *BAK1* and *EGR2*) and the network "Cell Death, Hematological System Development and Function, and Gene Expression" with 48 genes from the gene list (including *ESRRB* and *ROS1*) were the two networks with the highest number of genes from the gene list involved.

5.3.5.4 Statistical testing of genes differently methylated according to response

The three top genes from the PAM analysis, *ESRRB*, *MYCT1* and *CCL2* were the genes showing the most different methylation level between the two response groups. To further test to which extent methylation levels can be associated with response status, a t-test was performed. *ESRRB*, *MYCT1* and *CCL2* all had a significantly higher methylation level in those with good response compared to the poor responders (p-value 1,7E-5, 2,4E-5 and 0,006, respectively). In addition, the genes *EGR2* (p=0,001), *RASGRP3* (p=0,007), *ROS1* (p=0,006), *RUNX3* (p=0,005) and *SFRP2* (p=0,010) reported in the literature as frequently abrogated in cancer, were also investigated further to study the association with response. Because of high standard deviation, a F-test was performed to make sure that right statistical test was chosen. The genes *RUNX3*, *SFRP2*, *ESRRB* and *CCL2* revealed a significant differently variance between the two response groups. These four genes were therefore tested using Mann Whitney U-test and were found also with this method to have a significant different methylation level between the two response group (p=0,013, 0,010, <0,001 and 0,001). The distribution of beta values in the genes *ESRRB*, *MYCT1, CCL2, RASGRP3, SFRP2* and *EGR2* are displayed as box plots in Figure 33.

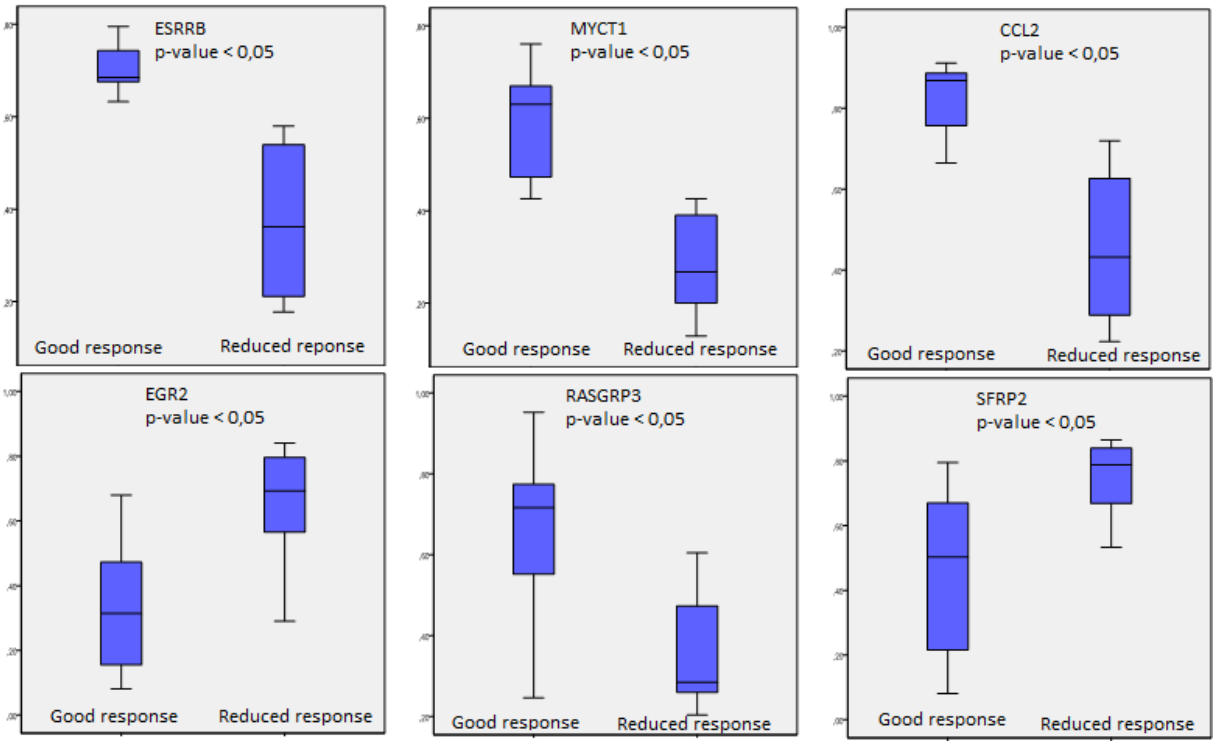

Figure 33 The genes *ESRRB*, *MYCT1*, *CCL2*, *EGR2, RASGRP3* and *SFRP2* were significant different methylated between the two response groups. The distribution of the beta values within the groups is visualized in the box plot. The box plots display the maximum and minimum value, upper and lower quantile, in addition to the median value for each gene.

5.4 Mutation and SNP analyses

Genetic alterations in some of the members of the p53 pathway, as well as mutations in *PIK3CA* are reported to influence the radiation response. Analyses of these elements are therefore included in this thesis.

5.4.1 Sequencing p53

All the coding exons and flanking introns of *TP53* were sequenced on ABI PRISM®3730. The results from the sequencing were aligned and analyzed in SeqScape v.2.5 (Applied Biosystem), and the sequences were checked independently by two persons. Figure 11 is an example of an electropherogram from sample number 116, exon 5. An overview of all mutations and polymorphisms found in this material are listed in Table 23 in appendix A. Mutation in *TP53* was found in six of the 22 tumor samples (27,3% of the tumor samples were mutated). No mutations were found in any of the 15 normal samples as expected.

In earlier studies, cDNA from the same tumor material had been sequenced. In order to check for aberrations in splice variants and introns, sequencing of gDNA was performed. Comparing data from cDNA with data from gDNA revealed some discrepancies. In cDNA, a mutation in sample number 105 codon 281 was reported, but this was not detected in gDNA. Sample number 109 was described with one unverified base pair deletion in codon 135 in cDNA, but this was not confirmed in gDNA. Sample number 116 (shown in Figure 11) had a heterozygote mutation in gDNA codon 254. This mutation was not reported in cDNA. The forward string of sample number 116 showed C/C in c.72, while G/C was revealed in the reverse string. The G allele was very weak in the tumor, but was confirmed in the blood sample. Sample number 109 carried different polymorphism in c.72 when comparing blood- and tumor results. Samples with inconsistent results were sequenced twice to ensure quality control.

5.4.2 SNP analysis in Mdm2, Mdm4 and PIK3ca

3 additional polymorphisms in the p53 pathway were analyzed; two SNPs in *MDM2 (*rs 2279744 (SNP 309) and rs 117039649 (SNP 285)) and rs1563828 in *MDM4*. In addition, *PIK3CA* was screened for mutations in exon9 and exon20. Associations between genetic variations and response to radiation treatment were investigated using crosstabs and Fisher's exact test. All polymorphisms sequenced in *PIK3CA* (exon 9 and 20) were counted in one group. SNP 285 was detected only in 1 sample and no further analysis on the impact of this SNP was performed.

The difference in distribution of the sequenced genotypes between normal and cancer samples were analyzed using Pearson Chi Square. Allele frequencies were grouped according to sample type (

Table 15). There was no significant difference in genotype frequencies between normal- and cancer samples for any of the analyzed SNPs.

Table 13 The polymorphisms in *MDM2* (rs 2279744 (SNP 309) and rs117039649 (SNP 285)), *MDM4* (rs1563828), *PIK3CA* and c.72 (rs1042522) and recorded response to treatment are listed in the table. The four response groups are: G=good, Pt= partial, No= no response and Pg= progression.

SNP/ haplotype	Patient-id										
	101	102	103	104	105	106	107	108	109	110	112
MDM2 SNP 309 rs 2279744	G/T	T/T	T/T	G/T	T/T	G/G	G/T	G/T	G/T	G/T	T/T
MDM2 SNP 285 rs117039649	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
MDM4 rs1563828	A/G	G/G	G/G	G/G	A/A	A/A	G/G	G/G	G/G	A/G	G/G
TP53 C.72 rs1042522	G/G	G/C	C/C	C/C	G/G	G/G	C/C	G/G	G/G	C/C	C/C
PIK3CA	G>A	WT	G>A	WT	WT	A>G	WT	WT	WT	WT	WT
Response	G	G	No	G	G	N	G	-	G	N	Pt
SNP/ haplotype	Patient-id										
	113	114	116	117	118	120	121	123	124	125	127
MDM2 SNP 309 rs 2279744	T/T	T/T	G/T	T/T	G/G	T/T	T/T	G/T	G/T	G/T	G/T
MDM2 SNP 285 rs117039649	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/C	G/G	G/G	G/G
MDM4 rs1563828	G/G	G/G	G/G	A/G	G/G	A/G	A/G	A/G	G/G	A/G	A/A
TP53 C.72 rs1042522	G/G	G/G	C/C	G/C	G/G	G/G	G/C	G/C	G/G	G/G	G/G
PIK3CA	WT	WT	WT	WT	WT	A>G	G>A	WT	WT	WT	WT
Response	N	G	Pt	Pg	-	N	Pt	G	G	G	Pt

Table 14 Overview of the polymorphisms studied in DNA from healthy tissue biopsies in *MDM2* ((rs 2279744 (SNP 309) and rs117039649 (SNP 285)), *MDM4* (rs1563828) and *TP53* c.72 and the resulting genotypes. Some of the samples failed to be analyzed and are therefore displayed with not determined (nd).

Norm. Samples	RP36	RP37	RP38	RP40	RP41	RP42	RP43	RP44	RP45	RP59	RP61	RP68	RP70	RP71	RP73	RP75
SNP 309	T/T	G/G	T/T	G/G	T/T	G/T	T/T	G/T	G/T	G/T	G/T	T/T	G/T	G/T	T/T	G/T
SNP 285	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
Mdm4	G/G	A/A	A/G	nd	nd	A/G	A/G	G/G	Nd	G/G	nd	G/G	A/G	G/G	G/G	nd
c.72	G/C	G/C	G/G	C/C	ND	G/G	G/C	G/G	G/G	G/C	G/G	G/C	G/C	C/C	G/G	G/G

Table 15 Genotype distribution and minor allele frequency (MAF) in *MDM2* (rs 2279744 (SNP 309)), *MDM4* dm4 (rs1563828) and *TP53* c .72 (rs1042522) for the material according to sample type (cancer= non-irradiated tumor tissue, normal= DNA from blood).

	Mdm2 (rs 2279744)			Mdm4 (rs1563828)			c.72 (rs1042522)		
	T/T	G/T	G/G	G/G	G/A	A/A	G/G	G/C	C/C
Frequency cancer	40,90 %	50,00 %	9,10 %	54,50 %	31,80 %	13,60 %	54,50 %	18,20 %	27,30 %
Frequency normal	37,50 %	50,00 %	12,50 %	66,70 %	26,70 %	6,70 %	50,00 %	37,50 %	12,50 %
MAF cancer			34,1%			29,5%			36,4%
MAF normal			37,5%			20,0%			31,3%

5.3. Response to treatment

The response was categorized into four groups; good response, partial response, no response and progression, as discussed in chapter 5.2. Because of few samples in this study, the samples were divided into two response groups, as described in chapter 5.3.5. Fisher's exact test was performed to identify relationship between p53/p53 pathway status and response to treatment. When harboring the minor allele, the polymorphisms in *MDM2* or *MDM4*are reported to give similar effect to response as a mutated *TP53 (29)*. The "p53 pathway" refers to the combined *TP53* mutation status and the unfavorable variant for the SNPs in *MDM2* and *MDM*:

p53 path WT = *TP53*wt Mdm2^{major}, Mdm4 ^{major}

p53 path mut = *TP53*mut or Mdm2^{minor},or Mdm4 ^{minor}

Table 16 Association between p53/p53 pathway mutation status and response to treatment Fisher's exact test revealed no significant association between p53/p53 pathways mutations and response to treatment

	p53 WT	p53 mut	p53 path WT	p53 path mut
Good response	83,3 %	16,7 %	66,7 %	33,3 %
Poor progression	60,0 %	40,0 %	40,0 %	60,o %
Fisher's exact test (p-value)		0,348		0,391

No significant difference was observed when analyzing the association between p53/p53 pathway mutation status and response to treatment Table 16. However, a tendency of lower response to treatment when harboring a mutation in p53 or p53 pathway is indicated.

Patients with ER negative status are reported to have increased levels of p53 mutations. Crosstabs and Fisher's exact test were performed to investigate any relationship between ER status and p53/p53 pathway mutations.

Table 17 Fisher's exact test was performed to investigate a correlation between ER status and p53/p53 pathway. The pathway mutations had a stronger relationship to ER negative than mutation in the gene *TP53*

	p53 WT	p53 mut	p53 path WT	p53 path mut
ER neg	66,70 %	33,30 %	33,30 %	66,70 %
ER pos	80,00 %	20,00 %	70,00 %	30,00 %
Total	73,70 %	26,30 %	52,60 %	47,40 %
P value	0,611		0,179	

No significant correlation was found when comparing p53 status with ER status. When taking the polymorphisms in *MDM2* and *MDM4* into account, a tendency towards stronger relationship between ER negative and an affected p53 pathway was revealed although not reaching statistical significance (Fisher's exact test p-value = 0,179).

5.4.2.1 SNP's in the p53 pathway and mutations in PIK3CA in relation to response to radiation

To investigate the difference in distribution of SNPs in the two radiation response groups, Chi square and Fisher's exact tests were performed. No significant difference in genotype distribution was identified between the different response groups (Table 18).

Table 18 Distribution of SNP's in MDM2-309 (rs 2279744), MDM4 ((rs1563828), TP53 codon 72 (rs1042522) and mutations in *PIK3CA* within the two response groups.

	Mdm2			Mdm4			c.72			PIK3CA	
	G/G	G/T	T/T	G/G	G/A	A/A	G/G	G/C	C/C	WT	Heteroz.
Good/partial response	41,7%	50,0%	8,3%	66,7%	25,0%	8,3%	66,7 %	16,7%	16,7%	91,7%	8,3%
no response/ progression	60,0%	30,0%	10,0%	40,0%	40,0%	20,0%	40,0%	20,0%	40,0%	60,0%	40,0%
P-value	0,632*			0,440*			0,40*			0,135§	

*Pearson correlation coefficient §Fisher's Exact test

Mutations in *PIK3CA* were studied for association with response to treatment by Fisher's exact test. This test did not reach statistical significance (p-value of 0,135), but a higher frequency of individuals carrying mutations in *PIK3CA* were found in the poor response group.

If the response groups were divided differently, those with good and partial response into one group, and those with no response and progression into the other group (termed modified response group), the results would off course be altered. *MDM2* SNP 309 revealed then a borderline significantly association with radiation response (Pearson P= 0,077). The tendency was that patients harboring the genotype (T/T) in SNP 309 had lower response to treatment than patients carrying the genotype G/G or G/T. However, since 4 cells had less than 5 observations, Fisher's exact test would be more reliable. The samples were divided into two groups; the genotype G/G in one group and the genotype T/T together with heterozygote genotype G/T in the second group. This analysis did not confirm the same tendency.

6 DISCUSSION

In this part the methods and results are discussed in the same order as they are presented under the result section.

6.1 DNA extraction from breast tissue

DNA from some of the tumor tissue samples and all of the normal breast tissue samples were extracted. When DNA was isolated, the concentration was measured both with absorbance and fluorescence. When using absorbance, all nucleic acids are measured, while fluorescence only measure dsDNA. Therefore, samples measured with fluorescence should remain lower than those measured with absorbance. The fluorescence method is only linear up to 100 ng/µl and samples with the highest concentrations may therefore be unreliable. In the low concentration area a better concordance was seen between the two measurements, than in higher area. Nearly all the samples had a low concentration of DNA, with both methods, which is understandable because the breast often consists of much fat and the tumor samples may have contained some necrosis, factors making it difficult to extract DNA.

6.2 Methylation analysis

In order to unravel more of the underlying molecular mechanism of radiation response, genes and pathways associated with the different methylation patterns between the groups of interests were investigated. This was performed between normal and non-irradiated tumor tissue, between non-irradiated and irradiated tumor tissue as well as in relation to response data.

6.2.1 Considerations regarding the bioinformatic and statistic analysis

The methylation arrays analysis generates a large amount of data calling upon the use of different bioinformatic and statistical tools traditionally used in the analysis of data from mRNA expression arrays to interpret the data. The beta values provided from methylation analysis range between 0 and 1 and refer to percent methylation. Log-transformation of the methylation data in order to reach a normal distribution is therefore not advisable. Based on this, Prediction Analysis for Microarrays (PAM) in R was chosen to analyze the data both because of its capacity to handle such a huge data set as well as it being a non-parametric method not requiring normally distributed values. In PAM analysis the data will be ranked. Through the PAM analysis a gene list with genes differentially methylated between the groups of interests, was extracted depending on the selected threshold. Threshold was chosen on basis of what gives the best separation of the categories and the resulting number of genes to include in downstream analysis. The most important factor was the separation, if several options of equally good separation were available, a threshold giving a long gene list was chosen. For this project, a gene list containing between 100 and 500 genes would be ideal. From experience, it is difficult to do a pathway analysis with less than 100 genes, but a too long gene list is difficult to handle and may contain genes with less clear information (“noisy genes”).

Hierarchical clustering was used in two different ways to analyze the methylation data, unsupervised and supervised. The unsupervised clustering can be performed with all the methylation probes available without giving any prior information on subgroups within the dataset. The supervised clustering uses predefined categories and only the subset of genes related to the categories are imported to the program. Thus, already in this step most of the genes

are excluded for further analysis. An unsupervised hierarchical clustering was performed on all samples in combination (normal samples and both non-irradiated and irradiated tumor samples) by the epigenetic group at Centre Nationale de Genotypage (CNG) in Paris, while supervised hierarchical clustering was performed in J-express on the basis of the selected gene sets provided from the PAM analysis on selected subsets of the samples. To perform an unsupervised clustering in J-express, probe filtering is required since the program is not able to cluster samples based on a large number of genes. The unsupervised clustering could have been performed in R, but this was not prioritized because of time limitations. Different statistical options can be chosen when performing a hierarchical clustering analysis. In the cluster analysis used in this project, the algorithm average linkage and Spearman Rank Correlation were chosen. Spearman Rank Correlation cluster genes and samples with correlated levels of methylation. This is visualized by the heatmap in the dendrogram where different color refers to the beta-values. Another option was to use Euclidian, which calculate the distance between the methylation levels (73).

The genes with highest score provided from PAM (indicating the largest difference in methylation levels between the studied groups), and some of the genes identified through the pathway analysis or which previously had been associated with breast cancer were investigated in a comparison t-test for significance. Some of the groups contained high standard deviation which was considered when interpretations of the results were discussed. High standard deviation is often seen with non-normally distributed data, making the t-test less suitable, as it requires normally distributed data (75). For genes uncovering a high variance in the F-test, a Man-Whitney U-test was used for comparison of the genes investigated. This is a non-parametric test, based on ranked ordering, and avoids the problem with high variance within the groups. A disadvantage of non-parametric methods is that they often are less powerful in detecting associations and tend to give less significant p-values (75). The t-test was therefore chosen in cases where the variance was acceptable for the parametric method.

In order to investigate a relation between changes in the level of methylation after irradiation and given dose, a linear regression was performed. This method indicates the strength of the association between the variables. Few samples in the test may influence the results and this must be considered when interpreting the data.

For the pathway analysis, the Ingenuity Pathway Analysis (IPA) was chosen. The same gene list as provided from PAM was used in IPA in order to find pathways and networks overrepresented within the gene list. The analysis is based on the number of molecules involved and does not take the measured values into account giving all genes the same weight in the analysis. This resulted in the fact that few genes with a high PAM score were found in the top canonical pathways. A weighting of the genes according to score may have given other interesting results. The numbers of genes involved, the p-value and FDR gives an indication whether the results are reliable.

6.2.1.1 The total methylation level

An overall global loss of methylation in the genome and hypermethylation within special regions are often observed in cancer (53). In this dataset, no significant differences in the overall methylation level between the normal, non-radiated and radiated tumor samples were observed. However, this does not imply that there were no differences in methylation level between different groups of samples, only that the overall degree was equal. LINE-1 and Alu-

elements are known to be hypermethylated in order to maintain genomic stability. Studies has shown that irradiation of cell lines result in hypomethylation of LINE-1 and Alu-elements (66). In the Illumina Infinium array used in this project, no probes covered these repetitive elements and these results could therefore not be validated.

6.2.1.2 Unsupervised Clustering

Overall sample relations were investigated using the unsupervised hierarchical clustering based on raw beta values. The analysis revealed a clear separation of normal samples and tumor samples. The tumor samples were distributed in many clusters, regardless of being irradiated. Eight of the samples clustered with its respective pair, and for one sample both the before and after sample were localized in the same cluster. These results tell us that some of the samples tend to cluster together independent of the amount of probes in the analysis. It sounds reasonable that tumor samples drawn from the same patient overall share the same methylation pattern and tend to cluster together. The paired tumor samples are in some way different from the tumor samples with separation of the before and after samples. If this is due to the received dose or other unknown factors must be further examined.

6.2.2 Methylation pattern in tumor and normal samples

The PAM analysis resulted in a gene list with 14 genes differentially methylated between the two tissue types. By using a threshold on 5.8, all the samples were classified correct according to the phenotype. This threshold resulted in a short gene list. One should expect that more genes were differentially methylated in normal and tumor tissue, knowing that a lot of mechanisms are regulated between the two tissue types.

Based on the gene list provided from PAM the samples were clustered using supervised hierarchical clustering in J-Express. In the resulting heatmap, tumor samples and normal samples showed distinct different methylation profiles with a clear separation into two main clusters. Tumor sample number 123 was the only misclassified sample and clustered together with the normal samples. One of the explanations may be contamination of normal tissue in the sample (low tumor percentage). On the other hand, this sample is also reported as triple negative and likely to be classified as subgroup basal-like (discussed in chapter1.8.1.1). An earlier study on methylation profiling with unsupervised cluster analysis showed that basal-like tumors tend to cluster in a mix group together with normal tissue samples (63).

Pathway analysis with the gene list with 14 genes from this study was not performed, because of too few genes. This analysis requires more genes for identification of pathways overrepresented in a set of genes.

Based on a t-test all the 14 genes were significantly differentially methylated when comparing the level of methylation between the normal tissue and the non-irradiated tumor tissue. Surprisingly, genes like *DLC1* and *MGC4399* were higher methylated in normal tissue than in tumor tissue. The gene *DLC1* are reported as a gene preventing cancer growth in several cancers including breast cancer, and are often deleted in breast cancer (76). However, hypermethylation of the gene *DLC1* is not reported as a frequent event in breast cancer (77). Down regulation of the gene *MGC4399* is reported in oligodendrogliomas (78) and the gene may be differentially regulated in breast cancer. Genes like *DDR2* and *TLX3* are reported as tumor promoting genes (79;80), but were in this data set higher methylated in tumor tissue than in normal tissue. The gene *DDR2* was in a previously study found significantly

hypomethylated in hepatocellular carcinoma (81). Regulation with methylation may vary between the types of tissue and this gene can therefore be regulated in a different manner in breast cancer. The gene *TLX3* are frequently overexpressed in T-cell acute lymphocytic leukemia (80), but have so far not been associated with breast cancer or radiation response. These results can be explained by several mechanisms. Methylation is associated with gene silencing, but other mechanisms like acetylation, miRNAs and modifications in chromatin remodeling complex (discussed in chapter1.8), also influences on the expression. Some genes have several alternative promoters (82), which can remain unmethylated and lead to expression. Methylation is also describes as a dynamic process and may change in a cyclic dependent manner (83). Other genes like *NID2* and *NR2E1* were higher methylated in the tumor samples compared with the normal samples. This is in line with previous findings reported in the literature. The gene *NID2* was found to be highly methylated in urine bladder cancer (84), and the gene *NR2E1* was found highly methylated in 11 of 21 early stage breast cancer patients (85). The heatmap also indicate that the degree of methylation vary between the tumor samples. When comparing the average methylation for each group, the diversity between the samples is not taken into account. The standard deviation is quite high for some of the genes, indicating a high variance in the group. Last, but not least, the different tumors may harbor different kinds of dysregulated genes.

6.2.3 Methylation pattern in tumor samples before and after radiation

Tissue being exposed to radiation is known to activate different pathways in order to maintain stability in the cells, such as described in chapter 1.5. By comparing irradiated tumor tissue with non-irradiated tumor tissue, different methylation pattern may be discovered and give important information about pathway regulations associated with radiation. This information may be useful when it comes to choice of treatment.

In the first PAM analysis (before removing eight after samples), the misclassification figure indicated a threshold of 1,7 resulting in a gene list with 140 genes, or a threshold on 0,8 giving a gene list with over 3000 genes. The first threshold was chosen because it gave the lowest overall error rate and 3000 genes are many genes to handle in further analysis. When analyzing the modified sample set (where eight after samples were removed) in PAM, a threshold of 2,8 was chosen. This threshold gave a list of 84 genes, which best separated the groups of interests. A lower threshold would have been preferred in order to provide a longer gene list, if this not had led to a higher misclassification error. A threshold of 1 indicated a low error rate, but also a gene list with over 5000 genes.

The first cluster analysis was performed on the dataset containing all the before and after samples. This analysis gave two clusters, where the first contained mostly before samples and the second contained mostly after samples. Still, a mix of before and after samples was seen. Interestingly, seven of the before samples clustered together or in the same cluster as its respective after samples, indicating that these shared more similarities with each other than with their categorized group. These samples were the same clustering together in the unsupervised cluster analysis. This means that after samples clustering together with their before sample do not show the same changes in methylation as those with a separation of the before and after samples Further investigation of the treatment information led to the discovery that the after samples had been collected at different time-points in the treatment regimen. All the after samples clustering together with their respective before sample, were collected after five to eight days, and

had therefore received a lower radiation dose before the second biopsy was taken. This is a very interesting finding indicating that the changes in methylation may be dose dependent. A correlation between methylation and the degree of absorbed dose *in vivo* has so far not been described in the literature. However, cell culture models have reported of dose dependent relationships with methylation (66). Another explanation to this phenomenon may also be the different number of days between the two sample drawings. The methylation machinery is described as a slow mechanism and methylation changes may be acquired at later post-irradiation times (66).

6.2.3.1 Removal of after samples

In order to obtain a better separation of the before and after samples, eight after samples were removed. Six of the removed after samples were clustered with their respective before sample, one were clustered next to it's before sample, but in different clusters, and the last sample was in the same cluster as the before sample, but not clustered as pair. It is uncertain if the two latter samples should have been removed, because they were not paired like the six other samples. On the other hand, they received a lower radiation dose than the remaining after samples.

6.2.3.2 Analysis of non-irradiated and radiated tumor tissue

In PAM a new gene list with 84 genes was provided after 8 of the after samples had been removed. The threshold on 2.8 was chosen since this resulted in the lowest error rate. A longer gene list would have been preferable, but a lower threshold would also have led to a higher error rate. The cluster analysis performed on the modified before/after gene list showed a more distinct difference in the methylation pattern in the two tumor groups. Only one after samples (number 114) is classified as before sample, and 4 before samples (number 104, 108, 123 and 114) are classified as after samples.

Sample 114 before is clustered with the after samples and 114 after is clustered with the before samples in. This may indicate a switch of the samples, but the sample tubes were marked with biopsy collecting date and this should therefore not be the case.

Ingenuity Pathway Analysis was performed using the gene list with all before and after samples as well as the modified sample list. The five canonical pathways given in Table 11, indicates a significant overrepresentation of these pathways within the two datasets, but the false discovery rates are high, especially in the modified data set. None of the pathways contained a FDR lower than 8%, which indicate that a too high degree of the identified pathways are false positives. All the results from the pathways analysis should therefore be interpreted with caution. The ratio which indicates the number of molecules from the imported gene lists involved in the identified pathways is low. This may be due to short gene lists. A longer gene list may have improved the ratio as well as the FDR, but would also contained genes with weaker significance. This should be further investigated in the continuation of the project.

Several of the identified pathways indicate an alteration in inflammatory processes. Not surprisingly, the "Acute phase response signaling" was one of the canonical pathways significant in both analysis. This is associated with response to infection, inflammation, tissue injury, malignant growth, immunological disorder and radiation. Several pathways are involved in the LXR/RXR activation, but especially the NF-κB pathway had numerous elements

represented in the gene list, leading to transcription of inflammatory mediators. The" IL-10 signaling pathway" is also involved in inflammatory processes (74). This pathway leads to expression of inflammatory cytokines such as TNF-alpha, IL-6 and IL-1 activated by macrophages. "Serotonin Receptor Signaling" is related to transmitter release and depolarization of neurons (74). In all the significant molecules involved in the pathways are shown. Three genes, *IL1A*, *IL1R2*, and *LBP* are involved in all of the top five canonical pathways. Three of the pathways contained only these three genes from the gene list. This may indicate that some of the pathways achieved a high score because of these three genes, or it may also indicate that these genes drive the pathways.

The t-test comparison resulted in a significant p-value for the genes *IL1A* and *LBP* while *IL1R2* was not found significantly differently methylated between the before and after groups.

IL1A is a pro-inflammatory cytokine known to be involved in several pro-inflammatory pathways, as well as in regulation of the immune response. This gene has also shown to be upregulated by radiation and may be predictive for development of radiation induced pneumonitis (86). In this study IL1A showed a higher methylation level in the before samples compared to the after samples. In a study of rat lung irradiation, the mRNA level of IL1A is reported to be elevated both in irradiated and in the bystander region of lung tissue, and occurred in a cyclic dependent manner (87).

The lipopolysaccharide binding protein, LBP is involved in the acute phase response (88) and is shown to activate macrophages (89). Interestingly, this gene was found significantly higher methylated in the radiated tissue, than in the tissue not exposed to radiation. This is surprising when we know that radiation causes a lot of damage which normally would have triggered the macrophages and acute phase response. The normal tissue and the non-radiated tumor tissue had almost the same degree of methylation in this gene (28% and 21% respectively). The average methylation level increased from 21% to 40% after radiation. Two samples showed a decrease in methylation level of the gene LBP and two other samples had no changes after radiation. Theses samples were recorded as good response, no response, and partial response. A larger material could possible revealed if the silencing of the gene *LBP* also has an impact on the response to radiotherapy.

The genes *BCAN*, *H2AFY* and *PPGB* were the genes with the highest PAM score in the modified gene list. These genes are supposed to best separate the before and after group. None of these genes were associated with the top five canonical pathways. *BCAN* and *H2AFY* contributed the network "Amino Acid Metabolism, Cellular Assembly and Organization, Connective Tissue Development and Function". This may indicate that the gene *PPGB*, which was significant different methylated between the before and after sample, has some unknown function or is involved in processes not adequate elucidated.

BCAN, also termed brevican, a proteglycan and a member of the subfamily of hyalectans, was significantly higher methylated in non-irradiated tumor tissue, which decreased after exposure to radiation doses over 16 Gy. Eight of the after samples did not show a decrease in methylation. Not surprisingly, those samples were exposed to radiation under 16Gy. An upregulation of brevican is associated with aggressive glial tumors (90). Aberrant regulation of this

gene is not described in breast cancer. The gene *H2AFY* which codes for the macroH2A1, is reported to be an important factor in base excision repair, but is also suggested to play a role in double strand breaks (91). However, in a previous study it appeared that the protein was not essential for radio-sensitivity; mice lacking the gene *H2AF* were irradiated with a single dose on 6.5Gy with a dose rate on 0,68Gy/min. The mice showed no hypersensitivity to IR (91). In this project, the average methylation degree of the gene *H2AFY* was on 18% in normal samples, 67% in non-radiated tumors and decreased to 48,9% in radiated tumor samples. However, seven of the samples showed an increased degree of methylation after radiation. These samples were the same which clustered as pair in the cluster analysis. Two samples showed no changes in the degree of methylation after radiation. These samples were separated in the cluster analysis, but are recorded with 8 and 9 radiation days, corresponding to 16 and 18Gy. After removing the paired after samples, the average methylation degree decreased to 35% in the after samples, leading to a significant change in methylation. The gene *H2AFY* should be further investigated in terms of a possible association with sensitivity to radiotherapy.

The gene *PPGB*, also termed CTSA, encodes for CathepsinA. Little is known about its biological function although in a previous study, CathepsinA was demonstrated to play an activating role in recruitment of factors involved in assembly of elastic fibers. In their mice models, CathepsinA deficient mice showed a decrease in elastic fibers in dermis (92). In this project, the gene *PPGB* showed an average methylation degree of 14% in normal samples, 61% in non-radiated tumors and 44% in radiated tumor samples. Seven of the after samples showed increased methylation after radiation. None of these after samples were radiated with more than 16Gy. After removing the paired after samples, the average methylation degree in the after samples decreased to 29%, leading to a significant difference in the methylation level when comparing the before and after samples. This may indicate a dose dependent change in methylation. This gene should be further investigated in a larger study, and also evaluated with regards to involvement in development of skin complications such as fibrosis. Unfortunately, this study had inadequate information about skin complications and development of fibrosis in the patients enrolled in this study.

It should also be noted that the different groups in the t-tests revealed a high standard deviation, meaning that there are high variance within the groups. This must be considered when interpreting the data. A Mann Whitney U-test was performed on the same genes, leading to a higher p-value but there was still a significant difference in the methylation level between the before and after samples for all genes analyzed.

In order to find a correlation between given dose and changes in the methylation level after radiation, the genes *IL1A*, *LBP*, *H2AFY*, *PPGB* and *BCAN* were included in a linear regression analysis. This statistical test revealed a significant relation between changes in methylation level after radiation exposure in the genes *IL1A*, *H2AFY* and *PPGB*, and the given dose. This is an interesting finding, because dose dependent changes in methylation pattern have so far not been elucidated. These results should be validated in an independent study.

6.2.4 Response groups

To reduce the number of groups and obtain groups of better sizes, the patients were combined into two groups. All patients who responded well to the radiotherapy were put into the good response group and all the other patients,

meaning those with partial, no response and progression were put into the poor response group. The basis was the hypothesis that the good responders would show the largest difference in methylation pattern after radiation, and this also gave more balanced groups with regards to size. This choice is likely to have affected the results since many mechanisms also would be differently regulated in those with partial response, compared to those with progression. Another solution would have been to term those with good and partial response as good responders and combine those reported with no response or progression into poor responders. This option was tried out in the SNP analysis discussed in chapter 6.3.3, leading to quite different results but due to the extensive amount of work to do the statistical and bioinformatic analysis this was not done for the methylation data in this thesis.

6.2.5 Methylation pattern and response

Patients undergoing radiotherapy respond differently to the administered treatment. This may be due to various reasons. Patients with good response to radiation treatment may have genes differentially methylated than those with poor response. This can affect both signaling pathways and networks.

In PAM analysis, a threshold of 2,2 was chosen, leading to a gene list with 342 genes. This threshold was chosen because it gave zero misclassified samples, and the highest number of genes.

The cluster analysis performed in J-express, separated the two response classes of interest quite good. Interestingly, sample number 117, clustered in the main branch with poor response, but in a subgroup alone. This sample is the only one reported with progression of the disease after treatment. The positioning of sample 117 indicate that some of the genes are similar methylated in sample 117 as for those reported with a reduced response, but that it also has some genes differentially methylated when compared to the rest of the poor response group. Sample number 113 and number 120 were misclassified with good response even though they are reported with no response.

The two top canonical pathways identified as significantly overrepresented in the gene list provided from PAM, G-protein coupled signaling and cAMP signaling, contained much of the same genes from the imported gene list. This is not surprising since they are involved in much of the same mechanisms. GPCRs can alter the intracellular concentration of cyclic AMP by activate or inactivate adenylyl cyclase (2). Activation of cAMP signaling has been shown to inhibit DNA induced apoptotic response to radiation in the leukemic BCP-ALL cells (93). The other three canonical pathways in the top 5 list “Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F”, “Chemokine signaling” and “Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza” may explain the role of inflammatory mediators influenced on the radiation response. Also, chemokines are known to mediate its effect through the G-protein signaling pathway. The genes *CCL2* and *CCL4* were related to all three pathways involved in the immune system, which may indicate that CCL2 and CCL4 are central in these pathways. On the other hand, it may also indicate that some of the pathways are on the top canonical list, because of these two genes. The gene *CCL2*, discussed below, was one of the genes most significant differently methylated in the response groups, and may be a good predictor of response. The genes *CCL1*, *CCL4*, *CCL7* and *CCL13* belonging to the same chemokine family, were also higher methylated in those with good response.

The FDRs for the canonical pathways indicates that 17-19% of the identified pathways may be false positive. This is a high FDR and ideally, the FDR should be under 5%. The ratio shows that the number of molecules from the imported gene list is low compared to the total number of molecules known for the different pathways. On the other hand, a high number of aberrantly methylated genes are not necessarily needed to achieve a dysregulated pathway.

Some of the same genes from these pathways are also involved in the network "Cell Morphology, Cellular Development and Nucleic Acid Metabolism". This network was supported by genes like *RASGRP3*, *RASSF6*, *RunX3* and *SFRP*, and contained 96 genes from the imported gene list. In the network analysis, CCL1, CCL2, CCL4, CCL8 and CCL13 are associated with "Antigen Presentation, Cellular Movement, Hematological System Development and Function".

Eight of the genes from the response gene list were selected for further investigation. The three genes on the top of the gene list (*ESRRB*, *MYCT1* and *CCL2*) were chosen because these genes are supposed to be the most differentially methylated genes between the two response groups. The genes *EGR2*, *RASGRP3, ROS1*, *RUNX3* and *SFRP2* were selected because they are previously reported in association with treatment response (94-98) and because they also came up in the pathway analysis. The gene list provided from PAM contained 342 genes selected because of their potential for separating the two response group and may contain many other interesting genes involved in radiation response. This should be further investigated in the continuation of this project but was not possible within the time frame of this project.

T-test was performed on the eight genes selected. The secreted frizzled-related protein2 (SFRP2) has been reported to be frequently downregulated by epigenetic silencing in breast cancer, leading to upregulation of the Wnt signaling (99). SFRP2 suppress proliferation in breast cancer cells and is suggested to play a role as a tumor suppressor gene (99). Also, the methylation status of the gene *SFRP2* was reported to correlate with tumor relapse in a previous study (98). Poor response to treatment may lead to earlier tumor relapse. These findings are in concordance with my results, where the gene was significantly higher methylated in those with poor response, compared with those with good response. Proliferation is one of the hallmarks of cancer and is one of the advantages a tumor cell must acquire. The genes *EGR2* and *RUNX2* were both significantly higher methylated in those with poor response, compared with those with good response. The transcription factor early growth receptor 2 (EGR2), is identified as a downstream mediator of p53 and promotes apoptosis (94). Radiation will normally increase the apoptosis rate, but tumor cells resistant to radiation are known to evade apoptosis, which may lead to survival and more malignant tumor cells. Cancer cells with the gene *EGR2* methylated, which is associated with gene silencing, may evade apoptosis and thereby become more resistant to radiotherapy. The gene RUNX2 is reported frequently methylated in breast cancer and may also be associated with survival (97). This is consistent with what was found in this study. Those with low percentage of methylation in the gene *RUNX2* showed a tendency to be more sensitive to radiation, which is associated with increased loco-regional control and survival.

The gene *ROS1*, reported as a proto-oncogene (96), and *RASGRP3*, a mediator in activating Ras (95), were both found to be significantly higher methylated in those reported with good response compared with the poor responders. Ras in known to participate in subset of pathways, including the Akt pathway, discussed in chapter 1.7.2. The Akt

pathway is involved in three main of the main mechanisms related to radiation response: DNA repair, repopulation and hypoxia, all shown to be involved in radiation resistance (25). Demethylation of *RASGRP3* may lead to increased signaling and further upregulation of the radiation response mechanisms leading to radiation resistance.

The three genes with highest score from the gene list (*ESRRB*, *MYCT1* and *CCL2*) provided from PAM, were significantly higher methylated in the group with good responders compared with poor responders. The *ESRRB* gene is an orphan nuclear receptor, which probably is involved in the ER-mediated pathways in the development of breast cancer (100). A target of the oncogene *C-MYC*, *MYCT1* also termed MTLC, also appears to be important in carcinogenesis. The gene *C-MYC* is known to be involved in a lot of mechanisms such as differentiation and cell cycle progression, but the role of MYCT1 is not clear. Interestingly, the gene *MYCT1* has been suggested to promote apoptosis in gastric carcinoma, though this mechanism is unclear. In this project, the gene *MYCT1* was associated with response, favored those with *MYCT1* methylated. The gene *MYCT1* was significant higher methylated in the good responders than in the poor responders. This confirms the oncogenic role for MYCT1, at least in breast cancer patients. The gene *CCL2*, reported as an inflammatory chemokine, associated as a cancer supporting factor, and may contribute to development and metastasis in breast cancer patients (101). In this study, the gene was found highly methylated in the group with good responders and less methylated in the poor responders. This finding was significant and indicates that it also contributes to treatment response. A high standard deviation indicates that the variance within the groups is high, and interpreting the results must be performed with caution. The genes were also tested for significance, using the non-parametric method Man-Whitney U- test, and leading to a higher, but still significant p-value.

6.3 P53 mutations frequencies

Approximately 20-40% of all breast cancers are found to be altered in the gene TP53 (30). 6 mutations were identified in the GenX material, giving a frequency of 27,3%, which corresponds with what is previously reported. There are only 22 patients in the GenX material, so the power is not strong enough to make any conclusions.

Sequencing the GenX material revealed single base substitutions in codon 175, 195, 254, 222 and 248. One of the samples had a single base deletion in codon 110. These mutations are within the regions reported in the literature to be most prone to mutations in *TP53* (31) .

The discrepancies revealed when comparing the cDNA sequencing results with gDNA results may be due to the fact that cDNA is processed and that gDNA holds more information than cDNA. On the other hand, the mutations revealed in cDNA should have been detected also in gDNA. A tumor is a mosaic structure containing different clones which can be differently amplified because of selection. In that way, cDNA does not always reflect all the information gDNA contains, and different tumor cells may harbor unequal genetic information. One should also bear into mind the fact that cDNA is converted from mRNA and some error may have occurred during the reverse transcriptase process.

6.3.1 P53 mutations, response and ER status

Several researchers have found a strong correlation between p53 mutations and poor response to radiation treatment (102). A mutation in *TP53* or in the p53 pathway may lead to a p53 protein unable to induce apoptosis or induction of repair. Langerød *et.al* demonstrated in their research that p53 status was a strong prognostic marker which may be useful for distinguishing between choices of treatment (103). Our data revealed no significant correlation between mutations in p53 or the p53 pathway and response when using Fisher's exact test. On the other hand, a trend of better response in those with WT was seen, compared with those harboring mutations. Unfortunately, the material was too small for detecting any significant correlation. Two of the samples with mutated *TP53* were described with good response to treatment. This may be explained by the fact that not all p53 mutations leads to equal effect on the protein function and level, and carriers of a heterozygote mutation still have the other not-mutated allele in the tumor, which may be functional. The four other *TP53* mutated samples were reported with partial, no response or progression. The latter was a sample containing a base pair deletion in *TP53* associated with an unfunctional protein.

The p53 mutation status is reported to correlate with ER expression, and both ER negative status and loss of functional p53 are related to poor prognosis. No significant association between ER status and p53/p53 pathway status was found in the GenX material. For detecting any significant correlation a larger material is required. Anyway, a relationship between p53 mutation and ER negative status is seen. The tendency is strongest when looking at the p53 pathway and ER status, with p= 0,179 using Fisher's exact test. This may be explained by the fact that *MDM2* is a member of the p53 pathway and ERα expression are reported to be involved in the transcription of *MDM2* (104).

6.3.2 P53 polymorphisms

Over 80 different polymorphisms in *TP53* have been identified and validated in human populations, where the majority are located in introns, outside splice sites, or in noncoding exons (31). Four types of polymorphisms were revealed in the GenX material when sequencing the *TP53* gene, two intronic and two exonic. The polymorphism in codon 72 has been reported to have an impact on prognosis in breast cancer patients (36). In the GenX material a polymorphism in codon 36 G>A was only seen together with a polymorphism in intron IVS 9+12 T>C. This was seen in one tumor sample (1/20) and one normal sample (1/11). Another polymorphism in intron IVS3-29 C>A was seen together with the heterozygote variant of codon 72 G/C in two of the tumor samples (2/20). The LD between two SNP's was not calculated due to the low number of samples.

Two of the samples (109 and 116) had different *TP53* c.72 polymorphism in blood (C/G) when compared with tumor samples (G/G and C/C). Sample 116 revealed a weak G allele on the reverse string, but sample 109 did not show the C allele in tumor. This can be explained with loss of heterozygosity (LOH) where one of the alleles is amplified and the other allele remains in only one copy or is completely lost. This is a common feature in genetic instable cancer cells (2).

6.3.3 SNPs related to response

Several researchers have reported a link between harboring unfavorable SNP alleles and an impaired treatment response and outcome in breast cancer patients (36). In order to find a correlation between response and harboring some of the unfavorable SNP alleles, a Chi-square analysis was performed. Unfortunately, a low number of cases did this analysis unreliable and Fisher's exact test was preferred, were only two independent variables are legal, meaning that the genotype groups had to be reduced into two instead of three groups. When looking at *MDM2* related to response (the modified response group) with Chi-square, a borderline significant Pearson P was revealed (p= 0.077). Because of few observations with the minor allele homozygote, this group was merged together with those observed heterozygote. Fisher's exact test (with two allele groups) did not reveal the same tendency. Still, this supports the idea that harboring G/G results in higher Mdm2 levels, leading to stronger regulative effect on the p53 pathway (29).

Sample number 106 had *MDM2* SNP309G/G, *MDM4* A/A, and a wild type of *TP53.* This patient was reported to have no response to treatment, which support our hypothesis that tumors harboring the unfavorable allele in *MDM2* SNP309 or *MDM4* may have the same effect on response as a mutated *TP53.* On the other hand, two of the patients are reported with good response despite a mutated *TP53* and three patients without a *TP53* mutation are reported with no response. Two of the samples in the latter example may be explained by a mutation in *PIK3CA*. As mentioned earlier, there are many ways to inactivate p53 and we are not investigating all the possibilities in this project.

Mutations in exon 9 and 20 in *PIK3CA* are related to influence the response to radiotherapy, because the activating role of PIK3CA in the PI3-K/AKT pathway associated with main mechanisms in development of radiation resistance. A trend towards better response when carrying the wild type, and no response or progression in those harboring a mutation, was uncovered when using Fisher's exact test although this did not reach statistical significance (p=0,131).

No associations between response and harboring the minor allele in codon72 or *MDM4* were identified. This is not surprising knowing there are conflicting evidence whether the SNP investigated in *MDM4* can predict radiation response or if it is only correlated to tumorigensis and early onset of cancer. Some studies have reported a relationship between the SNP in codon 72 and response to adjuvant chemotherapy, when also harboring the minor allele in *MDM2* (36). On the other hand, such interaction is also been referred to as not existing (38). None of the samples in this project harbored this combination of SNPs.

Despite the difficulties according to small materials, a tendency towards less response when harboring unfavorable SNPs in elements important in the p53 pathway or Akt pathway was revealed.

7 Conclusions

Response to radiation treatment is correlated with the cells ability to activate genes and pathways essential in the response to radiation. In this study we see clear differences in methylation patterns both connected to the treatment with radiation (in the before-after analysis) and response to radiation. Methylation level of selected subset of genes alone could separate the non-irradiated tissue from the irradiated tissue and also samples from individuals with good and poor response to treatment with low error rates. Methylation levels in molecules involved in the immune system pathways such as IL-10 and IL-6 signaling were demonstrated to be altered after exposure to radiation. The interleukin *IL1A* and the lipopolysaccharide binding protein *LBP* were involved in several of the inflammatory responses including acute phase response. The immune response signaling appeared to be important also in the response comparison group, but involving other genes and inflammatory pathways such as Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F" and "Chemokine signaling. The chemokine *CCL2* was shown to be involved in several of these pathways. A tendency toward less response to treatment was seen in those with genetic alterations in the gene *PIK3CA*. The Akt pathway is earlier pointed out as a crucial pathway involving at least three main mechanisms to radiation response, and some genes associated with this pathway such as *RASGRP3*, did show an altered methylation level after irradiation.

Apoptosis is another important mechanism for achieving good response to radiotherapy. It's well known that those with mutated *TP53* also respond less to treatment like radiotherapy. In this thesis, we saw a tendency toward less response to treatment when other elements regulating p53, like *MDM2* and *MDM4*, is overexpressed. Increased knowledge about the epigenetic regulation of genes contributes to an emerging understanding of how genes can be downregulated or upregulated without disrupting the genetic code. In this project the gene *EGR2*, involved in the p53 pathway and associated with apoptosis, where found highly methylated in samples collected from individuals showing less response to treatment. Also, the activation of cAMP signaling, found overrepresented in the genes identified analyzing the different response groups, is shown to inhibit p53-induced apoptosis by preventing p53 accumulation.

Many genetic and epigenetic alterations are seen in cancerous tissue, and some of them seem to be predictable for the response to radiation treatment. In addition, a dose dependent methylation level was seen in some of the genes identified. If one could be able to pinpoint those with altered pathways associated with resistant to radiation, other options for treatment could have been chosen. As the field targeted treatment emerges, a more personalized treatment in combination with radiotherapy is a promising aspect.

8 Future aspects / further plans

The material available for this project is unique because of the opportunity to compare paired samples collected before- and after- radiation. Unfortunately, too few samples lead to low power in the statistical and bioinformatic analysis. The findings should therefore be validated in an independent sample set. As a continuation of this project there are plans to study the alterations in methylation level caused by irradiation in a panel of fibroblast cell lines where DNA is available from before and after irradiation. Results from the methylation analysis should also be confirmed by the more quantitative pyrosequencing analysis. In order to unravel more of the complexity of epigenetic regulation of genes, some of the genes will be compared with available gene expression data for this material.

Reference List

(1) The Norwegian Cancer Registry. Cancer in Norway 2008. kreftregisteret 2009 [cited 2011 Jan 13];Available from: URL: http://www.kreftregisteret.no

(2) Alberts, Johnson, Lewis, Raff, Roberts, Walter. Molecular Biology of The Cell. fifth ed. New York: Garland Science; 2008.

(3) Polyak K. Breast cancer: origins and evolution. J Clin Invest 2007 Nov;117(11):3155-63.

(4) Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011 Mar 4;144(5):646-74.

(5) Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. Nature 2007 May 24;447(7143):433-40.

(6) Deroo BJ, Korach KS. Estrogen receptors and human disease. J Clin Invest 2006 Mar;116(3):561-70.

(7) Brisken C, O'Malley B. Hormone action in the mammary gland. Cold Spring Harb Perspect Biol 2010 Dec 1;2(12):a003178.

(8) Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2001 Feb;2(2):127-37.

(9) breastcancer.org. What is breastcancer? breastcancer org 2010 [cited 2011 Feb 2];Available from: URL: http://www.breastcancer.org/symptoms/understand_bc/what_is_bc.jsp

(10) Sorlie T. Molecular classification of breast tumors: toward improved diagnostics and treatments. Methods Mol Biol 2007;360:91-114.:91-114.

(11) Norsk Bryst Cancer Gruppe. Postoperativ strålebehandling. nbcg no 2010 [cited 2011 Feb 1];Available from: URL: http://nbcg.no/nbcg.blaaboka.html

(12) Kreftforeningen; Behandling. 2009 [cited 2010 Nov 20]; Available from: URL: http://www.kreftforeningen.no/om_kreft/behandling

(13) Aaltonen P, Brahme A, Lax I, Levernes S, Naslund I, Reitan JB, et al. Specification of dose delivery in radiation therapy. Recommendation by the Nordic Association of Clinical Physics (NACP). Acta Oncol 1997;36 Suppl 10:1-32.:1-32.

(14) Michael Joiner AvdK. Basic Clinical Radiobiology. Fourth ed. Hodder Arnold; 2009.

(15) Royal College of Surgeons of Ireland (RCSI). Lymph Nodes of Breast and Axilla (Labeled). healcentral org 2011Available from: URL: http://www.healcentral.org/content/collections/RCSI/sde146.JPG

(16) Thwaites D, Scalliet P, Leer JW, Overgaard J. Quality assurance in radiotherapy. European Society for Therapeutic Radiology and Oncology Advisory Report to the Commission of the European Union for the 'Europe Against Cancer Programme'. Radiother Oncol 1995 Apr;35(1):61-73.

(17) Medical-dictionary. medical-dictionary thefreedictionary com 2011 [cited 2010 Dec 2];Available from: URL: http://medical-dictionary.thefreedictionary.co

(18) Walsh N, Rheaume D, Barnes P, Tremaine R, Reardon M. Postirradiation morphea: an underrecognized complication of treatment for breast cancer. Hum Pathol 2008 Nov;39(11):1680-8.

(19) Ho AY, Atencio DP, Peters S, Stock RG, Formenti SC, Cesaretti JA, et al. Genetic predictors of adverse radiotherapy effects: the Gene-PARE project. Int J Radiat Oncol Biol Phys 2006 Jul 1;65(3):646-55.

(20) Bourguignon MH, Gisone PA, Perez MR, Michelin S, Dubner D, Giorgio MD, et al. Genetic and epigenetic features in radiation sensitivity Part I: cell signalling in radiation response. Eur J Nucl Med Mol Imaging 2005 Feb;32(2):229-46.

(21) Kåresen Rolf WE. Kreftsykdommer. first ed. Oslo: Gyldendal; 2000.

(22) Riley PA. Free radicals in biology: oxidative stress and the effects of ionizing radiation. Int J Radiat Biol 1994 Jan;65(1):27-33.

(23) Hall EJ. The bystander effect. Health Phys 2003 Jul;85(1):31-5.

(24) Eishi Noguchi. The Cell Cycle and Checkpoint Controls. Drexel University College of Medicine 2006 [cited 2011 Dec 15];Available from: URL: http://homepage.mac.com/enognog/checkpoint.html

(25) Schuurbiers OC, Kaanders JH, van der Heijden HF, Dekhuijzen RP, Oyen WJ, Bussink J. The PI3-K/AKT-pathway and radiation resistance mechanisms in non-small cell lung cancer. J Thorac Oncol 2009 Jun;4(6):761-7.

(26) Lane D, Levine A. p53 Research: the past thirty years and the next thirty years. Cold Spring Harb Perspect Biol 2010 Dec;2(12):a000893.

(27) Brown CJ, Lain S, Verma CS, Fersht AR, Lane DP. Awakening guardian angels: drugging the p53 pathway. Nat Rev Cancer 2009 Dec;9(12):862-73.

(28) Boehme KA, Blattner C. Regulation of p53--insights into a complex process. Crit Rev Biochem Mol Biol 2009 Nov;44(6):367-92.

(29) Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 2004 Nov 24;119(5):591-602.

(30) Borresen-Dale AL. TP53 and breast cancer. Hum Mutat 2003 Mar;21(3):292-300.

(31) Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol 2010 Jan;2(1):a001008.

(32) Mendrysa SM, McElwee MK, Michalowski J, O'Leary KA, Young KM, Perry ME. mdm2 Is critical for inhibition of p53 during lymphopoiesis and the response to ionizing irradiation. Mol Cell Biol 2003 Jan;23(2):462-72.

(33) Bouska A, Eischen CM. Murine double minute 2: p53-independent roads lead to genome instability or death. Trends Biochem Sci 2009 Jun;34(6):279-86.

(34) Kulkarni DA, Vazquez A, Haffty BG, Bandera EV, Hu W, Sun YY, et al. A polymorphic variant in human MDM4 associates with accelerated age of onset of estrogen receptor negative breast cancer. Carcinogenesis 2009 Nov;30(11):1910-5.

(35) Atwal GS, Kirchhoff T, Bond EE, Montagna M, Menin C, Bertorelle R, et al. Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. Proc Natl Acad Sci U S A 2009 Jun 23;106(25):10236-41.
(36) Toyama T, Zhang Z, Nishio M, Hamaguchi M, Kondo N, Iwase H, et al. Association of TP53 codon 72 polymorphism and the outcome of adjuvant therapy in breast cancer patients. Breast Cancer Res 2007;9(3):R34.
(37) Langerod A, Bukholm IR, Bregard A, Lonning PE, Andersen TI, Rognum TO, et al. The TP53 codon 72 polymorphism may affect the function of TP53 mutations in breast carcinomas but not in colorectal carcinomas. Cancer Epidemiol Biomarkers Prev 2002 Dec;11(12):1684-8.
(38) Knappskog S, Bjornslett M, Myklebust LM, Huijts PE, Vreeswijk MP, Edvardsen H, et al. The MDM2 Promoter SNP285C/309G Haplotype Diminishes Sp1 Transcription Factor Binding and Reduces Risk for Breast and Ovarian Cancer in Caucasians. Cancer Cell 2011 Feb 15;19(2):273-82.
(39) Tu HF, Chen HW, Kao SY, Lin SC, Liu CJ, Chang KW. MDM2 SNP 309 and p53 codon 72 polymorphisms are associated with the outcome of oral carcinoma patients receiving postoperative irradiation. Radiother Oncol 2008 May;87(2):243-52.
(40) Shirley SH, Rundhaug JE, Tian J, Cullinan-Ammann N, Lambertz I, Conti CJ, et al. Transcriptional regulation of estrogen receptor-alpha by p53 in human breast cancer cells. Cancer Res 2009 Apr 15;69(8):3405-14.
(41) Westphal CH, Hoyes KP, Canman CE, Huang X, Kastan MB, Hendry JH, et al. Loss of atm radiosensitizes multiple p53 null tissues. Cancer Res 1998 Dec 15;58(24):5637-9.
(42) Wang YV, Leblanc M, Wade M, Jochemsen AG, Wahl GM. Increased radioresistance and accelerated B cell lymphomas in mice with Mdmx mutations that prevent modifications by DNA-damage-activated kinases. Cancer Cell 2009 Jul 7;16(1):33-43.
(43) Angeloni SV, Martin MB, Garcia-Morales P, Castro-Galache MD, Ferragut JA, Saceda M. Regulation of estrogen receptor-alpha expression by the tumor suppressor gene p53 in MCF-7 cells. J Endocrinol 2004 Mar;180(3):497-504.
(44) Phelps M, Darley M, Primrose JN, Blaydes JP. p53-independent activation of the hdm2-P2 promoter through multiple transcription factor response elements results in elevated hdm2 expression in estrogen receptor alpha-positive breast cancer cells. Cancer Res 2003 May 15;63(10):2616-23.
(45) Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: its functions and alterations in human cancer. Apoptosis 2004 Nov;9(6):667-76.
(46) Samuels Y, Waldman T. Oncogenic mutations of PIK3CA in human cancers. Curr Top Microbiol Immunol 2010;347:21-41.:21-41.
(47) Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, et al. High frequency of mutations of the PIK3CA gene in human cancers. Science 2004 Apr 23;304(5670):554.
(48) Qiu W, Tong GX, Manolidis S, Close LG, Assaad AM, Su GH. Novel mutant-enriched sequencing identified high frequency of PIK3CA mutations in pharyngeal cancer. Int J Cancer 2008 Mar 1;122(5):1189-94.
(49) Abubaker J, Bavi P, Al-Haqawi W, Jehan Z, Munkarah A, Uddin S, et al. PIK3CA alterations in Middle Eastern ovarian cancers. Mol Cancer 2009 Jul 28;8:51.:51.
(50) Gupta AK, McKenna WG, Weber CN, Feldman MD, Goldsmith JD, Mick R, et al. Local recurrence in head and neck cancer: relationship to radiation resistance and signal transduction. Clin Cancer Res 2002 Mar;8(3):885-92.
(51) Schmitz KJ, Otterbach F, Callies R, Levkau B, Holscher M, Hoffmann O, et al. Prognostic relevance of activated Akt kinase in node-negative breast cancer: a clinicopathological study of 99 cases. Mod Pathol 2004 Jan;17(1):15-21.
(52) Kim TJ, Lee JW, Song SY, Choi JJ, Choi CH, Kim BG, et al. Increased expression of pAKT is associated with radiation resistance in cervical cancer. Br J Cancer 2006 Jun 5;94(11):1678-82.
(53) Tost J. DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker. Mol Biotechnol 2010 Jan;44(1):71-81.
(54) Kouzarides T. Chromatin modifications and their function. Cell 2007 Feb 23;128(4):693-705.
(55) Esteller M. Epigenetics in cancer. N Engl J Med 2008 Mar 13;358(11):1148-59.
(56) Lee WJ, Zhu BT. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. Carcinogenesis 2006 Feb;27(2):269-77.
(57) Toyota M, Suzuki H. Epigenetic drivers of genetic alterations. Adv Genet 2010;70:309-23.:309-23.
(58) Aypar U, Morgan WF, Baulch JE. Radiation-induced genomic instability: Are epigenetic mechanisms the missing link? Int J Radiat Biol 2011 Feb;87(2):179-91.
(59) Barekati Z, Radpour R, Kohler C, Zhang B, Toniolo P, Lenner P, et al. Methylation profile of TP53 regulatory pathway and mtDNA alterations in breast cancer patients lacking TP53 mutations. Hum Mol Genet 2010 Aug 1;19(15):2936-46.
(60) Sullivan A, Lu X. ASPP: a new family of oncogenes and tumour suppressor genes. Br J Cancer 2007 Jan 29;96(2):196-200.
(61) Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, et al. DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables. Mol Oncol 2011 Feb;5(1):77-92.
(62) Martens JW, Margossian AL, Schmitt M, Foekens J, Harbeck N. DNA methylation as a biomarker in breast cancer. Future Oncol 2009 Oct;5(8):1245-56.
(63) Ronneberg JA, Fleischer T, Solvang HK, Nordgard SH, Edvardsen H, Potapenko I, et al. Methylation profiling with a panel of cancer related genes: association with estrogen receptor, TP53 mutation status and expression subtypes in sporadic breast cancer. Mol Oncol 2011 Feb;5(1):61-76.
(64) Luzhna L, Kovalchuk O. Modulation of DNA methylation levels sensitizes doxorubicin-resistant breast adenocarcinoma cells to radiation-induced apoptosis. Biochem Biophys Res Commun 2010 Feb 5;392(2):113-7.

(65) Ma S, Liu X, Jiao B, Yang Y, Liu X. Low-dose radiation-induced responses: focusing on epigenetic regulation. Int J Radiat Biol 2010 Jul;86(7):517-28.
(66) Aypar U, Morgan WF, Baulch JE. Radiation-induced epigenetic alterations after low and high LET irradiations. Mutat Res 2010 Dec 13.
(67) Yoo CB, Jones PA. Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov 2006 Jan;5(1):37-50.
(68) Gravina GL, Festuccia C, Marampon F, Popov VM, Pestell RG, Zani BM, et al. Biological rationale for the use of DNA methyltransferase inhibitors as new strategy for modulation of tumor response to chemotherapy and radiation. Mol Cancer 2010 Nov 25;9:305.:305.
(69) 2011 Illumina I. Infinium Methylation Assay. illumina 2011 [cited 2010 Nov 21];Available from: URL: http://www.illumina.com/technology/infinium_methylation_assay.ilmn
(70) Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res 2001 Jul 1;29(13):E65.
(71) Campan M, Weisenberger DJ, Trinh B, Laird PW. MethyLight. Methods Mol Biol 2009;507:325-37.:325-37.
(72) Tibshirani HNC. Prediction Analysis for Microarrays. Stanford edu 2002 [cited 2011 Mar 11];Available from: URL: http://www-stat.stanford.edu/~tibs/PAM/
(73) Molmine. J-express. Molmine com 2009 [cited 2011 Apr 16];Available from: URL: http://www.microarray.no/magma/
(74) Ingenuity® Systems. Ingenuity Pathway Analysis. Ingenuity® Systems 2011 [cited 2011 May 3];Available from: URL: http://www.ingenuity.com
(75) Altman DG. Practical statistics for medical research. First ed. Boca Raton: Chapman&Hall; 1999.
(76) Yuan BZ, Zhou X, Durkin ME, Zimonjic DB, Gumundsdottir K, Eyfjord JE, et al. DLC-1 gene inhibits human breast cancer cell growth and in vivo tumorigenicity. Oncogene 2003 Jan 23;22(3):445-50.
(77) Teramoto A, Tsukuda K, Yano M, Toyooka S, Dote H, Doihara H, et al. Less frequent promoter hypermethylation of DLC-1 gene in primary breast cancers. Oncol Rep 2004 Jul;12(1):141-4.
(78) Tews B, Felsberg J, Hartmann C, Kunitz A, Hahn M, Toedt G, et al. Identification of novel oligodendroglioma-associated candidate tumor suppressor genes in 1p36 and 19q13 using microarray-based expression profiling. Int J Cancer 2006 Aug 15;119(4):792-800.
(79) Zhao G, Chen J, Deng Y, Gao F, Zhu J, Feng Z, et al. Identification of NDRG1-regulated genes associated with invasive potential in cervical and ovarian cancer cells. Biochem Biophys Res Commun 2011 Apr 2.
(80) Borghini S, Vargiolu M, Di DM, Ravazzolo R, Ceccherini I. Nuclear factor Y drives basal transcription of the human TLX3, a gene overexpressed in T-cell acute lymphocytic leukemia. Mol Cancer Res 2006 Sep;4(9):635-43.
(81) Hernandez-Vargas H, Lambert MP, Le Calvez-Kelm F, Gouysse G, McKay-Chopin S, Tavtigian SV, et al. Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors. PLoS One 2010 Mar 17;5(3):e9749.
(82) Ayoubi TA, Van De Ven WJ. Regulation of gene expression by alternative promoters. FASEB J 1996 Mar;10(4):453-60.
(83) Metivier R, Gallais R, Tiffoche C, Le PC, Jurkowska RZ, Carmouche RP, et al. Cyclical DNA methylation of a transcriptionally active promoter. Nature 2008 Mar 6;452(7183):45-50.
(84) Renard I, Joniau S, van CB, Collette C, Naome C, Vlassenbroeck I, et al. Identification and validation of the methylated TWIST1 and NID2 genes through real-time methylation-specific polymerase chain reaction assays for the noninvasive detection of primary bladder cancer in urine samples. Eur Urol 2010 Jul;58(1):96-104.
(85) Tommasi S, Karm DL, Wu X, Yen Y, Pfeifer GP. Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. Breast Cancer Res 2009;11(1):R14.
(86) Chen Y, Hyrien O, Williams J, Okunieff P, Smudzin T, Rubin P. Interleukin (IL)-1A and IL-6: applications to the predictive diagnostic testing of radiation pneumonitis. Int J Radiat Oncol Biol Phys 2005 May 1;62(1):260-6.
(87) Calveley VL, Khan MA, Yeung IW, Vandyk J, Hill RP. Partial volume rat lung irradiation: temporal fluctuations of in-field and out-of-field DNA damage and inflammatory cytokines following irradiation. Int J Radiat Biol 2005 Dec;81(12):887-99.
(88) Stein T, Morris JS, Davies CR, Weber-Hall SJ, Duffy MA, Heath VJ, et al. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. Breast Cancer Res 2004;6(2):R75-R91.
(89) Chen Z, Soo MY, Srinivasan N, Tan BK, Chan SH. Activation of macrophages by polysaccharide-protein complex from Lycium barbarum L. Phytother Res 2009 Aug;23(8):1116-22.
(90) Theocharis AD, Skandalis SS, Tzanakakis GN, Karamanos NK. Proteoglycans in health and disease: novel roles for proteoglycans in malignancy and their pharmacological targeting. FEBS J 2010 Oct;277(19):3904-23.
(91) Boulard M, Storck S, Cong R, Pinto R, Delage H, Bouvet P. Histone variant macroH2A1 deletion in mice causes female-specific steatosis. Epigenetics Chromatin 2010 Apr 1;3(1):8.
(92) Seyrantepe V, Hinek A, Peng J, Fedjaev M, Ernest S, Kadota Y, et al. Enzymatic activity of lysosomal carboxypeptidase (cathepsin) A is required for proper elastic fiber formation and inactivation of endothelin-1. Circulation 2008 Apr 15;117(15):1973-81.
(93) Naderi EH, Findley HW, Ruud E, Blomhoff HK, Naderi S. Activation of cAMP signaling inhibits DNA damage-induced apoptosis in BCP-ALL cells through abrogation of p53 accumulation. Blood 2009 Jul 16;114(3):608-18.
(94) Yokota I, Sasaki Y, Kashima L, Idogawa M, Tokino T. Identification and characterization of early growth response 2, a zinc-finger transcription factor, as a p53-regulated proapoptotic gene. Int J Oncol 2010 Dec;37(6):1407-16.
(95) Yang D, Kedei N, Li L, Tao J, Velasquez JF, Michalowski AM, et al. RasGRP3 contributes to formation and maintenance of the prostate cancer phenotype. Cancer Res 2010 Oct 15;70(20):7905-17.

(96) Acquaviva J, Wong R, Charest A. The multifaceted roles of the receptor tyrosine kinase ROS in development and cancer. Biochim Biophys Acta 2009 Jan;1795(1):37-52.
(97) Jiang Y, Tong D, Lou G, Zhang Y, Geng J. Expression of RUNX3 gene, methylation status and clinicopathological significance in breast cancer and breast cancer cell lines. Pathobiology 2008;75(4):244-51.
(98) Hill VK, Ricketts C, Bieche I, Vacher S, Gentle D, Lewis C, et al. Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity. Cancer Res 2011 Mar 3.
(99) Suzuki H, Toyota M, Carraway H, Gabrielson E, Ohmura T, Fujikane T, et al. Frequent epigenetic inactivation of Wnt antagonist genes in breast cancer. Br J Cancer 2008 Mar 25;98(6):1147-56.
(100) Lu D, Kiriyama Y, Lee KY, Giguere V. Transcriptional regulation of the estrogen-inducible pS2 breast cancer marker gene by the ERR family of orphan nuclear receptors. Cancer Res 2001 Sep 15;61(18):6755-61.
(101) Soria G, Ofri-Shahak M, Haas I, Yaal-Hahoshen N, Trejo-Leider L, Leibovich-Rivkin T, et al. Inflammatory mediators in breast cancer: Coordinated expression of TNFalpha & IL-1beta with CCL2 & CCL5 and effects on epithelial-to-mesenchymal transition. BMC Cancer 2011 Apr 12;11(1):130.
(102) Fei P, El-Deiry WS. P53 and radiation responses. Oncogene 2003 Sep 1;22(37):5774-83.
(103) Langerod A, Zhao H, Borgan O, Nesland JM, Bukholm IR, Ikdahl T, et al. TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. Breast Cancer Res 2007;9(3):R30.
(104) Bond GL, Hirshfield KM, Kirchhoff T, Alexe G, Bond EE, Robins H, et al. MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner. Cancer Res 2006 May 15;66(10):5104-10.

APPENDIX A Supplementary tables and figures

Table 19 analyzing the difference between normal tissue and breast cancer tissue, the samples were divided into two groups. The normal group contained the normal samples and the before group contained non-radiated tumor samples. A gene list with 14 genes was provided from PAM at a threshold of 5.8. The genes on the top have the highest score and are those which are most differentially methylated between the groups of interests.

ID	Genes	Before-score	Normal-score
1	DLC1_cg05226008	-0.2147	0.4771
2	MGC4399_cg13631913	-0.1937	0.4304
3	LTV1_cg02885771	0.191	-0.4245
4	KA35_cg09088193	0.1103	-0.2452
5	NR2E1_cg03958979	0.0855	-0.1901
6	SRC_cg22437284	0.0552	-0.1227
7	HLXB9_cg20420433	-0.0438	0.0974
8	NID2_cg22881914	0.0371	-0.0824
9	C9orf45_cg07763768	0.0347	-0.0772
10	HDC_cg02329886	0.0311	-0.069
11	TLX3_cg25720804	0.0163	-0.0361
12	INA_cg25764191	0.0126	-0.0281
13	H2AFY_cg24628744	0.0081	-0.018
14	DDR2_cg22740835	0.002	-0.0044

Table 20 The tumor samples were divided into two groups, the before group containing samples drawn before radiation, and the after group containing samples drawn after radiation. In the PAM analysis a threshold of 1.7 resulted in a list of 142 genes differentially methylated between the two groups.

ID	Genes	Score af	Score be	ID	Genes	Score af	Score be
1	HTR3D_cg14483391	-0.0919	0.0874	71	ALDH1A3_cg19224278	-0.0206	0.0196
2	RRAD_cg19428417	-0.0845	0.0803	72	RGS13_cg05023691	0.0206	-0.0196
3	KSP37_cg08132711	0.0792	-0.0753	73	PTPN22_cg00916635	-0.0205	0.0194
4	GALNT1_cg05714729	-0.0789	0.0749	74	ABCC11_cg08446111	0.0201	-0.019
5	VSIG4_cg26561773	0.0789	-0.0749	75	SLC18A1_cg22441882	0.02	-0.019
6	IL1R2_cg17142183	-0.0788	0.0749	76	RAB11FIP4_cg04764624	-0.0194	0.0185
7	ARMC3_cg11673092	-0.0776	0.0738	77	OR10J1_cg25076881	0.0188	-0.0178
8	KRTHA6_cg02780988	-0.0741	0.0704	78	PHLDA1_cg27182761	-0.0184	0.0175
9	MBNL1_cg12360736	-0.074	0.0703	79	ELAVL4_cg26227005	-0.0181	0.0172
10	H2AFY_cg24628744	-0.0736	0.0699	80	FGF19_cg26096837	-0.018	0.0171
11	GNB5_cg14120436	-0.061	0.058	81	TMEM109_cg10735607	0.0179	-0.017
12	GYPE_cg13143729	0.0598	-0.0568	82	ABHD8_cg08145177	-0.0176	0.0167
13	GCET2_cg25462303	-0.0556	0.0528	83	ABL2_cg18433086	-0.0175	0.0167
14	GPR123_cg21607649	0.0549	-0.0522	84	LOC342897_cg10092957	-0.0174	0.0166
15	CD8A_cg02170525	-0.0548	0.0521	85	RASGRP2_cg14170423	-0.0174	0.0165
16	LBP_cg18979491	0.0547	-0.052	86	POR_cg20748065	-0.0173	0.0164
17	BCAN_cg21475402	-0.0533	0.0506	87	SLC39A12_cg19856444	0.0172	-0.0163
18	MARCH1_cg07259382	0.0492	-0.0468	88	SPRR2A_cg26059632	0.0166	-0.0158
19	SLAMF8_cg04275881	-0.0481	0.0457	89	MS4A7_cg10853416	-0.0164	0.0156
20	IL5RA_cg08404225	0.0474	-0.045	90	37135_cg19663795	-0.0158	0.015
21	HESX1_cg10608341	-0.0466	0.0443	91	TNFRSF1B_cg26189983	-0.0156	0.0148
22	PNOC_cg19391527	0.0445	-0.0423	92	SLC9A11_cg15975283	0.0155	-0.0148
23	FGG_cg01593385	0.0438	-0.0417	93	GPX2_cg20764656	-0.0153	0.0146
24	KRTAP21-2_cg23581186	0.0434	-0.0412	94	P2RY12_cg05094216	-0.0151	0.0144

25	FAM79B_cg19682367	0.0433	-0.0411	95	SLC7A3_cg20622056	-0.0145	0.0138
26	FLJ20184_cg25370441	0.0413	-0.0392	96	FLJ31951_cg14686321	-0.0143	0.0136
27	HIST1H4L_cg10536916	-0.0408	0.0387	97	PPGB_cg19067730	-0.0138	0.0131
28	GALNAC4S-6ST_cg08082692	-0.0401	0.0381	98	SLC17A8_cg06836849	-0.0132	0.0126
29	CTSZ_cg01623438	-0.0396	0.0377	99	LILRB1_cg01720520	0.0124	-0.0118
30	GARNL3_cg00342530	0.0395	-0.0376	100	C4orf17_cg26647453	0.012	-0.0114
31	DYSF_cg15491567	-0.0389	0.037	101	CORIN_cg26018901	-0.0105	0.01
32	IL21R_cg19423311	-0.0384	0.0365	102	SPOCK2_cg10983208	-0.0102	0.0097
33	PRELP_cg07947930	0.037	-0.0352	103	DNAJC6_cg09082287	-0.0102	0.0097
34	TNXB_cg13823701	0.037	-0.0352	104	BAPX1_cg20073553	-0.0102	0.0097
35	TRIM38_cg22502502	-0.0366	0.0348	105	NPTX1_cg17775235	-0.0099	0.0094
36	SERPINA3_cg06190732	0.0364	-0.0346	106	HTR2A_cg00308665	0.0098	-0.0093
37	MBNL1_cg14423778	-0.036	0.0342	107	LBR_cg14607642	-0.0093	0.0088
38	TFF2_cg12456510	0.036	-0.0342	108	DLC1_cg00933411	0.0093	-0.0088
39	PCK1_cg13904968	0.0356	-0.0338	109	ATP4A_cg06123346	0.0084	-0.008
40	PART1_cg18704047	0.0353	-0.0336	110	MS4A2_cg10414946	0.0081	-0.0077
41	CATSPER1_cg14894216	0.0351	-0.0334	111	ZNF502_cg21672276	-0.0081	0.0077
42	PDCD1_cg00795812	-0.0318	0.0302	112	LTC4S_cg11394785	-0.0079	0.0075
43	STAT5A_cg03001305	-0.0318	0.0302	113	DTL_cg03938043	0.0078	-0.0074
44	CCR1_cg13144783	-0.0317	0.0301	114	ITGB1BP1_cg07974891	0.0077	-0.0073
45	OR7C1_cg11328541	0.0317	-0.0301	115	SERPINE2_cg00514407	-0.0074	0.0071
46	SERPINB7_cg17251713	0.0317	-0.0301	116	CD79A_cg04790874	-0.0071	0.0068
47	_cg13279585	0.0313	-0.0297	117	TEX15_cg20939319	0.0068	-0.0065
48	BNIPL_cg11584936	0.0313	-0.0297	118	HEM1_cg17605084	-0.0056	0.0053
49	NCF1_cg17468997	-0.0305	0.029	119	NCR1_cg12952132	0.0054	-0.0051
50	MAGEA5_cg06313930	0.0305	-0.029	120	BPIL1_cg13696012	0.0054	-0.0051
51	FAM78A_cg17936488	-0.0302	0.0287	121	CD248_cg00350296	-0.0054	0.0051
52	SCAND2_cg17866455	0.0301	-0.0286	122	SLC2A5_cg24480859	-0.0046	0.0044
53	PIGC_cg08587864	0.0294	-0.028	123	ANGPTL2_cg09427311	-0.0045	0.0043
54	DSG1_cg01337047	0.0289	-0.0274	124	C1orf182_cg24042452	0.0044	-0.0042
55	NMNAT3_cg01724150	-0.0279	0.0266	125	SNAPAP_cg07576541	0.0038	-0.0036
56	MMP10_cg00347729	0.027	-0.0257	126	AOC3_cg21602160	0.0033	-0.0031
57	TANK_cg23871659	-0.0265	0.0252	127	SPOCK3_cg06021171	0.0032	-0.003
58	PIK3CD_cg20994801	-0.0264	0.0251	128	CCT6A_cg23839680	-0.003	0.0028
59	CHML_cg15775914	-0.0262	0.0249	129	IL1B_cg07935264	-0.0027	0.0026
60	PTPN22_cg14385738	-0.0261	0.0248	130	LTC4S_cg16361890	-0.0024	0.0022
61	HRASLS5_cg00754253	-0.0261	0.0248	131	PODN_cg16028753	-0.0022	0.0021
62	MEGF10_cg26465611	-0.0254	0.0242	132	SLC15A2_cg18636558	0.0021	-0.002
63	ITGAM_cg15337006	-0.0253	0.0241	133	C1orf36_cg00658007	-0.0018	0.0017
64	MSX1_cg09573795	-0.0253	0.024	134	CCL4L2_cg15129294	0.0018	-0.0017
65	FXC1_cg25219333	0.0244	-0.0232	135	IGSF2_cg23953831	-0.0018	0.0017
66	TREML2_cg19005210	-0.024	0.0228	136	UBD_cg07326586	0.0015	-0.0014
67	FCER1G_cg13853198	-0.0226	0.0215	137	TNFAIP8L2_cg23612220	-0.0015	0.0014
68	MYF6_cg26711820	-0.0218	0.0207	138	NALP10_cg18484189	8,00E-04	-8,00E-04
69	SPI1_cg06147863	-0.0213	0.0202	139	SH3TC1_cg07816074	-5,00E-04	4,00E-04
70	OR51B4_cg06353345	0.021	-0.0199	140	UNC5D_cg00297600	4,00E-04	-4,00E-04

Table 21 After removing eight of the after samples, a new list was extracted from PAM. 84 genes were differentially methylated between the two groups at a threshold of 2.8.

ID	Genes	Score af	score be	ID	Genes	Score af	score be
1	BCAN_cg21475402	-0.2789	0.1534	43	CATSPER1_cg14894216	0.0362	-0.0199
2	H2AFY_cg24628744	-0.2506	0.1378	44	IL20_cg01103730	0.0331	-0.0182
3	PPGB_cg19067730	-0.1609	0.0885	45	SNAPAP_cg07576541	0.0316	-0.0174
4	PCK1_cg13904968	0.1524	-0.0838	46	_cg15051063	0.0286	-0.0157
5	_cg13279585	0.1443	-0.0794	47	LOC116123_cg24272559	0.0285	-0.0157
6	LTC4S_cg11394785	-0.1149	0.0632	48	REG3A_cg24240626	0.0273	-0.015
7	FAM79B_cg19682367	0.1134	-0.0624	49	BACE2_cg16334795	-0.027	0.0148
8	IL5RA_cg08404225	0.1109	-0.061	50	TAL2_cg06119575	0.0269	-0.0148
9	BNIPL_cg11584936	0.0981	-0.0539	51	RBP1_cg23363832	-0.0252	0.0138
10	GYPE_cg13143729	0.0971	-0.0534	52	GRIK5_cg09555879	0.0225	-0.0124
11	KSP37_cg08132711	0.0956	-0.0526	53	C1orf114_cg13958426	-0.0224	0.0123
12	TRDN_cg14462830	0.0942	-0.0518	54	OR12D3_cg20856834	0.0216	-0.0119
13	SLC2A10_cg27610561	0.0937	-0.0515	55	BCAS1_cg08927738	0.0205	-0.0113
14	SPOCK2_cg10983208	-0.0883	0.0485	56	ENPP2_cg14409958	-0.0197	0.0108
15	FGG_cg01593385	0.0839	-0.0461	57	DOPEY2_cg00673191	0.0148	-0.0081
16	GNB5_cg14120436	-0.0821	0.0452	58	KRTAP13-4_cg14062083	0.0138	-0.0076
17	HIST1H4L_cg10536916	-0.0801	0.044	59	ALDH1A3_cg27652350	-0.0138	0.0076
18	BAPX1_cg20073553	-0.0769	0.0423	60	RETNLB_cg14659547	0.0126	-0.0069
19	FXC1_cg25219333	0.0753	-0.0414	61	EPB41L1_cg20993403	0.0121	-0.0067
20	CNTN4_cg10503138	0.0746	-0.041	62	IL1A_cg00839584	-0.0118	0.0065
21	TINAG_cg12397274	0.0702	-0.0386	63	BRS3_cg15016628	0.011	-0.006
22	IL1R2_cg17142183	-0.0699	0.0384	64	MMP10_cg00347729	0.0107	-0.0059
23	BPIL1_cg13696012	0.0631	-0.0347	65	KRTAP11-1_cg22643217	0.0105	-0.0058
24	ZNF436_cg10347418	0.0602	-0.0331	66	GPLD1_cg14023451	0.0098	-0.0054
25	SCAND2_cg17866455	0.0581	-0.0319	67	MS4A2_cg10414946	0.008	-0.0044
26	PPGB_cg08260891	-0.0559	0.0307	68	PNOC_cg19391527	0.0078	-0.0043
27	RRAD_cg19428417	-0.0553	0.0304	69	TMPRSS11F_cg02936740	0.0076	-0.0042
28	PART1_cg18704047	0.0548	-0.0301	70	FLJ42393_cg21909391	0.0075	-0.0041
29	KRTAP13-1_cg02764897	0.0542	-0.0298	71	UBD_cg07326586	0.0071	-0.0039
30	C6_cg11976616	0.0531	-0.0292	72	GYS2_cg06141025	0.0065	-0.0036
31	MAGEA5_cg06313930	0.0489	-0.0269	73	SLAMF8_cg04275881	-0.0059	0.0032
32	SPATA19_cg18457737	0.0475	-0.0261	74	PRG3_cg24459209	0.0057	-0.0031
33	ZNF80_cg03109316	0.0468	-0.0257	75	SERPINB12_cg03468463	0.004	-0.0022
34	REG3A_cg27342801	0.0466	-0.0256	76	CLEC2A_cg27190239	0.0036	-0.002
35	DOC1_cg06436504	0.0465	-0.0256	77	ATP4A_cg06123346	0.0032	-0.0018
36	RGPD5_cg02148642	0.0459	-0.0252	78	TCP10_cg07256847	0.0029	-0.0016
37	LTC4S_cg16361890	-0.0434	0.0239	79	C12orf34_cg01335367	-0.0026	0.0014
38	TMEM109_cg10735607	0.0387	-0.0213	80	DNAJC6_cg09082287	-0.0023	0.0012
39	BPIL1_cg10968815	0.0386	-0.0212	81	KRTAP11-1_cg07014174	0.0019	-0.001
40	GALNT1_cg05714729	-0.0385	0.0212	82	PTPN22_cg14385738	-0.0018	0.001
41	MPI_cg13828047	-0.038	0.0209	83	LBP_cg18979491	0.001	-6,00E-04
42	SLCO1C1_cg18109798	0.0363	-0.02	84	FLJ25410_cg05215575	0.001	-5,00E-04

Table 22 The samples were divided into two groups according to response to treatment. A gene list was extracted from PAM, containing 342 genes differentially methylated between the two groups at a threshold of 2.2.

id	Gene	Resp— score	Resp+- score	id	Gene	Resp— score	Resp+- score
[1,]	ESRRB_cg07864297	-0.5033	0.4026	[175,]	FLJ40235_cg21930712	-0.062	0.0496
[2,]	MYCT1_cg02830467	-0.4699	0.3759	[176,]	C6orf118_cg02064106	0.0619	-0.0495
[3,]	CCL2_cg21109025	-0.4377	0.3501	[177,]	F2RL2_cg00415993	0.0618	-0.0494
[4,]	SIRPD_cg25737664	-0.4342	0.3474	[178,]	KCNA10_cg07830847	-0.0617	0.0493
[5,]	GABRA6_cg07592353	-0.395	0.316	[179,]	RASGRP3_cg01109219	-0.0613	0.0491
[6,]	PRODH2_cg27404050	-0.3941	0.3153	[180,]	C15orf32_cg17978274	-0.0604	0.0484
[7,]	MASP1_cg21831174	-0.3611	0.2889	[181,]	PRICKLE2_cg23047271	0.0598	-0.0479
[8,]	PRSS1_cg12878228	-0.3427	0.2742	[182,]	OR2W1_cg05779068	-0.0596	0.0477
[9,]	TUB_cg05492113	-0.3409	0.2727	[183,]	LRTM1_cg11532513	-0.0585	0.0468
[10,]	CLDN16_cg27235662	-0.3398	0.2718	[184,]	ZNF157_cg03957435	-0.0583	0.0466
[11,]	PABPC5_cg04875162	0.3368	-0.2694	[185,]	VNN1_cg02184413	-0.0576	0.0461
[12,]	_cg23984130	-0.3276	0.2621	[186,]	PAX4_cg08886154	-0.0572	0.0458
[13,]	TUB_cg15480475	-0.3206	0.2564	[187,]	PFKM_cg27446233	-0.0572	0.0458
[14,]	OR10H1_cg24477636	-0.2929	0.2344	[188,]	TMEM22_cg02672493	0.0569	-0.0455
[15,]	CRTAM_cg10977115	-0.2862	0.229	[189,]	NOS3_cg25007250	-0.0563	0.0451
[16,]	CCNA1_cg16422907	0.2856	-0.2285	[190,]	ADCY4_cg12265829	0.0563	-0.0451
[17,]	DDC_cg04144768	-0.2821	0.2257	[191,]	CCL4_cg25659818	-0.0555	0.0444
[18,]	PTGFR_cg03495868	0.2799	-0.2239	[192,]	FATE1_cg08177445	-0.0553	0.0442
[19,]	TAB3_cg14186071	-0.2764	0.2211	[193,]	MOV10L1_cg18638931	0.0551	-0.0441
[20,]	CRTAM_cg22512531	-0.2746	0.2197	[194,]	MARCO_cg11009736	-0.0544	0.0435
[21,]	OR8B8_cg14620221	-0.2599	0.2079	[195,]	MRGPRX4_cg09691574	-0.0535	0.0428
[22,]	GFI1_cg22341104	0.2519	-0.2015	[196,]	MRGPRX1_cg24252809	-0.0535	0.0428
[23,]	CNR1_cg23276695	-0.2487	0.199	[197,]	ACMSD_cg02812142	-0.0532	0.0425
[24,]	PAQR9_cg00970325	0.2454	-0.1963	[198,]	ANXA1_cg01894895	-0.0531	0.0425
[25,]	FLJ27255_cg13126790	-0.2453	0.1962	[199,]	ATXN3_cg02028524	-0.0528	0.0422
[26,]	SLC6A15_cg03064067	0.243	-0.1944	[200,]	RAB32_cg20098887	0.052	-0.0416
[27,]	GABRA5_cg24387380	-0.2347	0.1878	[201,]	FBN2_cg27223047	0.0519	-0.0416
[28,]	C15orf32_cg25455753	-0.2338	0.1871	[202,]	SSX2_cg03712237	-0.0512	0.0409
[29,]	CCL13_cg02706575	-0.2271	0.1817	[203,]	SIGLEC7_cg23458892	-0.0511	0.0409
[30,]	COL23A1_cg10730712	0.2257	-0.1806	[204,]	CPNE8_cg23495733	0.0501	-0.0401
[31,]	ATP2B2_cg14547335	-0.2205	0.1764	[205,]	MAPK4_cg26946769	-0.0495	0.0396
[32,]	OTUD7_cg01421985	-0.2202	0.1762	[206,]	SPOCK2_cg10983208	0.0495	-0.0396
[33,]	CCL8_cg27000831	-0.2174	0.1739	[207,]	ROS1_cg21166999	-0.0493	0.0394
[34,]	FLJ25530_cg13018903	-0.2117	0.1694	[208,]	TPSD1_cg01375871	-0.0483	0.0387
[35,]	DYDC1_cg17703212	0.2099	-0.168	[209,]	VSIG9_cg13669740	-0.0477	0.0381
[36,]	ZNF264_cg16636110	0.2079	-0.1664	[210,]	PCDHAC1_cg12629325	0.0473	-0.0378
[37,]	CALD1_cg02382666	-0.2048	0.1638	[211,]	OR10A5_cg22951794	-0.0472	0.0378
[38,]	PDYN_cg14944362	-0.196	0.1568	[212,]	PRAMEF1_cg09196942	-0.0458	0.0367
[39,]	EGR2_cg19355190	0.1929	-0.1544	[213,]	_cg00718513	-0.0454	0.0363
[40,]	PRSS2_cg04958389	-0.1906	0.1524	[214,]	GYPB_cg27214365	-0.0445	0.0356
[41,]	ADRA1B_cg21575929	-0.189	0.1512	[215,]	XKR6_cg10947146	0.0441	-0.0353
[42,]	XAGE5_cg25993152	-0.1884	0.1507	[216,]	GYPA_cg09841009	-0.0441	0.0353
[43,]	EPDR1_cg27641018	-0.1874	0.1499	[217,]	CX62_cg25217765	-0.0433	0.0347
[44,]	KRTAP13-4_cg14062083	-0.1847	0.1478	[218,]	CLEC4M_cg01532771	-0.0429	0.0343

[45,]	GABRA6_cg22672790	-0.1825	0.146	[219,]	OR7A17_cg04645174	-0.0424	0.0339
[46,]	MRGPRX1_cg23001457	-0.182	0.1456	[220,]	CSPG3_cg13118849	-0.0419	0.0335
[47,]	TRY1_cg14153740	-0.1815	0.1452	[221,]	HLA-G_cg21529533	0.0413	-0.0331
[48,]	CX62_cg08947964	-0.1747	0.1398	[222,]	CNGA2_cg17032587	-0.0404	0.0323
[49,]	FGF2_cg17214107	0.1743	-0.1395	[223,]	IFNA14_cg00474004	-0.0399	0.0319
[50,]	KLHL21_cg19884658	0.1737	-0.139	[224,]	RSNL2_cg21972382	0.0394	-0.0316
[51,]	SLN_cg17971003	-0.1711	0.1369	[225,]	GUCY2F_cg22901146	-0.0389	0.0311
[52,]	FOXG1B_cg02681442	0.1708	-0.1366	[226,]	CDK9_cg08999352	-0.0386	0.0308
[53,]	OR8B8_cg16612699	-0.1673	0.1339	[227,]	GUCY1A3_cg02210887	-0.0385	0.0308
[54,]	NELL1_cg17371081	0.1642	-0.1314	[228,]	FCN1_cg17357062	-0.0383	0.0307
[55,]	BTNL2_cg25391023	-0.1613	0.129	[229,]	VCX_cg09018040	-0.0383	0.0306
[56,]	LILRA2_cg19486673	-0.1592	0.1274	[230,]	ABR_cg25374854	0.0372	-0.0298
[57,]	MUC15_cg03087937	-0.159	0.1272	[231,]	SSX3_cg13601079	-0.0372	0.0298
[58,]	FOXB1_cg08583049	0.1583	-0.1267	[232,]	FLJ43582_cg11653466	0.037	-0.0296
[59,]	FATE1_cg01423840	-0.157	0.1256	[233,]	SPATA8_cg02423618	-0.037	0.0296
[60,]	ACOT8_cg08101264	0.157	-0.1256	[234,]	REG1B_cg00579393	-0.0368	0.0295
[61,]	ADAMTS8_cg01033938	0.1555	-0.1244	[235,]	ERAF_cg14387505	-0.0366	0.0293
[62,]	OR3A3_cg05674036	-0.155	0.124	[236,]	MARCO_cg02431964	-0.0365	0.0292
[63,]	NEUROD1_cg22359606	0.155	-0.124	[237,]	SYCP2_cg22214414	-0.0365	0.0292
[64,]	LUZP4_cg08693325	-0.1547	0.1238	[238,]	APOH_cg19058765	-0.0365	0.0292
[65,]	CLDN17_cg13792279	-0.1536	0.1229	[239,]	ABCA3_cg00949442	0.0357	-0.0285
[66,]	PRSS2_cg13944141	-0.1527	0.1221	[240,]	PLS3_cg08990057	0.0351	-0.0281
[67,]	CNGA2_cg04613080	-0.1517	0.1214	[241,]	AMOT_cg18389752	0.0348	-0.0278
[68,]	FAM38B_cg21165219	-0.1494	0.1195	[242,]	SYCP2_cg07347645	-0.0347	0.0278
[69,]	LOC283537_cg02233559	0.1493	-0.1194	[243,]	DEFB126_cg20305726	-0.0347	0.0277
[70,]	CUBN_cg10707565	-0.1491	0.1193	[244,]	SPP2_cg21137417	-0.0346	0.0277
[71,]	PROX1_cg22176895	0.1482	-0.1186	[245,]	SLC1A6_cg16377872	-0.0337	0.027
[72,]	KIAA1913_cg18172186	-0.1482	0.1185	[246,]	ASPA_cg07732644	-0.0335	0.0268
[73,]	C6orf118_cg05799317	-0.1456	0.1165	[247,]	MYOD1_cg18555440	0.0323	-0.0258
[74,]	PLEKHA4_cg08077345	0.1453	-0.1162	[248,]	UTS2D_cg11158430	-0.0315	0.0252
[75,]	SIRPB1_cg09577651	-0.1448	0.1158	[249,]	KRT2A_cg00463848	-0.0314	0.0251
[76,]	SH2D3C_cg26920757	0.1446	-0.1157	[250,]	OR12D3_cg20856834	-0.031	0.0248
[77,]	MGAM_cg01476044	-0.1415	0.1132	[251,]	PCDHAC1_cg18902090	0.0306	-0.0245
[78,]	OR1G1_cg06882926	-0.1404	0.1123	[252,]	CLTCL1_cg07251788	0.0304	-0.0243
[79,]	COL21A1_cg13830624	0.1376	-0.1101	[253,]	TUBB6_cg07307078	0.0302	-0.0241
[80,]	PGBD5_cg19560210	-0.1373	0.1098	[254,]	SSX8_cg00962799	-0.0299	0.0239
[81,]	SIRPB2_cg11061975	-0.1372	0.1098	[255,]	RASSF6_cg03996822	-0.0293	0.0234
[82,]	MS4A3_cg17173423	-0.1369	0.1095	[256,]	SIRPD_cg17423978	-0.0291	0.0233
[83,]	ANKRD33_cg19948393	0.1365	-0.1092	[257,]	UNQ739_cg00333226	0.0288	-0.0231
[84,]	ARHGAP4_cg06791102	0.1344	-0.1075	[258,]	CSAG2_cg03033367	-0.0287	0.023
[85,]	HBE1_cg08970694	-0.1328	0.1063	[259,]	AHSG_cg07361385	-0.0285	0.0228
[86,]	ELSPBP1_cg19404979	-0.1317	0.1054	[260,]	GRM8_cg02946850	-0.0278	0.0223
[87,]	SCN3B_cg15457899	0.1315	-0.1052	[261,]	OR1D2_cg02721374	-0.0276	0.0221
[88,]	EDG2_cg14563260	0.1315	-0.1052	[262,]	RBM13_cg23663332	0.0275	-0.022
[89,]	SSX8_cg01632517	-0.1311	0.1049	[263,]	MSX1_cg03199651	0.0275	-0.022
[90,]	NEUROD1_cg02836529	0.1304	-0.1043	[264,]	NXF2_cg00280894	-0.0272	0.0218
[91,]	CDX2_cg02055963	0.1276	-0.1021	[265,]	SFRP2_cg23207990	0.0264	-0.0211

[92,]	GYPE_cg16998872	-0.1262	0.101	[266,]	GDA_cg18286698	-0.0264	0.0211
[93,]	GABRA5_cg08099701	-0.1258	0.1007	[267,]	EFCBP2_cg02899772	0.0262	-0.0209
[94,]	HSPB3_cg20169062	-0.1246	0.0996	[268,]	DNASE1L2_cg12619509	0.0261	-0.0209
[95,]	SCUBE3_cg21604042	0.123	-0.0984	[269,]	CLEC7A_cg26066361	-0.0255	0.0204
[96,]	CER1_cg01446692	-0.1227	0.0981	[270,]	CD80_cg21572897	-0.0252	0.0202
[97,]	LPAL2_cg15398520	-0.1226	0.0981	[271,]	POU4F2_cg24199834	0.025	-0.02
[98,]	BNC2_cg14613546	-0.1213	0.097	[272,]	GPR141_cg24995381	-0.0242	0.0194
[99,]	CDKL2_cg24432073	0.1191	-0.0953	[273,]	HIST1H2AB_cg19430897	0.0238	-0.019
[100,]	SLC6A1_cg11021744	0.1183	-0.0947	[274,]	SYP_cg16370737	-0.0233	0.0187
[101,]	FLJ45964_cg10832945	-0.1181	0.0945	[275,]	REG3A_cg27342801	-0.0222	0.0178
[102,]	CDX2_cg01424107	0.1176	-0.0941	[276,]	BNC1_cg19988449	0.0216	-0.0173
[103,]	SAT_cg20308511	0.1171	-0.0936	[277,]	SBSN_cg23680518	-0.021	0.0168
[104,]	IFNA1_cg11959435	-0.1166	0.0933	[278,]	KLHL1_cg20523861	0.021	-0.0168
[105,]	RUNX3_cg00117172	0.1122	-0.0897	[279,]	CD300LB_cg06407137	-0.0208	0.0166
[106,]	CLEC5A_cg18463686	-0.1117	0.0893	[280,]	THEDC1_cg04001802	-0.0207	0.0165
[107,]	CCL7_cg02936263	-0.1109	0.0887	[281,]	ATP8A2_cg12111714	0.0199	-0.0159
[108,]	DEFB129_cg02046532	-0.1093	0.0875	[282,]	OR5P3_cg01469547	-0.0197	0.0157
[109,]	CLDN16_cg11984608	-0.1086	0.0869	[283,]	NALP14_cg02347487	-0.0195	0.0156
[110,]	UGT1A3_cg23464269	-0.1083	0.0866	[284,]	RTP1_cg11504646	-0.0193	0.0154
[111,]	CXorf20_cg19845843	-0.1067	0.0853	[285,]	HBQ1_cg07703401	0.0189	-0.0151
[112,]	KRTAP26-1_cg18822544	-0.1066	0.0853	[286,]	PPEF1_cg17198372	-0.0188	0.015
[113,]	TAT_cg22136365	0.1059	-0.0847	[287,]	FGF16_cg02096520	-0.0184	0.0147
[114,]	FAM107B_cg02876062	-0.1042	0.0833	[288,]	SLC36A3_cg22445920	-0.0181	0.0145
[115,]	PAGE1_cg23937047	-0.1033	0.0826	[289,]	FSHB_cg27420123	-0.0178	0.0143
[116,]	HRG_cg07749074	-0.1028	0.0823	[290,]	TRPA1_cg06493386	0.0178	-0.0142
[117,]	GRIA4_cg19343464	0.1022	-0.0818	[291,]	FCRL2_cg11921829	-0.017	0.0136
[118,]	OR10H3_cg25843439	-0.0995	0.0796	[292,]	MASP1_cg20725021	-0.0165	0.0132
[119,]	KIAA0367_cg19282250	-0.0991	0.0792	[293,]	CMTM5_cg00174500	0.0161	-0.0129
[120,]	TRY1_cg10466917	-0.0969	0.0775	[294,]	COX7A1_cg24335895	0.0159	-0.0127
[121,]	MAS1L_cg01078434	-0.0969	0.0775	[295,]	SERPINA10_cg05788638	-0.0158	0.0126
[122,]	KCNK10_cg10935723	-0.0962	0.077	[296,]	HECW1_cg17628717	-0.0156	0.0125
[123,]	ATP4A_cg04713352	-0.0956	0.0765	[297,]	MGC16291_cg21096915	0.0144	-0.0115
[124,]	DNAI2_cg11856697	-0.0938	0.075	[298,]	DEFB118_cg03014957	-0.0143	0.0114
[125,]	SPATA22_cg06862644	-0.0937	0.0749	[299,]	GRASP_cg04034767	0.0141	-0.0113
[126,]	CD33_cg11122968	-0.0931	0.0745	[300,]	CD38_cg26043257	0.0139	-0.0112
[127,]	TAAR1_cg15582891	-0.0928	0.0742	[301,]	GALP_cg23075286	-0.0132	0.0106
[128,]	RGS9_cg15595739	-0.0927	0.0741	[302,]	OLFML1_cg22243733	0.0132	-0.0105
[129,]	BAI1_cg21210789	-0.0914	0.0731	[303,]	HIST1H3E_cg07922606	0.0123	-0.0098
[130,]	SERPINA10_cg19937039	-0.0894	0.0715	[304,]	LGI1_cg00532335	-0.0123	0.0098
[131,]	C10orf81_cg11204562	-0.0892	0.0714	[305,]	CAMK2A_cg27244482	-0.0115	0.0092
[132,]	HRH2_cg20277670	-0.0892	0.0714	[306,]	ABCC9_cg20025970	-0.0114	0.0091
[133,]	CCL1_cg20556988	-0.0885	0.0708	[307,]	NALP10_cg20311730	-0.0112	0.0089
[134,]	CD274_cg19724470	0.0884	-0.0707	[308,]	GPC5_cg09896445	0.0109	-0.0087
[135,]	SLC39A12_cg26550234	-0.0875	0.07	[309,]	DOCK2_cg02251134	-0.0108	0.0086
[136,]	ADRA1B_cg09038885	0.0867	-0.0694	[310,]	PDE10A_cg24133115	0.0104	-0.0084
[137,]	PRND_cg09906458	-0.0863	0.069	[311,]	NTF3_cg04740359	-0.0104	0.0083
[138,]	NPTX2_cg00548268	0.0859	-0.0687	[312,]	CLEC10A_cg21550483	-0.0101	0.0081

[139,]	FPR1_cg05376954	-0.084	0.0672	[313,]	MOSPD2_cg11648471	-0.0101	0.008
[140,]	CRISPLD1_cg01410472	0.0835	-0.0668	[314,]	PAQR4_cg13644052	-0.0099	0.0079
[141,]	MYH4_cg23400451	-0.083	0.0664	[315,]	SCGB1D1_cg13916742	-0.0097	0.0077
[142,]	SGNE1_cg15787039	0.0821	-0.0657	[316,]	CXorf23_cg25257360	-0.0094	0.0075
[143,]	TOLLIP_cg22424444	-0.0805	0.0644	[317,]	SERPINB12_cg11435943	-0.0092	0.0073
[144,]	CCNA1_cg12571423	0.0798	-0.0638	[318,]	KRT5_cg23645091	-0.0089	0.0071
[145,]	DPP6_cg26738880	-0.0792	0.0633	[319,]	ITM2A_cg06208111	-0.0085	0.0068
[146,]	ASB5_cg11698653	-0.0792	0.0633	[320,]	KRTAP13-1_cg02764897	-0.0084	0.0067
[147,]	KRTAP20-1_cg25388528	-0.0767	0.0614	[321,]	SIGLEC7_cg01193293	-0.0082	0.0066
[148,]	BAK1_cg07679836	0.0764	-0.0611	[322,]	GDF5_cg07378350	-0.0079	0.0063
[149,]	SLN_cg12237269	-0.0764	0.0611	[323,]	PKIA_cg04689061	0.0068	-0.0054
[150,]	SLC22A12_cg07220939	-0.0763	0.061	[324,]	OTUD7_cg22458082	-0.0067	0.0054
[151,]	PHLDA1_cg27182761	0.0761	-0.0608	[325,]	LCE5A_cg01868128	-0.0064	0.0051
[152,]	GPR1_cg19132372	-0.0758	0.0606	[326,]	PSG1_cg25839766	-0.0063	0.005
[153,]	HBB_cg14544583	-0.0756	0.0605	[327,]	C11orf38_cg23743472	-0.006	0.0048
[154,]	CLCN5_cg20062122	-0.0741	0.0593	[328,]	C1orf158_cg24338843	-0.0058	0.0046
[155,]	TRPM6_cg22161874	-0.0739	0.0591	[329,]	REG1B_cg07841014	-0.0056	0.0045
[156,]	FLJ46365_cg03167883	-0.0727	0.0581	[330,]	ABP1_cg07897701	-0.005	0.004
[157,]	LPL_cg08918749	0.0722	-0.0578	[331,]	OR1F1_cg07879977	-0.0048	0.0038
[158,]	KRTAP13-3_cg16431978	-0.0709	0.0567	[332,]	MAGEA5_cg06313930	-0.0045	0.0036
[159,]	CD300C_cg17677397	-0.0709	0.0567	[333,]	TREM2_cg20095587	-0.0042	0.0034
[160,]	NRG1_cg19162158	0.0707	-0.0565	[334,]	C11orf38_cg07747336	-0.0029	0.0023
[161,]	ADRA1D_cg11934695	0.0695	-0.0556	[335,]	PARP12_cg07937272	0.0027	-0.0022
[162,]	OR6A2_cg03780486	-0.0676	0.0541	[336,]	GHRHR_cg05058973	-0.0026	0.0021
[163,]	MS4A3_cg25944100	-0.0672	0.0537	[337,]	CDO1_cg07644368	0.0022	-0.0018
[164,]	SCN3B_cg13765785	0.0668	-0.0534	[338,]	LYNX1_cg23180489	-0.0019	0.0015
[165,]	ANXA4_cg22792910	-0.0665	0.0532	[339,]	DEFB103A_cg25214366	-0.0019	0.0015
[166,]	ACE2_cg08559914	-0.066	0.0528	[340,]	KRT1_cg10766289	-0.0016	0.0013
[167,]	CNTNAP4_cg06793062	-0.0658	0.0526	[341,]	HIST1H1A_cg14652095	4,00E-04	-3,00E-04
[168,]	UGT1A10_cg18098286	-0.0656	0.0525	[342,]	KRT23_cg06378617	-2,00E-04	2,00E-04
[169,]	GGTLA1_cg15448245	0.0649	-0.0519				
[170,]	PCBP3_cg07685034	-0.0638	0.051				
[171,]	GPR83_cg27634151	0.0637	-0.0509				
[172,]	MID1_cg25710140	-0.0631	0.0505				
[173,]	HBG1_cg01598642	-0.063	0.0504				
[174,]	SPARC_cg25913233	0.0627	-0.0502				

Table 23 An overview of all mutations and polymorphisms in p53 found in the tumor samples

Sample	P53 mutation	Polymorphism exon	C.72 gDNA (blood)	Polymorphism intron
101	C.248 CGG>TGG,het		C.72:G/G (G/G)	
102	WT		C.72:G/C (G/C)	
103	WT	c.36 CCG>CCA, het	C.72:C/C (C/C)	IVS 9+12 T>C het
104	WT		C.72:C/C (C/C)	
105	WT		C.72:G/G (G/G)	
106	ND		C.72:G/G (G/G)	
107	C.175 CGC>CAC,het		C.72:C/C (C/C)	
108	WT		C.72:G/G (G/G)	
109	WT		C.72:G/G ** (G/C)	
110	C.195 ATC>ACC, het		C.72:C/C (C/C)	
112	WT		C.72:C/C (C/C)	
113	WT		C.72:G/G (G/G)	
114	WT		C.72:G/G (G/G)	
116	C.254 ATC>AAC, het		C.72:C/C * (G/C)	
117	C.110, 1 basedel.		C.72:G/C (G/C)	
118	WT		C.72:G/G (G/G)	
120	WT		C.72:G/G (G/G)	
121	C.222 CCG>CCT het		C.72:G/C (G/C)	IVS3-29C>A het
123	WT		C.72:G/C (G/C)	IVS3-29C>A het
124	WT		C.72:G/G (G/G)	
125	WT		C.72:G/G (G/G)	
127	WT		C.72:G/G (G/G)	
rp36	WT		C.72:G/C	
rp37	ND		C.72:G/C	
rp38	WT		C.72:G/G	
rp40	WT	c.36 CCG>CCA, het	C.72:C/C	IVS 9+12 T>C het
rp42	WT		C.72:G/G	
rp43	WT		C.72:G/C	
rp44	WT		C.72:G/G	
rp45	WT		C.72:G/G	
rp59	WT		C.72:G/C	
rp61	WT		C.72:G/G	
rp68	WT		C.72:G/C	
rp70	WT	c.36 CCG>CCA, het	C.72:G/C	IVS 9+12 T>C het
rp71	WT	c.36 CCG>CCA, hmo	C.72:C/C	IVS 9+12 T>C hmo
rp73	WT		C.72:G/G	
rp75	WT		C.72:G/G	

APPENDIX B Recipes and reagent list

1xTAE-buffer

Reagents:
Trizma®base, Sigma-Aldrich Norway (cat. AS/93362)
Glacial acetic acid, VWR International (cat. 1.00063.1000)
EDTA dinatriumsalt, Titriplex®III, VWR International (cat. 1.08418.0250)

Procedure:
Scale in 242g Trizma®base
Add 500ml MQ-water
Add 100ml 0.5M EDTA (pH=8) and 57,1 ml Glacial acetic acid (fume hood)
Add MQ-water to adjust to 1 liter

Gel loading buffer 0,1% Bromophenol blue

Reagents:
Bromophenol Blue, Bio-Rad Laboratories (cat. 161-0404)
Ficoll™PM 400, Sigma-Aldrich Norway (cat. AS/46327)
1xTAE (Trizma®base/Glacial acetic acid/EDTA dinatriumsalt, Titriplex®III-buffer)

Procedure:

Measure off 50ml 1xTAE-buffer

Scale in 10g Ficoll™PM 400 and 0.05g Bromophenol Blue into a 50ml Sarstedt tube
Fill up with 1x TAE. The Ficoll™ need at least a night to be dissolved. Fill up with the remaining 1xTAE-buffer

Long term storage: Fridge/ Short term storage: Room Temperature

0.1M NaOH

Reagent:
NaOH-pellets, VWR International (cat. 1.06498.1000)

Procedure:
Scale in 0,4g NaOH–pellets and fill up with 0.1L MQ-water.

95% formamide/1mM EDTA

Reagents:
Formamide >99%, Merck (cat. 109684)

EDTA dinatriumsalt, Titriplex®III, VWR International (cat. 1.08418.0250)

NaOH-pellets, VWR International (cat. 1.06498.1000)

Procedure:
0.50M EDTA:

Scale in 93,05g EDTA dinatriumsalt, Titriplex®III
Add 400ml MQ-water
Add 11,1g NaOH-pellets

EDTA will not dissolve until pH 8.0. Adjust to correct pH with NaOH
Add MQ-water up to 0.5L when dissolved
Store in cold room.

Procedure:
95% formamide/1mM EDTA:

Add 24ml MQ-water to 475ml 100% formamide and 1ml 0,5M EDTA

Table 24 All reagents with supplier and catalogue number, used in the analysis

Reagents	Supplier	Cat.nr/Prod.nr.
Maxwell 16 Tissue DNA Purification Kit, 48 samples, Promega	Nerliens Meszansky	AS1030
Quant-iT™ dsDNA Assay Kit, Broad Range, 1000 assays *2-1000 ng*, Promokode: P453571	Invitrogen	Q-33130
Reverse primer MDM2 (6pM/µl)	Eurofins MWG Operon	10685851
Forward primer MDM2 (6pM/µl)	Eurofins MWG Operon	10685850
Primer PIK3CA		
TP53 Primer	MWG	
dNTP mixture	Takara BIO inc.	4030
Hot Star Taq Polymerase	Qiagen	203205
Agarose pulver	BioRad Laboratories	161-3102
Ethidium Bromid	VWR	1.11608.0030
GelRed™ Nucleic Acid Stain	Biotum	41003
GeneRuler™ DNA ladder mix	Fermentas	SM0333
MultiScreen PCRµ96 filter plate	Millipore	LSKMPCR50
epTIPS Motion spisser uten filter 20-300ul	Eppendorf Nordic	0030 003.969
BigDye® Terminator v1.1 cycle sequencing kit	Applied Biosystem	4337451
Multiscreen®HV filter plates	Millipore	MAHVN4550
Sephadex® G-50, 100g	GE Healthcare	17-0041-01
MicroAmp Optical 96-well plate with Barcode	Applied Biosystem	4306737
1stk HumanMethyation27 DNA Analysis BeadChip Kit, Rev B (48 samples) Quotation number: 2010008LB100	Illumina, Inc	WG-311-2202
ALU-C4M Sequence Detection Primer: 80,000 pmol	Applied Biosystems	4304971
ALU-C4M TaqMan® MGB Probe 20,000 pmol	Applied Biosystems	4316033
EpiTect®Bisulphite Kit 48	Qiagen	59104
EpiTect® Control DNA, methylated (100)	Qiagen	59655
TaqManC Genotyping Master Mix, 1-Pack (1 x 10 mL)	Applied Biosystems	4371355
TaqMan SNP assay, human, 188ul,C_9493064_10	Applied Biosystems	4351379
MicroAmp optical adhesive film	Applied Biosystems	4311971