

**Entwicklung einer Methode zur genomweiten Analyse
von DNA-Methylierung
und
Identifizierung von
DNA-Methylierungsbiomarkern
in Brustkrebs**

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes

von
Dipl.-Ing. Anne Faßbender

Saarbrücken
2008

Tag des Kolloquiums:

Dekan:

Berichterstatter:

Danksagung

In erster Linie möchte ich Prof. Dr. Jörn Walter für die Betreuung dieser Arbeit und für die Möglichkeit, als externe Doktorandin an der Universität des Saarlandes eine Doktorarbeit durchzuführen.

Bei der Epigenomics AG danke ich Dr. Ralf Lesche und Dr. Matthias Schuster, die verschiedene Phasen dieser Arbeit betreut haben. Beide haben wesentlich zum Gelingen beigetragen und mich stets durch hilfreiche Tipps und Ideen unterstützt.

Weiterhin danke ich Thomas König, Dr. Jörn Lewin und Tamas Rujan, die die biostatistischen Analysen und Entwicklungen mit mir durchkämpft haben.

Allen ehemaligen Doktoranden, Claudia, Dimo, Manuel, Philipp, Rene und Reimo danke ich für die Unterstützung in allen Lebenslagen: sowohl beim Durchleuchten wissenschaftlicher Untiefen als auch dem Feiern von Erfolgen und noch mehr dem Überwinden vorübergehender Aussichtslosigkeit.

Besonderer Dank gilt Michal Bagrowski und Karen Freyberg sowie den vielen anderen Labormitarbeiter, die einen hervorragenden Job machen und mir viel zur Seite gestanden haben.

Für das Lesen des Manuskriptes danke ich neben Dr. Matthias Schuster und Dr. Ralf Lesche auch Dr. Jürgen Distler, Dr. Dimo Dietrich und meinem Bruder Tobias Faßbender.

Großer Dank gilt meiner Familie. Meinem Vater Till, der zwar nicht mehr hier ist, der aber trotzdem immer da ist und der einen großen Anteil an meinem technischen Interesse hat. Meiner Mutter Hanne, ohne die ich besonders den langen Endspurt nicht geschafft hätte, was auch für meinen Bruder Tobias und seine Familie gilt. Meinem Sohn Julius, weil er dem Leben soviel Sinn gibt.

Sehr wichtig für das Gelingen dieser Arbeit waren auch gute Freunde wie Claudia, Christoph, Elke, Hanene, Melissa, Susanne, Stephanie und Yvonne. Danke für den Beistand in allen Lebenslagen.

Die Arbeit wurde gefördert durch das Bundesministerium für Bildung und Forschung
Foerderprojekt: Technologien zur Charakterisierung des humanen regulatorischen Proteom
und das EU-Projekt: EU-funded sixth framework program project (LSHC-CT-2003-504586):
DNA Methylation Profiling in Breast Cancer.

INHALTVERZEICHNIS

1. EINLEITUNG	1
1.1. Brustkrebs: Arten, Diagnose, molekulare Marker	1
1.1.1. Epidemiologie von Brustkrebs	1
1.1.2. Arten und Klassifizierung von Brustkrebs	2
1.1.3. Diagnose und Behandlung von Brustkrebs	4
1.1.3.1. Anti-Hormontherapie	5
1.1.3.2. Chemotherapie	6
1.1.4. Anwendung molekularer Biomarker	7
1.1.5. Molekulare Biomarker in Brustkrebs	8
1.1.6. Methylierung und Brustkrebs	9
1.1.6.1. Die Funktion der Methylierung in gesunden Zellen	10
1.1.6.2. DNA-Methylierung und Brustkrebs	11
1.1.6.3. DNA-Methylierung als molekularer Biomarker.	13
1.2. Methoden für die Analyse der DNA – Methylierung	15
1.2.1. Methoden zur Bestimmung von DNA-Methylierungsmustern und DNA-Methylierungsniveaus	17
1.2.2. Genomweite Analysemethoden von DNA-Methylierungsprofilen	18
1.2.2.1. Methoden ohne Mikroarrays	19
1.2.2.2. Methoden mit Mikroarrays	21
1.2.3. Zusammenfassung – Brustkrebs und DNA-Methylierung	24
1.3. Zielsetzung	25
2. METHODEN	26
2.1. Allgemeine molekularbiologische Methoden	26
2.1.1. DNA Präparation	26
2.1.2. UV Quantifizierung von DNA	26
2.1.3. Herstellung eines DNA Standards mit 0% DNA-Methylierung	26
2.1.4. Herstellung eines DNA Standards mit 100% DNA-Methylierung	27
2.2. Herstellung einer Repräsentation methylierter DNA Fragmente (DMH-Amplicon)	27
2.2.1. Fragmentierung genomischer DNA	27
2.2.2. Ligation von Adaptern	28
2.2.3. Herstellung des Adapters	28
2.2.4. DNA-Methylierung - sensitive Restriktion	28
2.2.5. PCR Amplifikation	29
2.2.6. Fragmentierung und Markierung des PCR-Produktes	29
2.2.7. Hybridisierung, Waschen, Färben und Scannen	29
2.2.8. In Silico Analyse des DMH-PCR-Produktes	30
2.2.9. Mikroarraydesign	31
2.2.9.1. Design des Entwicklungs-Mikroarrays	31
2.2.9.2. Design des 50.000 Fragmente Mikroarrays	32
2.2.10. Datenverarbeitung	33
2.3. Direkte Bisulfitsequenzierung	34
2.3.1. Herstellung von Bisulfit-DNA	34
2.3.2. PCR Amplifikation von Bisulfit-DNA	34
2.3.3. Sequenzierung von Bisulfit-DNA	34
2.3.4. Datenanalyse der Bisulfitsequenzierung	35

3. ERGEBNISSE	36
3.1. Optimierung und Etablierung einer Methode für die genomweite Analyse der DNA – Methylierung	36
3.1.1. Optimierung und Etablierung des DMH-Protokolls	37
3.1.1.1. Nicht-methylierungssensitive Fragmentierung	37
3.1.1.2. Ligation	38
3.1.1.3. Methylierungssensitive Restriktion	38
3.1.1.4. DMH-Amplicon PCR	39
3.1.1.5. Fragmentierung, Markierung und Hybridisierung	41
3.1.2. Analyse der Komplexität, Größenverteilung und Anzahl informativer Fragmente	42
3.1.3. Design des Entwicklungs-Mikroarrays	44
3.1.4. Entwicklung der Datenverarbeitung	46
3.1.5. Charakterisierung der optimierten DMH-Methode	48
3.1.5.1. Herstellung von Amplicons aus PBL und Brustkrebszelllinie	48
3.1.5.2. Reproduzierbarkeit	49
3.1.5.3. Überprüfung mittels Bisulfitsequenzierung	51
3.1.6. Der 50.000-Fragmente-Mikroarray	52
3.2. Analyse der DNA – Methylierung in Brustkrebs	54
3.2. Analyse der DNA – Methylierung in Brustkrebs	54
3.2.1. Verwendete Proben und Gruppeneinteilung	54
3.2.2. Studienaufbau	55
3.2.3. Datenqualität und Vergleichbarkeit der zwei Studien	56
3.2.4. Einteilung in drei DNA-Methylierungsgruppen mit Hilfe der 0% und 100% Kontrollen	58
3.2.5. Verteilung der Methylierung in gesundem Gewebe	59
3.2.6. Kriterien für die Auswahl von Markerkandidaten	61
3.2.7. Unterschiede zwischen den gesunden Gewebekontrollen und den Tumorproben	61
3.2.8. Unterschiede zwischen Chemotherapie-Ansprechenden und Nicht-Ansprechenden	64
3.2.9. Bisulfitsequenzierung ausgewählter Kandidaten	67
4. DISKUSSION	74
4.1. Genomweite Analyse von DNA-Methylierungsunterschieden	74
4.1.1. Vergleich der DMH-Methode mit anderen Technologien	74
4.1.2. Oligonukleotid-Mikroarrays	75
4.1.3. Verbesserung des Protokolls und der Datenauswertung	76
4.1.4. Verbesserung der Datenauswertung	76
4.1.5. Einfluss genomischer Veränderungen	77
4.1.6. Datenqualität und quantitative Auswertung	78
4.1.7. Auswirkungen der veränderten Enzymkombinationen	79
4.2. Methylierungsmarker in Brustkrebs	80
4.2.1. Die verwendeten Patientinnen	81
4.2.2. Die Methylierung in gesundem Brustgewebe	83
4.2.3. Methylierungsunterschiede in Brustkrebs und deren Verwendung als Früherkennungsmarker	85
4.2.4. Markerkandidaten zur Vorhersage des Ansprechens auf eine Chemotherapie	89
4.3. Zusammenfassung	92
5. ANHANG	94
5.1. Firmen	94
5.2. Geräte	95
5.3. Chemikalien	95
5.4. Puffer und Lösungen	96

5.5.	Enzyme, Kits und molekularbiologische Reagenzien	96
5.6.	Bisulfitprimer - Brustkrebsstudie	97
5.7.	Fragmentsequenzen – Brustkrebsstudie	97
5.8.	Proben	98
5.9.	Listen für Kandidaten der Tumormarker	100
5.10.	Listen für Kandidaten der Tumormarker	112
5.11.	Abkürzungsverzeichnis und Gennamen	121
5.12.	Literaturverzeichnis	126
5.13.	Lebenslauf	135
5.14.	Veröffentlichungen	136

Zusammenfassung

Ziel der Arbeit war die Identifizierung von DNA-Methylierungsbiomarkern für das Ansprechen von Brusttumoren auf eine Chemotherapie. Hierzu wurde die Differentielle Methylierungs Hybridisierung (DMH) optimiert. Durch veränderte Enzymkombinationen wurde der Anteil analysierbarer CpGs erhöht und die Komplexität reduziert. Veränderungen des Protokolls führten zur Senkung der benötigten DNA Ausgangsmenge. Ein wesentlicher Schritt war die Einführung eines für die entstehenden Fragmente spezifischen Mikroarrays und von Kontrollen für die Normalisierung von Komplexitätsunterschieden. Die Reproduzierbarkeit des gesamten Prozesses zeigte einen Korrelationskoeffizienten von 94%. 111 Fragmente wurden direkt bisulfitsequenziert. Der Vergleich beider Methoden zeigte eine hohe Korrelation von 71%. Daher erlaubt die verbesserte Methode eine quantitative genomweite Methylierungsanalyse.

Für die Suche nach DNA-Methylierungsbiomarkern wurden jeweils 14 Brusttumore verwendet, deren Metastasen auf eine Chemotherapie ansprachen bzw. nicht-ansprachen. Der Vergleich mit gesundem Gewebe ergab 145 Kandidaten für Aufmethylierung und 202 für Demethylierung in den Tumoren. Beim Vergleich von Ansprechenden und Nicht-Ansprechenden wurden 140 Kandidaten identifiziert. Für die Weiterentwicklung dieser Kandidaten wäre zunächst die Bestätigung der Unterschiede mit einer unabhängigen Methode und in unabhängigen Proben notwendig. Ein erster Schritt war die direkte Bisulfitsequenzierung von 12 Kandidaten. Ein Fragment wies ein p-value von 0,03 beim Vergleich der insgesamt sequenzierten 29 Ansprechenden mit 21 Nicht-Ansprechenden Proben auf, d.h. es zeigte einen statistisch signifikanten Unterschied.

Summary

Focus of the study was to identify candidates for markers that predict the response of breast cancer patients to chemotherapy. Therefore the method differential methylation hybridisation (DMH) was chosen to be optimized. The introduction of a different combination of restriction endonucleases resulted in an increased amount of analyzed CpGs and the reduction of the complexity. A reduction of the DNA input was achieved by modifications of the protocol. An important step was the establishment of an oligonucleotide microarray specifically designed for DMH fragments and of control fragments for the normalization of differences of the complexity. The reproducibility of the complete procedure was high (94 %). The analysis of 111 fragments via direct bisulphite sequencing resulted in a good correlation of 71% with the DMH data. This means the optimized method makes quantitative analysis possible.

In order to identify methylation biomarkers in breast cancer 14 chemotherapy responders and 14 non-responders were analysed. The comparison of the tumor and the healthy samples identified 145 tumor marker candidates as hyper-methylated and 202 as down-methylated. Looking at the differences between responders and non-responders 140 candidates were found. The next steps would be the analysis of the methylation with an independent method and the verification of the differences with independent samples. A first step towards this was the analysis of 12 tumor marker candidates by bisulphite sequencing. One fragment showed a significant difference with a p-value of 0.03 when looking at all sequenced samples, which mean 29 responders and 21 non-responders. Therefore, this fragment is a very promising marker candidate.

1. Einleitung

1.1. Brustkrebs: Arten, Diagnose, molekulare Marker

1.1.1. Epidemiologie von Brustkrebs

Brustkrebs stellt die häufigste Krebserkrankung von Frauen dar. Weltweit erkrankten nach Schätzungen der WHO im Jahr 2000 über eine Million Frauen und 370.000 verstarben daran. In Deutschland betrug im Jahr 2003 die Zahl der Neuerkrankungen etwa 47.500 und die der Todesfälle circa 17.200. Brustkrebs stellt daher mit etwa 24 % den größten Anteil aller Krebsneuerkrankungsfälle und mit circa 18% auch den größten Anteil an der Krebssterblichkeit von Frauen [1]. Sowohl die Neuerkrankungsrate als auch die Sterberate von Brustkrebs sind abhängig vom Alter. Das Erkrankungsrisiko nimmt mit fortschreitendem Alter besonders aber zwischen 30 und 50 Jahren zu. Bei etwa 40% aller Frauen wird die Diagnose allerdings schon vor Erreichen des 60-sten Lebensjahr gestellt, was im Vergleich zu vielen anderen Krebserkrankungen früh ist [1].

Die Heilungschance von Krebserkrankungen wird mit der 5- bzw. 10-Jahres-Überlebensrate ausgedrückt. Die 5-Jahres-Überlebensrate für Brustkrebs betrug im Zeitraum zwischen 1990 und 1994 für in Deutschland erhobene Daten circa 79%. Dieser Wert ist im Vergleich zu anderen Krebsarten hoch. Zum Beispiel liegen die 5-Jahres-Überlebensrate von Magenkrebs bei nur 31 % und die von Lungenkrebs bei 15 %. Auch die 10-Jahres-Überlebensrate von Brustkrebs ist mit circa 65% im Vergleich zu anderen Krebsarten hoch [1]. Die 5- und 10-Jahres-Überlebensraten sind in den letzten Jahren erheblich angestiegen. Allerdings kann nicht eindeutig geklärt werden, ob dieser Anstieg auf eine verringerte Sterblichkeit aufgrund der verbesserten Früherkennung oder Therapie zurückzuführen ist, da die vermehrte Früherkennung der Tumore automatisch zu einer verlängerten Überlebenszeit führt [2]. Ausserdem bleibt Brustkrebs trotz der erhöhten Überlebensraten aufgrund seiner hohen Erkrankungsrate eine der Haupttodesursachen von Frauen.

Wesentliche Aufgaben für die Brustkrebsforschung sind daher die Verbesserung der Früherkennung und die Verfeinerung der Diagnose mit dem Ziel, maßgeschneiderte Therapien zu entwickeln. Brustkrebs ist nach einem Rückfall, d.h. dem Auftreten einer Metastase nicht mehr heilbar. Die Vorhersage des Krankheitsverlaufs (prognostisch) oder des Ansprechens eines Tumors auf eine Therapie (prädiktiv) sind daher weitere wesentliche Fragestellungen.

1.1.2. Arten und Klassifizierung von Brustkrebs

Die weibliche Brust (siehe Abbildung 1) besteht zum größten Teil aus Fett- und Drüsengewebe sowie aus Bindegewebe, Lymph- und Blutgefäßen. Die Brustdrüse ist aus einem duktalem und einem lobulärem Anteil aufgebaut. Der duktale Anteil besteht aus Milchgängen (Ductus lactifer), die kleinere und größere milchabführende Gänge (Duktuli) bilden und in die Brustwarze (Mamilla) münden. Beim lobulären Anteil handelt es sich um Drüsenläppchen (Lobuli), welche durch kleine Milchgänge zu Drüsenlappen verbunden werden. Die Drüsenlappen stellen die milchproduzierenden Strukturen des Milchgangssystems dar. Sowohl der duktale als auch der lobuläre Anteil bestehen hauptsächlich aus Epithelzellen. Die Drüsenläppchen bestehen aus einem inneren einschichtigen Epithel, das durch eine Basalmembran vom umgebenden Bindegewebe getrennt ist. Zwischen den Epithelzellen und ihrer Basalmembran liegen Korbzellen, die die Fähigkeit zur Kontraktion besitzen (Myoepithelien). Kleine Milchgänge haben ebenfalls ein einschichtiges Epithel. Die größeren Milchgänge besitzen ein ein- bis zweischichtiges Epithel und in ihrer Wand glatte Muskulatur, die dem Milchtransport dient.

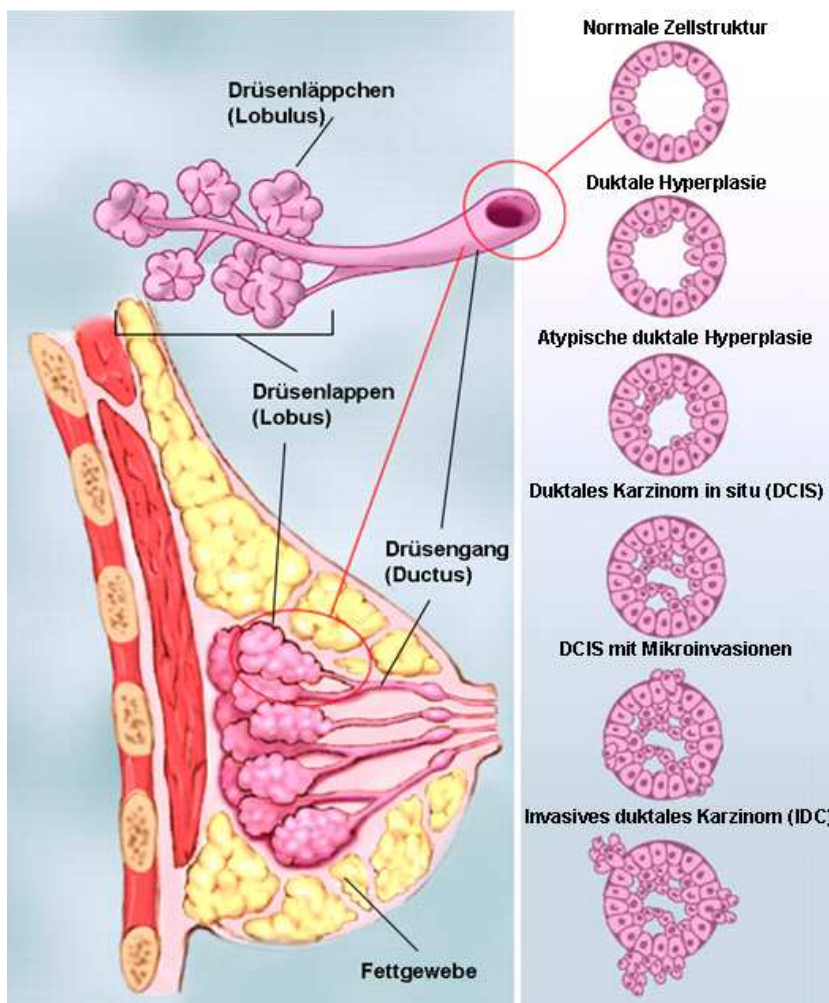


Abbildung 1: Schematische Darstellung verschiedener Stadien der Progression von soliden Tumoren (aus <http://www.breastcancer.org>)
 A: normales Epithel, B: benigne Läsion (Hyperplasie), C: in situ Karzinom, D: invasives Karzinom, E: metastasierendes Karzinom.

In der Medizin bezeichnet der Begriff Krebs einen bösartigen (malignen) Tumor, der durch Körperzellen gekennzeichnet ist, die wachsen, sich teilen und gesundes Gewebe verdrängen, infiltrieren und zerstören. Die meisten Tumore sind allerdings gutartig (benigne), d.h. sie bleiben von einer Basalmembran umgeben und infiltrieren umliegendes Gewebe nicht. Brustkrebs ist definiert als ein von Epithelzellen ausgehender maligner Tumor (Karzinom). Wachsen die entarteten Zellen noch innerhalb der Struktur ihres Ursprungsgewebes spricht man von einem Karzinom in situ (siehe Abbildung 1). Die malignen in situ Karzinome werden als unmittelbare Vorstufe eines invasiven Karzinoms gesehen. Beim invasiven Karzinom haben die Tumorzellen die Basalmembran durchbrochen und stehen in direktem Kontakt mit der Umgebung. Das häufigste Brustkarzinom ist das invasive-duktales Karzinom (IDC; 70-80% aller Brustkarzinome), das sich in den kleinen Milchgängen entwickelt. Die nicht-invasive Form ist das duktales Karzinom in situ (DCIS). Den zweitgrößten Anteil macht das invasiv-lobuläre Karzinom (ILC; 10-20% aller Brustkarzinome) aus. Die nicht-invasive Form ist das lobuläre in situ Karzinom (LCIS). Neben diesen Hauptgruppen gibt es zahlreiche seltenere Varianten, z.B. muzinöse oder medulläre Karzinome, die sich aufgrund ihrer Histologie von den duktales und lobulären Karzinomen unterscheiden [3].

Bei der feingeweblichen (histologischen) Klassifizierung von Brusttumoren werden verschiedene Parameter der Erkrankung herangezogen. Zunächst ist das Ziel, gutartige Veränderungen, wie vom Bindegewebe ausgehende Tumore (Fibroadenome), mit Flüssigkeit gefüllte Hohlräume (Zysten) oder Fettgeschwulste (Lipome) von malignen Tumoren abzugrenzen. Bei malignen Tumoren wird dann zunächst bestimmt, ob es sich um nicht-invasive DCIS oder LCIS, um bereits invasive ILC oder IDC bzw. um eine der seltenen Formen handelt.

Tabelle 1: TNM-Klassifikation von malignen Brusttumoren nach WHO

TNM-Klassifikation	
T (Tumorgröße)	Stadium I
T0: kein Tumor nachweisbar	I T1, N0, M0
Tis: Carcinoma in situ, nicht-invasiv	Stadium II
T1: Der Tumor ist nicht größer als 2 cm	IIA: T0-1, N1, M0 oder T2, N0, M0
T2: Tumor mit einem Durchmesser von > 2 bis 5 cm	IIb: T2, N1, M0 oder T3, N0, M0
T3: Der Tumor ist größer als 5 cm	Stadium III
T4: Tumor jeder Größe mit Ausdehnung auf die Brustwand oder Haut	IIIA: T3, N1, M0 oder T0-3, N2, M0
N (Nodes = Befallene Lymphknoten)	IIIB: T4, jedes N-Stadium, M0
N0: keine befallenen Lymphknoten	IIIC: jedes T-Stadium, N3, M0
N1: Metastasen in 1-3 Lymphknoten der Achsel	Stadium IV
N2: Metastasen in 4-9 Lymphknoten der Achsel	jedes T-Stadium, jedes N-Stadium, M1
N3: Metastasen in 10 oder mehr Lymphknoten der Achsel oder unter/über dem Schlüsselbein	
M (Fern-Metastasen)	
M0: keine Metastasierung in andere Organe	
M1: Fern-Metastasen vorhanden, meist in Knochen, Lunge, Leber oder Gehirn	

Maligne Tumore werden weiterhin anhand der TNM Klassifikation gemäß WHO aufgrund makroskopischer Gesichtspunkte in Gruppen eingeteilt. Hierzu dienen die drei wesentliche Faktoren: Größe des Primärtumors zum Diagnosezeitpunkt (T); Ausbreitung des Tumorgewebes regional in die Lymphknoten (N) der Brust oder bis zur Achselhöhle und Vorliegen von Metastasen(M) in anderen Organen (siehe: Tabelle 1) [4]. Weiterhin wird die Differenzierung des Tumors für das „Grading“ verwendet [5]. Dabei spielen beispielsweise Faktoren wie Kernmorphologie und Mitoserate eine Rolle um den Tumor in die „Grades“ 1-4 einzuteilen, wobei G1 noch gut differenziert, d.h. dem Ursprungsgewebe ähnlich und G4 undifferenziert, d.h. dem Ursprungsgewebe nicht mehr ähnlich, bedeutet. Eine weitere Ebene der Klassifikation bilden molekulare Biomarker, d.h. Eigenschaften eines Tumors auf der molekularen Ebene wie das Vorhandensein bestimmter Moleküle oder DNA Sequenzen. Bereits etabliert sind der Hormonrezeptorstatus der Hormone Progesteron und Östrogen [6], [7]. Über 70 % aller Brustkrebstumore sind ER positiv [8]. Ein positiver Östrogen- und Progesteronrezeptor-Status ist assoziiert mit gut differenzierter Histologie, negativem Lymphknoten-Status, Diploidie, niedriger Proliferationsrate und einer Tendenz für einen positiven klinischen Verlauf [6, 8, 9]. Der ER und PgR Status wird für die Wahl einer adjuvanten Hormontherapie verwendet [10]. Entsprechend der biologischen Funktion des Östrogenrezeptors sprechen ER und PgR positive Patientinnen in der Regel gut auf adjuvante Hormontherapien an [9].

1.1.3. Diagnose und Behandlung von Brustkrebs

Der größte Teil der Brusttumore wird von Frauen selbst ertastet. In diesen Fällen haben sie häufig schon eine Größe von mehr als 2 cm Durchmesser erreicht. Im Rahmen von Früherkennungsprogrammen mittels Mammografie oder Ultraschalluntersuchung werden auch wesentlich kleinere Tumore entdeckt. In der Regel wird dann zunächst mittels einer Biopsie eine Gewebeprobe entnommen und untersucht, ob es sich um einen malignen Tumor handelt. Bei einem malignen Tumor ist grundsätzlich eine operative Entfernung der erste Schritt der Therapie. Hierbei ist das Ziel, das gesamte Tumormaterial zu entfernen und gleichzeitig die Brust zu erhalten. Ist dies aufgrund der Tumorgöße und Tumorausbreitung nicht möglich, kann eine neo-adjuvante (unterstützende, zusätzliche) systemische (den gesamten Körper betreffende) Therapie der Operation vorangestellt werden. Hierbei handelt es sich meist um eine Chemotherapie, die den Tumor auf eine operable Größe schrumpfen lassen soll. Bei der operativen Entfernung des Tumors, werden auch Lymphknoten aus der Achselhöhle entnommen. Sind in den Lymphknoten Krebszellen zu finden, muss davon ausgegangen werden, dass die Tumorzellen zu wandern begonnen haben und eventuell eine Metastasierung bereits vorliegt.

Aus den in Abschnitt 1.1.2 beschriebenen Merkmalen des Tumors und weiteren prognostischen Faktoren wie Alter der Frau, menopausaler Status, Östrogenrezeptor- (ER-)

und Progesteronrezeptor- (PgR) Status wird die längerfristige Prognose, d.h. der wahrscheinliche Verlauf der Erkrankung ohne eine Behandlung, abgeschätzt. Je nach Prognose kommt es nach der Operation zu einer Chemo-, Strahlen- oder Hormontherapie bzw. Kombinationen dieser Behandlungen. Liegen noch keine detektierbaren Metastasen an anderen Organen vor, ist das Ziel der anschließenden systemischen Therapie, eventuell nicht entfernte Tumorzellen sowie nicht detektierte Mikro-Metastasen abzutöten oder ihr Wachstum zu verhindern bzw. zu verlangsamen. Diese Therapie wird als adjuvante (unterstützende oder zusätzliche) Therapie bezeichnet. Über die Art der adjuvanten Therapie entscheidet der Arzt anhand der Prognose der Patientin. Einerseits sollte ein Rückfall vermieden werden, andererseits stellt jede Therapie eine erhöhte Belastung für die Patientin dar und birgt das Risiko von Folgeerkrankungen. Sind bereits Metastasen vorhanden, dient die Therapie der Lebensverlängerung und Verminderung der Symptome und wird als palliativ (krankheitsmildernd) bezeichnet. Während die Prognose der Entscheidung über Notwendigkeit und Form einer Therapie dient, werden für die Auswahl der Therapie prädiktive Marker benötigt, welche die Wirksamkeit vorhersagen. Die Wirksamkeit hängt davon ab, ob der Tumor bzw. seine Metastasen auf die Therapie ansprechen oder gegenüber der Therapie resistent sind.

1.1.3.1. Anti-Hormontherapie

Wie bereits erwähnt sind Östrogenrezeptor- (ER) und Progesteronrezeptor- (PgR) Status wichtige prognostische Faktoren. Gesunde Brustzellen besitzen Rezeptoren für diese weiblichen Hormone, welche ihr Wachstum steuern. Auch Brustkrebszellen besitzen häufig diese Rezeptoren. Dadurch ist ihr Wachstum von diesen Hormone abhängig. Daher lässt sich durch Hormonentzug das Wachstum eventuell nicht entfernter oder gewanderter Zellen, die Metastasen bilden würden, stark verlangsamen bzw. stoppen.

Es gibt verschiedene Wege, die wachstumsfördernde Wirkung der Hormone auszuschalten. Am weitesten verbreitet ist das Anti-Östrogen Tamoxifen. Eine weitere Möglichkeit ist das Ausschalten der Bildung von Östrogen. Bei Frauen vor der Menopause wird dies durch die Gabe von GnRH (Gonadotropin-Releasing-Hormon)-Analoga erreicht, wodurch die Eierstöcke die Östrogenproduktion einstellen. Bei Frauen nach der Menopause, wird Östrogen durch die Umwandlung von Androgen in Östrogen in anderen Geweben gebildet. Für die Umwandlung wird das Enzym Aromatase benötigt. Daher werden bei Frauen nach der Menopause Aromatasehemmer eingesetzt.

Bei der Antihormontherapie handelt es sich um eine relativ milde Therapieform, d.h. die Nebenwirkungen sind im Vergleich zu einer Chemotherapie wesentlich geringer. Trotzdem hat sie eine sehr hohe Wirksamkeit und ist in vielen Fällen als adjuvante Therapie ausreichend. Weiterhin wird sie auch als Teil der palliativen Therapien eingesetzt.

1.1.3.2. Chemotherapie

Bei der Chemotherapie kommen Zellgifte (Zytostatika) zum Einsatz. Sie stören die Zellteilung, d.h. sie greifen sich teilende Zellen, z.B. die schnell wachsender Zellverbände eines Tumors, an. In Tabelle 2 sind die Hauptgruppen der heute verwendeten Zytostatika und ihre Wirkungsweisen aufgeführt. Bei Brustkrebs werden meist Polychemotherapien durchgeführt bei denen Zytostatika aus verschiedenen Wirkungsgruppen in Intervallen verabreicht werden. Die Intervallbehandlung setzt sich aus Behandlungsphase und Behandlungspause zusammen. Ziel der Behandlungspausen ist die Regeneration der Patientinnen. Die Intervallgabe ist nötig, da nur Zellen, die sich in der Teilungsphase befinden durch Zytostatika abgetötet werden können während Zellen, die sich in anderen Stadien des Zellzyklus befinden, nicht angegriffen werden. Grund für die Anwendung verschiedener Zytostatika ist die dadurch erhöhte Wirksamkeit bzw. die verringerte Wahrscheinlichkeit für eine Resistenz des Tumors gegen jeden der verschiedenen Wirkungsmechanismen.

Tabelle 2: Verschiedene Zytostatikagruppen, Wirkungsweise und Beispiele

Zytostatika		
Alkylantien (Alkylanzien)		
<i>Übertragung von Alkylgruppen auf die DNA. Vernetzung von DNA-Strängen. Verhinderung korrekter DNA-Replikation.</i>	Stickstoff-Lost-Derivate:	Cyclophosphamid, Ifosfamid, Trofosfamid, Melphalan, Chlorambucil
	Akylsulfonate:	Busulfan, Treosulfan
	Nitrosoharnstoffe:	Carmustin, Lomustin, Nimustin, Estramustin, Streptozotocin
	weitere:	Procarbazin und Dacarbazin, Temozolomid, Thiotepa
Platinanaloga		
<i>Quervernetzungen der DNA durch kovalente Bindung des Platinatoms an zwei Nukleinbasen. Verhinderung korrekter DNA-Replikation.</i>	weitere:	Cisplatin, Carboplatin, Oxaliplatin
Interkalantien		
<i>Nichtkovalente Bindung an die DNA. Verhindern Anbindung von Enzymen. Verhinderung/Störung von Replikation und Transskription.</i>	Anthracycline	Doxorubicin (Adriamycin), Daunorubicin, Epirubicin, Idarubicin
	weitere:	Mitoxantron, Amsacrin
Antibiotika		
<i>Die Wirkmechanismen sind unterschiedlich, meist DNA-Vernetzung durch Interkalation oder Alkylierung. Verhinderung korrekter DNA-Replikation.</i>	weitere:	Bleomycin, Actinomycin D (Dactinomycin), Mitomycin
Mitosehemmer		
<i>Diese Stoffe binden an Tubulin. Verhinderung korrekter Zellteilung.</i>	Alkaloide der Vinca rosea:	Vinorelbin, Vincristin (Oncovin), Vinblastin, Vindesin
Taxane		
<i>Bildung von anomalen Molekülen im Zellskelett. Verhinderung korrekter Zellteilung.</i>	weitere:	Paclitaxel, Docetaxel
Topoisomerasehemmer		
<i>Bewirken irreguläre, nicht behebbare DNA-Brüche und DNA Vernetzungen.</i>	Topoisomerase-I-Inhibitoren:	Camptothecin, Topotecan, Irinotecan
	Topoisomerase-II-Inhibitoren:	Etoposid, Teniposid
Antimetabolite		
<i>Werden als falsche Bausteine in die DNA oder RNA eingebaut oder verhindern den Einbau der korrekten Bausteine. Störung von Zellteilung und Stoffwechsel.</i>	Folsäureantagonist:	Methotrexat
	Pyrimidinanaloga:	5-Fluorouracil, Capecitabin, Cytosinarabinosid (Cytarabin), Gemcitabin
	Purinanaloga:	6-Thioguanin, Pentostatin, Azathioprin, 6-Mercaptopurin, Fludarabin

Bei den Polychemotherapien gibt es verschiedene Schemata, welche Zytostatika als Kombination verwendet werden. Standard-Schemata sind z.B. das CMF Schema (Cyclophosphamid, Methotrexat und 5-Fluorouracil) und die EC oder AC Schemata (Epirubicin oder Adriamycin und Cyclophosphamid). Eine Abwandlung der EC/AC Schemata ist das FEC/FAC Schema (5-Fluorouracil, Epirubicin oder Adriamycin und Cyclophosphamid). Die Auswahl des Schemata, der Dosierung und der Intervalle hängt - falls vorhanden - von prädiktiven Markern sowie von der Prognose der Patientin und ihrem allgemeinen Gesundheitszustand ab.

1.1.4. Anwendung molekularer Biomarker

Es gibt vier wesentliche Entwicklungsschwerpunkte der Brustkrebsforschung. Zunächst ist man auf der Suche nach Prädispositionsmarkern, mit denen sich das Risiko einer Frau, an Brustkrebs zu erkranken, abschätzen lässt. Tumormarker können für die Früherkennung und Klassifizierung der Tumore verwendet werden. Um den Verlauf einer Krebserkrankung vorherzusagen, werden prognostische Marker und für die Vorhersage der Wirksamkeit einer Therapie prädiktive Marker gesucht.

Nur etwa 5% der Brusttumore sind erblich, d.h. familiären Ursprungs [11]. Bisher ist nicht geklärt, wie viele verschiedene Gene beteiligt sind. Die zwei bekanntesten sind die Brustkrebsgene und Tumorsuppressoren *BRCA1* [12] und *BRCA2* [13]. Grundsätzlich gilt, dass die Früherkennung einer Krebserkrankung die Prognose eines Patienten wesentlich verbessert. Die Mehrheit der Brusttumore sind aber sporadische, d.h. sie sind nicht auf einen erblichen Faktor als nachweisbares erhöhtes Risiko zurückzuführen. Weiterhin treten erst in späten Stadien Symptome auf. Daher wird für die Früherkennung nach Tumormarkern gesucht, welche früh während seiner Entwicklung auftreten.

Das krankheitsfreie Überleben und somit auch das Gesamtüberleben von Patientinnen mit Tumoren gleicher klinischen Parametern können sehr unterschiedlich sein, d.h. die bisher verwendeten Parameter für die Erstellung der Prognose sind in vielen Fällen nicht ausreichend [7, 14]. Weiterhin haben die in den vergangenen Jahren verbesserten bildgebenden Verfahren (Mammographie und Ultraschall) zum Auffinden einer erhöhten Zahl verschiedener Vorstufen von Brustkrebs bzw. von Gewebeveränderungen, deren Entwicklung ungewiss ist geführt. Besonders in diesen Fällen ist eine gute Klassifizierung wesentlich um eine aggressive Veränderung, die eine Behandlung notwendig macht, von einer gutartigen Veränderung zu unterscheiden. Andererseits beruht die Mortalität von Brustkrebs auf der Bildung von Metastasen und Brustkrebs ist im metastasierenden Stadium nicht mehr heilbar. Daher ist das Ziel der Klassifizierung und der Erstellung einer Prognose genau zu bestimmen, welche Therapieform notwendig ist, um das Auftreten von Metastasen zu verhindern und gleichzeitig jede unnötige Therapie zu vermeiden.

Hauptsächlich soll die Aggressivität des Tumors bestimmt werden, da ein aggressiver Tumor schneller Metastasen bildet. Bei aggressiven Tumoren, ist die Wahrscheinlichkeit hoch, dass sich bereits nicht detektierbare Metastasen im Körper der Patientin befinden, was eine weitere Therapie notwendig macht. Handelt es sich im Gegensatz dazu um einen nicht-aggressiven Tumor, ist die Wahrscheinlichkeit groß, dass das gesamte Tumormaterial entfernt wurde und keine bzw. eine milde Therapie ausreichend ist. In diesem Fall haben die Patientinnen wesentliche Nachteile durch eine Chemotherapie, da sowohl die während der Chemotherapie auftretenden Nebenwirkungen die Lebensqualität stark verringern als auch die Langzeitfolgen, wie z.B. die Erkrankung an einer Leukämie lebensbedrohlich sein können [15, 16].

Ein anderes Problem ist die Resistenz von Tumoren gegenüber bestimmten Therapien bzw. die unterschiedliche Wirksamkeit einer bestimmten Therapie bei verschiedenen Patientinnen. Molekulare Marker, die das Verhalten eines Tumors gegenüber einer Therapie vorhersagen, nennt man prädiktive Marker. Sie sind sowohl für die adjuvante als auch für die palliative Therapie von Bedeutung. In beiden Fällen möchte man die Durchführung einer Chemotherapie verhindern, wenn der Tumor auf diese nicht anspricht bzw. gezielt die Therapie auswählen können, welche die größte Wirkung bei einer Patientin zeigt. Da jede Therapie eine große Belastung für die Patientinnen darstellt, gibt es häufig Zweifel an der Notwendigkeit eine Therapie durchzuführen insbesondere wenn es sich um eine palliative Therapie handelt. Ist diese nämlich nicht wirksam, hat sie nicht die Verbesserung der Lebensqualität einer Patientin sondern eine Verschlechterung zur Folge. In anderen Fällen ist nach einer fehlgeschlagenen Therapie aufgrund der Nebenwirkungen die Anwendung einer weiteren möglicherweise wirksamen Therapie nicht mehr möglich. Das Auffinden prädiktiver Marker ist daher von besonderer Bedeutung.

1.1.5. Molekulare Biomarker in Brustkrebs

Zur Entstehung von malignen Tumoren kommt es, wenn das genetisch geregelte Gleichgewicht zwischen Zellzyklus, Zelldifferenzierung und Zelltod (Apoptose) gestört ist. Dieser Vorgang ist mit der Aktivierung von Onkogenen („gain-of-function“) und Inaktivierung von Tumorsuppressorgenen („loss-of-function“) verbunden [17]. Tumorgewebe ist durch eine Vielzahl genetischer Veränderungen gekennzeichnet. Dabei handelt es sich um Amplifikationen, Deletionen oder Insertionen einzelner Gene oder größerer Bereiche, den Verlust der Heterozygotie (LOH – loss of heterozygosity), welcher auf der Deletion eines Allels beruht, chromosomale Umordnungen bei denen durch Genfusionen neue Onkogene entstehen, Chromosomen Abberationen bis hin zur allgemeinen Aneuploidie und Punktmutationen die zu Einzelnukleotidpolymorphismen (SNPs – single nucleotide polymorphism) führen. Diese Veränderungen sind häufig für eine veränderte Expression oder Funktionalität der betroffenen Gene verantwortlich, wodurch häufig auch Veränderungen auf der Ebene der Proteine gefunden werden. Die Eigenschaften, die Zellen erwerben müssen damit sich aus ihnen ein bösartiger Tumor entwickelt, lassen sich in folgende sechs Gruppen aufteilen [17]: Unabhängigkeit von Wachstumsfaktoren; Unempfindlichkeit gegenüber Wachstumsinhibitoren; Umgehen der Apoptose; Neoangiogenese (Neubildung von Blutgefäßen); Fähigkeit zur unbegrenzten Teilung und die Fähigkeit zur Gewebeeinfiltration und Metastasenbildung.

Tumormarker für die Früherkennung und Klassifizierung von Brustkrebserkrankung sind bislang noch nicht in der klinischen Routine etabliert. Gesucht werden vor allem Marker, die in leicht zugänglichen Körperflüssigkeiten (z.B. Blut) zu einem frühen Zeitpunkt der Erkrankung nachgewiesen werden können.

Sowohl prognostische als auch prädiktive Biomarker können meist auf Gewebeproben analysiert werden. Häufig ist keine klare Abgrenzung dieser Biomarker möglich, da ein guter Krankheitsverlauf und das Ansprechen auf eine bestimmte Therapie miteinander verbunden sind. Die bekanntesten prognostisch-prädiktiven Biomarker sind der bereits erwähnte Östrogenrezeptor (ER) und der Progesteronrezeptor (PgR) (siehe Abschnitt 1.1.2), welche seit vielen Jahren standardisiert in der klinischen Routine eingesetzt werden.

ER und PgR negative Tumoren sind oftmals mit anderen Biomarkern assoziiert. Dazu gehören z.B. die Amplifikation der Onkogene *HER-2/neu*, *c-Myc* und *int-2*, mit Mutationen des Tumorsuppressorgens *p53* und Überexpression von Invasions- und Metastasen-assoziierten Wachstumsfaktoren, deren Rezeptoren und einige Proteasen, welche mit einer schlechte Prognose verbunden sind [6, 10]. Die Bestimmung der *HER-2/neu* Expression ist inzwischen eine Standarduntersuchung in der klinischen Routine [18]. Die Überexpression des entsprechenden Proteins aufgrund einer Amplifikation des *HER-2/neu* Locus ist verbunden mit einer schlechten Prognose, sowohl bei Lymphknoten-negativen, als auch Lymphknoten-positiven Patientinnen [19]. Die Expressionsdaten sind ebenfalls prädiktiv für das Ansprechen auf die Antikörpertherapie mit dem entsprechenden *HER-2/neu* Antikörper (Trastuzumab, Herceptin®).

Weiterhin gibt es zurzeit lediglich zwei weitere Tests mit FDA Zulassung. Dabei handelt es sich um den CellSearch® Test (Veridex, USA), mit dem zirkulierende Tumorzellen im Blut von Brustkrebspatientinnen nachgewiesen werden und eine prognostische Aussage über den Krankheitsverlauf gemacht wird [20, 21]. Daneben gibt es seit 2007 einen RNA Chip-basierten Test (MammaPrint®, Agendia, NL), mit dem eine Prognose mit Hilfe des Expressionsmusters von 70 Genen erstellt wird [22-24]. Einen guten Überblick über diese und weitere prognostische und prädiktive Biomarker, den entsprechenden Wirkstoffen und die klinischen Relevanz der Marker gibt der Übersichtsartikel von Ross et al. 2005 [7].

1.1.6. Methylierung und Brustkrebs

Betrachtet man die Komplexität der Merkmale bzw. Veränderungen, die für die Entwicklung einer Tumorzelle notwendig sind, stellt sich die Frage, wie sie in einer zunächst gesunden Zelle entstehen können. Die Mutation einzelner Gene wäre dafür nicht ausreichend, zumal es ein komplexes System in der Zelle gibt, das Mutationen und karyotypische Veränderungen repariert oder eliminiert. Daher ist es unwahrscheinlich, dass die Vielzahl von Veränderungen, die in einem Tumor vorliegen, sich während der normalen Lebenszeit eines Menschen in Form von Mutationen ansammelt [17].

Neben den genetischen gibt es epigenetische Veränderungen, welche für die Entwicklung von Krebszellen verantwortlich sind. Als epigenetisch bezeichnet man eine Eigenschaft, die vererbbar ist und nicht durch Unterschiede in der DNA-Sequenz hervorgerufen wird.

Epigenetische Veränderungen treten in Krebszellen häufiger auf als genetische [25] und sind daher sehr gut als molekulare Marker geeignet. Eine epigenetische Modifikation ist die DNA-Methylierung der DNA. Veränderungen der DNA-Methylierung werden in Tumoren vielfach beobachtet. Dies hat zu einer Ergänzung von Knudsons Hypothese geführt, wonach eine Krebszelle zwei Möglichkeiten hat, einen Tumorsuppressor auszuschalten, den Verlust der Heterozygotität oder die homozygote Deletion. Zu diesen beiden kommt nun als dritte Möglichkeit das Ausschalten der Expression durch eine epigenetische Veränderung.

1.1.6.1. Die Funktion der Methylierung in gesunden Zellen

Methylierung wird durch die enzymatische Modifikation von Cytosin hervorgerufen. Dabei wird eine Methylgruppe an die fünfte Position des Cytosinringes angehängt. Das entstandene 5-Methylcytosin wird bei der Replikation und Transkription als Cytosin erkannt [26]. In Säugetieren werden hauptsächlich die Cytosine von 5'-CpG-3' Dinukleotiden methyliert. Der Anteil dieses Dinukleotides ist im Säugetiergenom gegenüber der erwarteten Häufigkeit bei Gleichverteilung aller vier Basen stark vermindert [27]. Grund dafür ist die überdurchschnittliche Mutation von Methylcytosin zu Thymin, welches als natürlich vorkommende Base keinem Reparaturmechanismus unterliegt. Es gibt allerdings Regionen im Genom deren CpG-Anteil den zu erwartenden Wert für die Häufigkeit dieses Dinukleotids annähernd erreicht [28, 29]. Solche Sequenzabschnitte werden als CpG-Inseln bezeichnet. DNA-Methylierungsmuster werden im Anschluss an die Replikation wieder vervollständigt und sind daher mitotisch stabil. Andererseits wird die DNA-Methylierung der DNA in Embryonen fast vollständig „neu-programmiert“ [30], d.h. DNA-Methylierung ist eine Modifikation, die einerseits während der Zellteilung stabil weitergegeben werden kann und andererseits potentiell veränderlich ist.

Wachsende Erkenntnisse ergeben ein komplexes Bild eines Zusammenspiels von Methylierung mit Histon-Modifikationen, den dazugehörigen Proteinen und der aktiven Transkription einer Region. Dieses Zusammenspiel legt die Chromatinstruktur der betreffenden Region fest. Wobei methylierte DNA Abschnitte in eine dichte Chromatinstruktur (Heterochromatin) verpackt werden, deren Expression reprimiert ist und nicht methylierte Abschnitte als offenes Chromatin (Euchromatin) vorliegen [31, 32, 33]. Durch die beschriebenen Eigenschaften ist DNA-Methylierung an verschiedenen biologischen Funktionen beteiligt, wie der Inaktivierung eines X-Chromosoms in weiblichen Zellen und dem genomischen Imprinting, d.h. der Expression oder Repression bestimmter Gene in Abhängigkeit von ihrer elterlichen Abstammung [34]. Weiterhin wird angenommen das Methylierung auch an der Kontrolle der gewebespezifischen Genexpression und deren Festlegung im Laufe der Entwicklung und Differenzierung beteiligt ist [31, 32]. Eine andere wesentliche Funktion ist die Unterdrückung beweglicher DNA Elemente und der homologer Rekombinationen [31].

Circa 70 % aller CpGs im Genom gesunder Zellen sind methyliert. CpG-Inseln befinden sich in der Nähe des Transkriptionsstartes (TSS) von circa der Hälfte aller humanen Gene. Grundsätzlich wird davon ausgegangen, dass ausschließlich CpG-Inseln in gesunden Zellen nicht-methyliert sind. Ausnahmen bilden solche, die sich in Imprinting Regionen oder auf dem inaktivierten X-Chromosom befinden. In den letzten Jahren wurde aber auch deutlich, dass auch autosomale CpG-Inseln, die nicht dem Imprinting unterliegen, methyliert sein können. Bei diesen CpG-Inseln nimmt man an, dass sie der Regulation gewebespezifischer Expression dienen [35]. Im Gegensatz dazu sind die DNA-Abschnitte, welche der Kontrolle von Transposons oder der homologen Rekombination dienen, in gesunden Zellen methyliert. Es handelt sich dabei sowohl um CpG Inseln als auch um CpG ärmere Sequenzabschnitte, welche in den verschiedenen Regionen des Genoms wie in Introns oder in genarmen Sequenzabschnitten liegen können.

1.1.6.2. DNA-Methylierung und Brustkrebs

Abnormale DNA-Methylierung und epigenetische Veränderung sind an der Entstehung von Brustkrebs wesentlich beteiligt [25]. Es gibt Hinweise, dass die DNA-Methylierung von CpG-Inseln mit dem Alter zunimmt [36, 37] während sie generell mit dem Alter sinkt [38]. Ähnlich lassen sich die Veränderungen in Krebszellen beschreiben, in denen diese Effekte jedoch stärker ausgeprägt sind.

In den letzten Jahren konnte für viele in gesunden Zellen unmethylierten CpG-Inseln verschiedener Gene, eine unterschiedlich starke Hypermethylierung in Krebs nachgewiesen werden [39]. Bei globaler Betrachtung sinkt jedoch der Anteil methylierter CpGs in Krebszellen im Vergleich zu gesundem Gewebe ab [40]. Man spricht von globaler Hypomethylierung.

Wie schon erwähnt, steigt das Risiko einer Frau, an Brustkrebs zu erkranken, mit zunehmendem Alter. Brustkrebs ist bei Frauen unter 30 Jahren sehr selten. Danach steigt die Neuerkrankungsrate kontinuierlich aber ab der Menopause (40–50 Jahre) langsamer an. Diese Beobachtung lässt sich durch die Beteiligung von Sexualhormonen an der Entwicklung von Brustkrebs erklären [41]. Darauf weisen auch andere epidemiologische Beobachtungen hin, wie z.B. das erhöhte Risiko von Frauen mit früher erster Menstruationsblutung und später Menopause oder das erhöhte Risiko übergewichtiger postmenopausaler Frauen, da nach der Menopause Fettgewebe die Hauptquelle für Östrogen ist, sowie die Verringerung des Risikos durch eine postmenopausale Hormontherapie. Östrogen ist ein Wachstumsfaktor, der direkt die Epithelzellen aus denen sich der Tumor bildet beeinflusst. Zusammenfassend ergibt sich also folgendes Bild. Mit zunehmendem Alter akkumulieren Zellen krankhafte sowohl genetische als auch epigenetische Veränderungen. Das dauerhafte Vorhandensein von Wachstumsfaktoren hauptsächlich Östrogen im Brustgewebe steigert die Wahrscheinlichkeit für solche Veränderungen und deren Ausprägung durch die permanente Stimulation der Zellteilung in den Epithelzellen.

Hypermethylierung von CpG-Inseln ist die am besten untersuchte epigenetische Veränderung, die z.B. die Unterdrückung der Expression von Tumorsuppressorgenen zur Folge hat. Die von Hanahan und Weinberg [17] eingeführte Unterteilung, der von Krebszellen erworbenen Eigenschaften, wird 2002 von Widschwendter und Jones [25] auf veränderte DNA-Methylierung in Brustkrebs angewendet. In dem Übersichtsartikel beschreiben sie Gene für welche direkt oder indirekt nachgewiesen wurde, dass deren veränderte DNA-Methylierung zur Entwicklung von Brustkrebs beiträgt. Es wird deutlich, dass die Hypermethylierung von CpG-Inseln einer Vielzahl von Genen einen wesentlichen Anteil an der Ausprägung der jeweiligen Eigenschaften hat. Bei der zum anderen beschriebenen Hypomethylierung handelt es sich um einen Verlust der DNA-Methylierung auch außerhalb von CpG-Inseln. Während die Rolle hypermethylierter CpG-Inseln hauptsächlich dem Ausschalten von Tumorsuppressor-Genen zugeordnet wird, erscheint die Rolle der Hypomethylierung noch wesentlich weniger verstanden [42]. Betroffene Regionen sind in vielen der bisher beschriebenen Fälle CpG-Inseln in Promotorregionen aber auch andere Regionen des Genoms, wie z.B. Regionen innerhalb von Genen oder repetitive Sequenzen. Die Hypomethylierung von CpG-Inseln in Promotorregionen hat vermutlich die Aktivierung von Onkogenen zur Folge. Der globale Verlust der DNA-Methylierung führt zur Reaktivierung beweglicher DNA-Elemente, zur gehäuften Rekombination und somit zur Instabilität des Genoms, welche häufig in Krebszellen beobachtet wird [43].

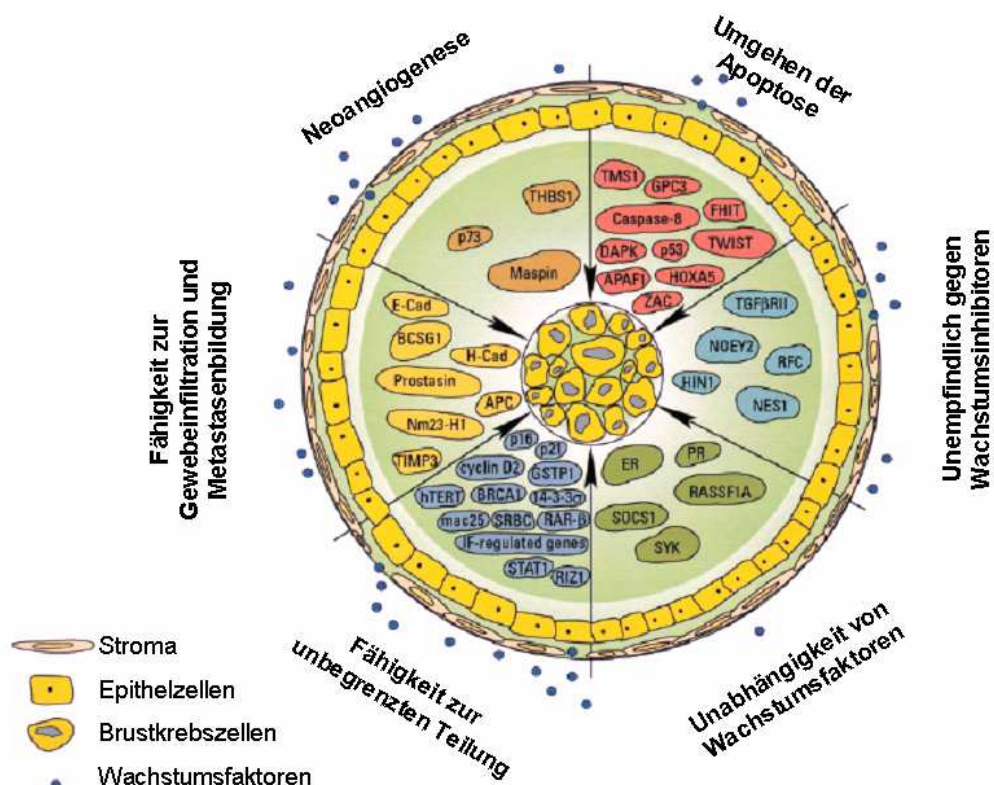


Abbildung 2: Gene die fehlerhafte Methylierung in Brustkrebszellen aufweisen unterteilt in die ausgelösten Eigenschaften nach Widschwendter und Jones [25].

1.1.6.3. **DNA-Methylierung als molekularer Biomarker.**

Veränderungen der DNA-Methylierung sind für alle Fragestellungen der Krebsforschung relevant. Sie treten schon früh während der Entstehung eines Tumors auf [44] und sind oft tumorspezifisch [35], wodurch DNA-Methylierung sehr gut für die Früherkennung eingesetzt werden kann. Die DNA-Methylierung hat gegenüber anderen Biomarkern wie Expressionsunterschieden, veränderten Proteinkonzentrationen oder Mutationen einige Vorteile. Zunächst basiert sie auf der kovalenten Bindung der Methylgruppe an DNA. DNA ist ein wesentlich stabileres Molekül als RNA und daher weniger anfällig für Degradierung. Auch die Bindung der Methylgruppe ist sehr stabil, d.h. im Vergleich zu RNA, ist die DNA-Methylierungsinformation wesentlich weniger empfindlich. Das bedeutet auch, dass eine Analyse auch noch nach einer histopathologischen Routinebehandlung, wie der Fixierung in Formalin, möglich ist. Ein weiterer Vorteil ist, dass in Körperflüssigkeiten enthaltene TumordNA durch die sensitive Detektion tumorspezifischer DNA-Methylierungsmarker nachgewiesen werden kann [35]. DNA-Methylierung betrifft meist mehrere CpGs gleichzeitig (Komethylierung) was einen Vorteil gegenüber Punktmutationen darstellt, da die Untersuchung bzw. Anreicherung erleichtert wird. Da es sich um DNA handelt, kann ein DNA-Methylierungsmarker amplifiziert werden, d.h. er ist schon in sehr geringen Mengen detektierbar. DNA-Methylierung ist zum einen relativ stabil auf der anderen Seite unterliegt sie aber Veränderungen durch Umwelteinflüsse, Alter, etc. Daher sind DNA-Methylierungsmarker viel wahrscheinlicher in allen nicht erblichen Krebsarten als z.B. Keimbahnmutationen bzw. vererbten Punktmutationen.

In den letzten Jahren wurden Marker oder Markergruppen identifiziert, die die Unterscheidung von Subgruppen verschiedener Tumoren möglich machten. Diese Subgruppen weisen meist Unterschiede in der Aggressivität der untersuchten Krebsart auf, d.h. es handelt sich um prognostische Marker. Bisher konnten DNA-Methylierungsmarkergruppen verwendet werden, um z.B.: Neuroblastome [45], Nierenkrebs [46] und Lungenkrebs [47-49] in klinisch relevante Subgruppen zu unterteilen. Die Daten einer der ersten genomweiten Analysen mittels DMH von Wei et al. ermöglichten die Unterscheidung von aggressivem und weniger aggressivem Eierstockkrebs [50, 51].

Auch für einzelne Gene wurde bereits gezeigt, dass ihr DNA-Methylierungsmuster als prognostischer Marker verwendbar ist. Hypermethylierung von *E-Cadherin* ist verbunden mit einem kürzeren krankheitsfreien Überleben bei Magenkrebs [52, 53] sowie bei Mundhöhlenkarzinomen [54]. Hypermethylierung von *E-Cadherin* und *H-Cadherin* ist ein Marker für ein hohes Risiko für einen Rückfall bei Gebärmutterhalskrebs. Auch die Kombination von *E-Cadherin* mit *APC* macht die Identifizierung einer Gruppe von Magenkrebspatienten mit einer negativen Prognose möglich [55]. Das DNA-Methylierungsmuster von *Laminin-5* ist ein Marker für eine negative Prognose in Prostata- [56] und Blasenkrebs [57]. Die DNA-Methylierung von *p16* [58, 59]), *IGFBP-3* [60], *APC* [61] oder *RASSF1A* [59] korreliert mit einer verkürzten Überlebensrate bei Lungenkrebs. Hypermethylierung von *p16* und *p15* sind

Marker für eine schlechte Prognose bei akuter myeloischer Leukämie [62] und Hypermethylierung von *Calcitonin* [63] und *p21* [64] ist verbunden mit einem erhöhten Rückfallrisiko bei akuter lymphatischer Leukämie. Hypermethylierung von *GSTP1*, *APC* oder *cyclin D2* konnte zur Identifizierung von Patienten mit einem erhöhten Rückfallrisiko bei Prostatakrebs verwendet werden [65].

Eines der bekanntesten Gene, für welches DNA-Methylierung in Brustkrebs als prognostischer Faktor gefunden wurde, ist der *Plasminogenaktivator vom Typ Urokinase (uPA)* [66], welches auch in klinischen Brustkrebsproben gezeigt werden konnte [67-69]. Die Promotormethylierung von sechs mit der Tumorentstehung assoziierten Genen (*RASSF1A*, *APC*, *TWIST*, *CDH1*, *GSTP1* und *RARB*) in 151 Brusttumoren wurde von Shinozaki et al. [70] analysiert und gezeigt, dass die Methylierung dieser Gene mit dem bekannten prognostischen Faktor des Lymphknotenstatus korreliert ist. Die Methylierung von *COX-2* konnte bei 110 Tumoren mit prognostischen Faktoren wie Tumorgroße und Anzahl der befallenen Lymphknoten korreliert werden [71]. Die prognostische Eigenschaft der DNA-Methylierung von *SFRP1* wurde kürzlich an 130 Brustkrebspatientinnen gezeigt [72]. Eine Analyse der Promotormethylierung von *PITX2* bei 415 unbehandelten Patientinnen ergab, dass die Promotormethylierung dieses Genes als prognostischer Marker verwendet werden kann [73].

Das Gebiet der Chemotherapieresistenz ist komplexer als die Prognose. Trotzdem konnten auch für die Chemotherapieresistenz bereits einige Marker gefunden werden. Hypermethylierung der *O-6-methylguanin-DNA-methyltransferase (MGMT)* und des mismatch repair Gens *hMLH1* sind als prädiktive Marker etabliert [74]. Aber auch weitere Gene wurden in jüngster Zeit gefunden. Zum Beispiel führte die veränderte DNA-Methylierung des DNA Reparatur Gens *ATM* zu einer erhöhten Sensitivität von Darmkrebszelllinien gegen eine Strahlentherapie [75] und die Sensitivität von Magenkrebs gegenüber mitosehemmenden Medikamenten ist abhängig von der Hypermethylierung des Gens *CHFR* [76]. Eine Unterbrechung der Apoptose durch Hypermethylierung des Gens *Caspase-8* wurde mit der Resistenz verschiedener Tumorzelllinien gegen eine zytotoxische Chemotherapie in Verbindung gebracht [77] genauso wurde gezeigt, dass auch die Expression des Gens *MDR1* (multi drug resistance) durch DNA-Methylierung beeinflusst ist. Diese Studien sind größtenteils funktionale Studien, d.h. meist wird mit Hilfe von Zelllinien gezeigt, dass das entsprechende Gen hypermethyliert ist und die Reaktivierung durch Behandlung mit einem Demethylierungsagens wie 5-Azacytidin zur Wiederherstellung der Sensitivität bzw. zur Resistenz führt. Die Respons von Brustkrebspatientinnen auf eine Hormontherapie wurden jedoch bereits in Patientengruppen analysiert und Hypermethylierung der Gene *PSAT1* [78] oder *ESR1* [79] wurde als prädiktiver Marker für die Vorhersage eines Erfolgs einer Tamoxifen Behandlung gefunden. Weiterhin wurde nachgewiesen, dass die Hypermethylierung von *RASSF1A* mit der Respons von Patientinnen auf eine adjuvante Tamoxifen Behandlung verbunden ist [80].

1.2. Methoden für die Analyse der DNA – Methylierung

Im vorangegangenen Kapitel wurde gezeigt, dass DNA-Methylierungsmarker Vorteile gegenüber anderen molekularen Markern aufweisen und vielseitig anwendbar sind. Bei einer DNA-Methylierungsanalyse wird grundsätzlich eine Probe genomischer DNA analysiert, die einen heterogenen Mix verschiedener DNA Moleküle aus verschiedenen Zellen des analysierten Gewebes darstellt. Die DNA-Methylierungsanalyse wird in die in Abbildung 3 dargestellten prinzipiellen Analysearten unterteilt: A) Globale Analyse: Bestimmung des Gesamtanteils von Methylcytosin in der analysierten DNA, B) DNA-Methylierungslevel: durchschnittliche DNA-Methylierung eines einzelnen CpGs des DNA-Gemisches, C) DNA-Methylierungsprofile: DNA-Methylierung verschiedener einzelner CpGs bzw. Loci und D) DNA-Methylierungsmuster: Analyse mehrerer cis-ständiger CpGs [81; 35].

Grundsätzlich ist auch noch die Analyse eines DNA-Methylierungsmusterprofils möglich, wobei die DNA-Methylierungsmuster cis-ständiger CpGs verschiedener unabhängiger Loci untersucht werden. Die Analyse von DNA-Methylierungsmustern stellt die größte Herausforderung dar, da hierfür nicht die durchschnittliche DNA-Methylierung von CpGs bestimmt werden kann, sondern die DNA-Methylierung einzelner CpGs, wie es z.B. durch die Bisulfitequenzierung von Klonen möglich ist.

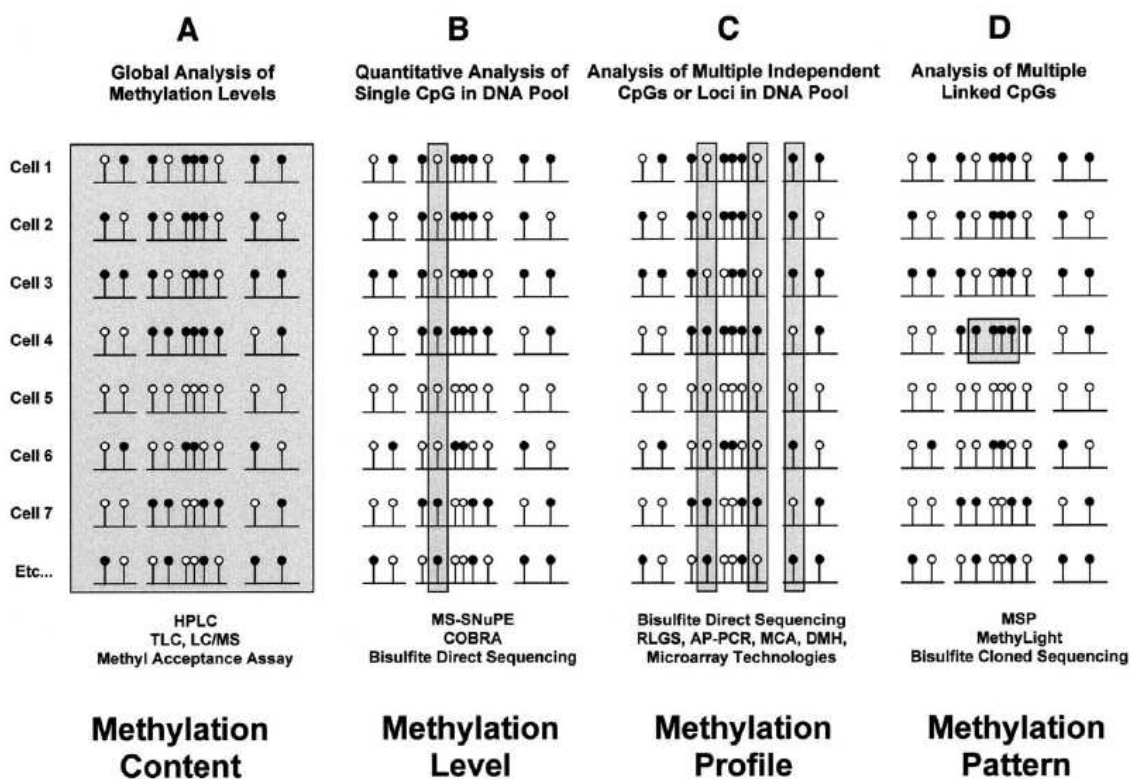


Abbildung 3: Vier Arten der Methylierungsanalyse aus Siegmund and Laird 2002 [81]: Horizontale Linien stellen das gesamte haploide Genom einer einzelnen Zelle, aufgeteilt in drei repräsentative Chromosomen mit verschiedenen CpG-Positionen, dar. Weiße Kreise sind unmethylierte und schwarze Kreise sind methylierte CpGs.

Auf der anderen Seite wird in vielen Fällen hauptsächlich nach einem bestimmten DNA-Methylierungsmuster nämlich der Komethylierung mehrerer CpGs gesucht. Geht man vom Auftreten von Komethylierung aus, kann die Analyse eines einzelnen CpGs dazu dienen, auf die DNA-Methylierung der benachbarten CpGs zu schließen. Besonders für die Suche nach neuen Markerkandidaten werden bevorzugt genomweite Methoden verwendet, d.h. es handelt sich um die Bestimmung von DNA-Methylierungsprofilen. Die Abgrenzung zwischen Levelprofilen und Musterprofilen, wie sie von Peter Laird [35] vorgenommen wird, ist aus den beschriebenen Gründen häufig zweitrangig.

Methoden für die Bestimmung des Gesamtanteils von Methylcytosin in der DNA der Zellen bzw. des untersuchten Gewebes sind hauptsächlich chromatografische Verfahren wie Dünnschichtchromatografie [82] und HPLC [83, 84] aber auch immunochemische Analysen [85] und Kapillarelektrophorese [86].

Die anderen Analysearten basieren in den meisten Fällen auf der Verwendung methylierungssensitiver Restriktionsenzyme, der Konvertierung unmethylierter Cytosine in Uracil durch die Reaktion mit Bisulfit oder einer Kombination aus beiden. Die Entdeckung methylierungssensitiver Restriktionsenzyme macht es erstmals möglich, sequenzspezifisch DNA-Methylierung einzelner CpGs zu analysieren. Der größte Teil der methylierungssensitiven Restriktionsendonukleasen sind Typ II Endonukleasen, die in ihrer Erkennungssequenz mindestens ein CpG enthalten und bei DNA-Methylierung der entsprechenden CpGs gehemmt sind. Das Restriktionsenzym *McrBC* hingegen schneidet nur dann, wenn die Erkennungssequenz methyliert ist. Ein großer Fortschritt war die Entwicklung der Bisulfitkonvertierung genomischer DNA durch Frommer et al. [87]. Durch die chemische Reaktion genomischer DNA mit Natriumbisulfit werden ausschließlich alle unmethylierten Cytosine in Uracil umgewandelt während methylierte Cytosine erhalten bleiben. Die Bisulfitreaktion erfolgt nur in Einzelstrang-DNA, dadurch kommt es zur Entstehung zweier nicht komplementärer Bisulfit-DNA-Stränge, in welchen DNA-Methylierung in einen Sequenzunterschied umgewandelt wird, d.h. aus unmethylierten CpGs werden UpGs, die sich von den erhaltenen CpGs an methylierten Positionen unterscheiden. Beide Bisulfit-Stränge sind nach der Umwandlung arm an Cytosin, man spricht vom C-armen Strang. In der anschließenden PCR verhält sich Uracil wie Thymin. Der entstehende komplementäre Strang ist folglich G-arm.

1.2.1. Methoden zur Bestimmung von DNA-Methylierungsmustern und DNA-Methylierungslevels

Die durch die Bisulfitreaktion bewirkte Umwandlung der DNA-Methylierung in eine Sequenzinformation wird bei vielen Methoden verwendet, wie z.B. der Bisulfitsequenzierung [87], MSP [88], COBRA (kombinierter Bisulfit-Restriktionsassay; engl: Combined Bisulfite Restriction Assay) [89], MethyLight [90-92] und Pyrosequenzierung [93]. Eine Liste mit weiteren Methoden ist z.B. in Peter Lairds Übersichtsartikel [35] zu finden. In den meisten Fällen handelt es sich um Methoden zur Analyse des DNA-Methylierungslevels einzelner CpGs bzw. der DNA-Methylierungsmuster von CpGs einzelner Sequenzabschnitte. DNA-Methylierungsprofile können durch diese Methoden nur in mittlerem Umfang mit Hochdurchsatz-Protokollen erreicht werden, da alle aufwendige Laborarbeit beanspruchen. Ein weiteres Hindernis ist die verminderte Komplexität der C-armen bzw. G-armen Bisulfit-DNA im Vergleich zu genomischer DNA, was für Hybridisierungen und beim Design von Primern wichtig ist.

Bisulfitsequenzierung

Bei der Bisulfitsequenzierung [87] werden Primer verwendet, die keine CpGs enthalten. Daher werden methylierte und unmethylierte DNA in gleichem Maße amplifiziert. Bei der direkten Sequenzierung wird das gesamte PCR-Produkt in die Sequenzierreaktion eingesetzt. Dadurch werden C/T Polymorphismen als Sequenzierergebnis erhalten, die sich in entsprechende durchschnittliche DNA-Methylierungswerte durch eine software-basierte Auswertung [94] übersetzen lassen. Eine andere Möglichkeit ist die Klonierung des PCR-Produktes und anschließende Sequenzierung einzelner Klone. Aus den Ergebnissen einer Vielzahl von Klonen kann wiederum die durchschnittliche DNA-Methylierung ermittelt werden. Bei der Pyrosequenzierung erfolgt die Quantifizierung durch die Detektion von Licht, das bei Zugabe eines Nukleotides durch eine Enzymkaskade freigesetzt wird.

COBRA

COBRA (kombinierter Bisulfit-Restriktionsassay; engl: Combined Bisulfite Restriction Assay) [89] ist eine weitere Methode, den bei einer PCR entstehenden C/T Polymorphismus zu nutzen, ist die Verwendung von Restriktionsenzymen mit CpGs in der Erkennungssequenz. Dadurch wird nur der methylierte Anteil der PCR-Produkte geschnitten. Bei der Analyse der geschnittenen PCR-Produkte durch Gelelektrophorese kann die DNA-Methylierung durch densitometrische Auswertung der Bandenstärke abgeschätzt werden.

MSP

Im Gegensatz zur Bisulfitsequenzierung werden bei der methylierungsspezifischen PCR (MSP) [88] Primer verwendet, die mindestens eine CpG Position abdecken. Dadurch wird entweder nur methylierte oder unmethylierte DNA amplifiziert. Der Anteil kann mit einem Agarosegel oder einer Realtime-PCR ermittelt werden. Für die Verwendung einer Realtime-PCR zur quantitativen Bestimmung von DNA-Methylierung führten Eads et. al die Bezeichnung MethyLight [90-92] ein. Hierbei kann es sich entweder um einen Realtime MSP-Assay handeln [91] oder die DNA-Methylierung wird mittels der Detektionssonde [90-92] selektiert.

Bisulfit-Mikroarrays

In den letzten Jahren wurden auch Methoden vorgestellt, um DNA-Methylierungslevel verschiedener Bisulfit-PCR-Produkte mit Hilfe von Mikroarrays parallel zu bestimmen [95-97]. Eine besondere Ausführung ist die Verwendung eines sog. BeadArrays [98]. Aufgrund der bereits erwähnten verringerten Komplexität von Bisulfit-DNA finden diese Methoden bisher ausschließlich Anwendung bei der Analyse einer mittleren Anzahl von Fragmenten.

1.2.2. Genomweite Analysemethoden von DNA-Methylierungsprofilen

Bei der Suche nach neuen Markern ist es vorteilhaft, möglichst viele Regionen parallel zu untersuchen, d.h. im besten Fall eine genomweite Analyse vorzunehmen. Besonders geeignet hierfür ist die Nutzung von Mikroarrays, da sie momentan das beste Verhältnis zwischen der Zahl parallel analysierbarer Fragmente und dem technischen und finanziellen Aufwand aufweisen. Vor der Einführung von Mikroarrays wurden andere Methoden, wie z.B. Sequenziergele für die methylierungsspezifische MS-AP-PCR [99], subtraktive Hybridisierung bei MCA [100] oder ein zweidimensionales Gel bei MS-RLGS [101, 102] verwendet. Auch einige neuere Methoden nutzen keine Mikroarrays. Dazu gehören ICEAMP [103], die Restriktionsmethode von Strichman-Almashanu et al. [104] oder MSDK [105]. Eine weitere Möglichkeit genomweite DNA-Methylierungsmuster zu bestimmen, ist die Vorhersage der DNA-Methylierung aufgrund spezifischer Sequenzeigenschaften [106].

Die erste Methode bei der ein Mikroarray verwendet wurde, wurde 2001 von Yan et al. publiziert [107]. Darin wurde die zuvor von dieser Gruppe beschriebene Methode, der Differentiellen DNA-Methylierungs Hybridisierung (englisch: differential methylation hybridisation, DMH, [108]) auf einen CpG-Insel-Mikroarray übertragen. Eine andere Methode wurde 2002 von Tompa et al. [109] beschrieben und in den letzten Jahren kamen weitere hinzu [110-115].

1.2.2.1. **Methoden ohne Mikroarrays**

MS-AP-PCR

Methylierungsspezifische Arbiträre Primer PCR [99] basiert auf der Verwendung kurzer Primer, die mehrere CpG-reiche Sequenzen amplifizieren und daher bei der Analyse des PCR-Produktes mittels eines Sequenziergels spezifische Bandenmuster ergeben. Für die DNA-Methylierungsanalyse wird die DNA zuvor mit einem methylierungssensitiven Enzym geschnitten. Der Vergleich mit einer ungeschnittenen Probe zeigt, welche Fragmente geschnitten wurden, d.h. nicht methyliert waren. Als Kontrollen dienen methylierungsunspezifische Isoschizomer und ungeschnittene DNA. Die Identifizierung der Sequenzen erfolgt durch Klonierung und Sequenzierung. Die Anzahl der analysierbaren Fragmente ist limitiert, da sie sich auf dem Gel noch gut auftrennen müssen. Grundsätzlich wäre es möglich, die Detektion mittels Gelelektrophorese durch einen Mikroarray zu ersetzen.

MCA

Bei der MCA Methode (englisch: methylated CGI amplification) [100] wird die DNA mit dem methylierungssensitiven Enzym *SMAI* gespalten, welches glatte DNA-Enden hinterlässt. Anschließend werden die verbliebenen methylierten Sequenzanteile mit dem nicht methylierungssensitiven Isoschizomer *XmaI* geschnitten, welches überhängende Enden erzeugt an die Adapter angehängt werden können. Bei einer mit Hilfe der Adapter vorgenommenen PCR werden nur CpG-reiche methylierte Sequenzen amplifiziert. Durch anschließende subtraktive Hybridisierung werden zwei Proben verglichen.

RLGS

RLGS steht für Restriktions Landmarken Genomisches Scannen (englisch: restriction landmark genomic scanning) [101, 102]. Bei der methylierungsspezifischen Ausführung wird die DNA zuerst mit einem methylierungssensitiven Enzym geschnitten, und die entstehenden Enden werden radioaktiv markiert. Dann wird die DNA durch ein weiteres Enzym noch einmal fragmentiert und anschließend auf einem zweidimensionalen Gel aufgetrennt. Da nur unmethylierte Sequenzen geschnitten und markiert wurden, ergeben auch nur solche Fragmente Signalpunkte auf dem Gel. Fehlende Signalpunkte beim Vergleich der Gele zweier verschiedener Proben zeigen daher unterschiedlich methylierte Fragmente auf. Auch in diesem Fall werden die Fragmente durch Klonierung und Sequenzierung identifiziert.

ICEAMP

Bei dieser Methode [103] wird eine Säule verwendet auf der die methylierte CpGs bindende Domäne von *MeCP2* immobilisiert wurde [116]. Angereicherte methylierte CpG-reiche Fragmente von zwei verschiedenen Proben werden anschließend subtraktiv hybridisiert und dann kloniert.

Restriktionsmethode von Strichman-Almashanu et al.

Bei der Restriktionsmethode von Strichman-Almashanu et al. [104] wird die DNA zunächst mit dem methylierungsunspezifischen Enzym MseI geschnitten, da die Erkennungssequenz TTAA relativ häufig im Genom und andererseits selten in CpG reichen Regionen vorkommt. Durch das Schneiden mit dem methylierungssensitiven Enzym Hpa II werden dann alle nicht methylierten CpG-reichen Regionen fragmentiert. Anschließend werden durch eine Gelfiltration Fragmente > 1 kb gewonnen, d.h. es werden hauptsächlich methylierte CpG-reiche Fragmente angereichert. Diese Fragmente werden kloniert und verlieren dadurch ihre DNA-Methylierung. In einem weiteren Anreicherungsschritt werden die demethylierten Fragmente mit einem 6 bp CpG-reichen Restriktionsenzym geschnitten und erneut kloniert. Bei der Analyse der erhaltenen Bibliothek finden die Autoren allerdings hauptsächlich repetitive DNA-Elemente. Das ist ein großes Problem besonders wenn Marker durch Sequenzierung von Klonbibliotheken identifiziert werden.

MSDK

Bei MSDK (englisch: methylation-specific digital karyotyping) [105] handelt es sich um eine Modifikation der digitalen Karyotypisierungstechnik, welche für die quantitative genomweite Analyse von Kopienzahlen entwickelt wurde. Die DNA wird mit einem methylierungssensitiven Restriktionsenzym geschnitten und an die entstehenden Enden werden biotinylierte Linker angehängt. Die DNA wird dann mit einem weiteren nicht methylierungssensitiven Enzym weiter fragmentiert und die entstehenden kurzen Fragmente mit Linker werden über Streptavidin gesammelt. Diese Methode reichert also nicht-methylierte Sequenzen an. Die Detektion erfolgt über Sequenzierung.

1.2.2.2. Methoden mit Mikroarrays

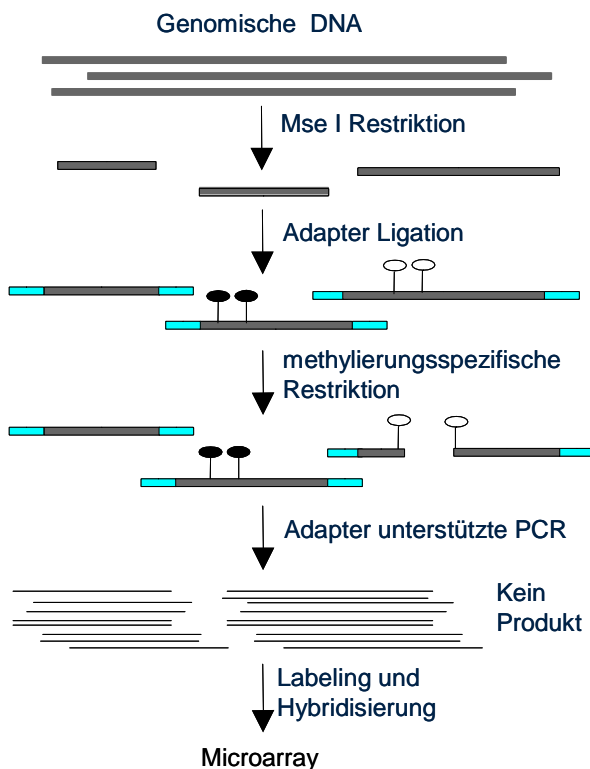


Abbildung 4: Fließschema DMH

DMH

Die Methode der Differentiellen Methylierungs Hybridisierung (englisch: differential methylation hybridisation, [108]) reichert methylierte CpG-reiche Fragmente in den Folgenden Schritten an: i) DNA Fragmentierung mit MseI (TTAA); ii) Adapter Ligation; iii) methylierungssensitive Restriktion mit HpaII und BstUI; iii) Adapter unterstützte Amplifikation.

Die Detektion erfolgt mit einem Mikroarray auf den Klone einer von Cross et al. [116] erstellten CpG-Insel-Bibliothek gespottet sind. Proben, die verglichen werden sollen, werden mit den Farbstoffen Cy-5 bzw. Cy-3 markiert und zusammen hybridisiert. Das Verhältnis aus den Signalen beider Proben wird für die Bewertung der DNA-Methylierung verwendet.

Restriktions Methode von Tompa et al.

Diese Methode [109] ist sehr einfach. DNA wird mit einem methylierungssensitiven Enzym geschnitten und anschließend nach Größen fraktioniert. Alle Fragmente die kleiner sind als 2,5 kb werden mit Cy-5 bzw. Cy-3 markiert und zwei Proben zusammen hybridisiert. Die Methode wurde zur Analyse von *Arabidopsis* Methylierungsmutanten auf einem Mikroarray mit zufällig ausgewählten Fragmenten verwendet. Für die Analyse von z.B. Tumorproben ist diese Methode nicht geeignet, da als Ausgangsmaterial zwischen 20-30 µg DNA notwendig sind.

Restriktions Methode von Lippman et al.

Bei dieser Methode [115] wird die DNA zunächst mechanisch zu Fragmente zwischen 1 kb bis 10 kb fragmentiert. Anschließend wird die DNA zu gleichen Teilen aufgeteilt. Ein Teil wird mit McrBC geschnitten, wodurch methylierte Fragmente abgereichert werden. Durch Ausschneiden aus einem Agarosegel werden von beiden DNAs Fragmente > 1 kb aufgereinigt, Cy-5 bzw. Cy-3 markiert und zusammen hybridisiert. Methylierte Fragmente zeigen ein vermindertes Signal gegenüber dem ungeschnittenen DNA-Anteil. Auch diese Methode benötigt mit mindestens 10 µg DNA zu viel Ausgangsmaterial.

Restriktionsmethode von Ching et al und Verwendung von BAC Mikroarrays

Bei dieser Methylierungsanalyse [110] werden die methylierungssensitiven Enzyme *NotI* oder *BssHII* verwendet. Die Hybridisierung erfolgt auf einen BAC Mikroarray. Durch vorherige Computeranalyse wurde ermittelt, dass 870 BACs des verwendeten Arrays eine oder mehrere *NotI* bzw. ~1700 BACs *BssHII* Erkennungssequenzen enthalten. Die DNA wird mit *EcoRV* und einem der methylierungssensitiven Enzyme geschnitten. Dann werden die überhängenden Enden der geschnittenen unmethylierten Fragmente mit Biotin markiert und anschließend über Bindung an Streptavidin isoliert. Die so isolierten Fragmente werden Cy-5 bzw. Cy-3 markiert und zwei Proben zusammen hybridisiert. Diese Methode arbeitet mit circa 2 µg Ausgangsmaterial. Die Abdeckung des Genoms ist jedoch nicht sehr groß.

Restriktionsmethode MIAMI (Integrated Analysis of Methylation by Isoschizomers)

Die MIAMI [114] Methode verwendet wiederum das methylierungssensitive Enzym *HpaII* und sein nicht-methylierungssensitives Isoschizomer *MspI*. Zunächst wird die DNA mit *HpaII* geschnitten, anschließend werden Adapter angehängt und die DNA amplifiziert. Das PCR-Produkt wird dann mit *MspI* geschnitten, worauf ein zweiter Amplifikationsschritt folgt. Diese Methode reichert also unmethylierte CpG-reiche Fragmente an. Durch Verwendung von *MspI* auch im ersten Schritt ist es möglich SNPs in der Erkennungssequenz zu erkennen. Zur Detektion werden Mikroarrays mit 60-mer Oligonukleotiden verwendet, die sich in Promotorregionen befinden. Da die Methode zwei Amplifikationsschritte mittels PCR beinhaltet, sollten kleine Mengen DNA als Ausgangsmaterial ausreichend sein. Allerdings ist die mehrfache Amplifikation auch eine Quelle für Artefakte.

Restriktions Methode von Yuan et al. – MSNP (SNP chip based methode)

Diese Methode [113] verwendet den 10K SNP-Mikroarray und den dazugehörigen Assay von Affymetrix in einer abgewandelten Form. Die DNA wird mit dem nicht-methylierungssensitiven Enzym *XbaI* fragmentiert und es werden Adapter angehängt. Eine Hälfte der DNA wird anschließend mit dem methylierungssensitiven *HpaII* geschnitten. Die anschließende PCR reichert methylierte Fragmente an. Die andere Hälfte der DNA wird entsprechend des SNP-Assays behandelt und zur Analyse genomischer Veränderungen verwendet. Bei der Hybridisierung auf dem 50K-SNP-Microarray von Affymetrix sind circa 14.800 Fragmente für die Methylierungsanalyse relevant. Die Verwendung des optimierten Affymetrix-Assays und Arrays hat den Vorteil einfacher Handhabung und hoher Reproduzierbarkeit. Nachteilig ist, dass nicht gezielt Regionen, die für die Methylierungsanalyse relevant sind untersucht werden.

Restriktions Methode von Schumacher et al.

Diese Methode [111] besteht aus der getrennten Anreicherung unmethylierter und methylierter Fragmente. Die Anreicherung unmethylierter Fragmente ist ähnlich der Methode von Hatada et al. und beginnt mit einem methylierungssensitiven schneiden genomischer DNA mit einem oder mehreren Enzymen wie *HpaII*, *Hin6I* oder *AcI*, wobei die Autoren zeigen, dass eine Kombination verschiedener Enzyme, die Abdeckung des Genoms erhöht. An die durch Schneiden der unmethylierten Sequenzen entstandenen Enden werden Adapter angehängt. Es wird davon ausgegangen, dass es sich bei den meisten Fragmente mit einer amplifizierbaren Länge um Fragmente aus vollständig unmethylierten Regionen handelt. Um jedoch verbleibende methylierte Fragmente abzureichern wird nach der Ligation noch mit *McrBC* geschnitten. Bei der abschließenden PCR werden also hauptsächlich DNA Fragmente aus unmethylierten Sequenzabschnitten amplifiziert.

Die Proben werden Cy-5 markiert und zusammen mit einer Cy-3 markierten genauso behandelten Kontrollprobe hybridisiert. Die Anreicherung der methylierten Fragmente verläuft ähnlich der DMH Methode, nur dass hier andere Enzyme bzw. Enzymkombinationen verwendet werden. Um SNPs zu detektieren wird eine dritte Hybridisierung vorgeschlagen bei der Phi29 amplifizierte DNA als unmethylierte Probe für die einzelnen Schritte verwendet wird. Diese Methode liefert eine sehr hohe Auflösung verschiedener DNA-Methylierungsmuster, ist aber ebenso aufwendig, da für jede zusätzliche Information auch eine zusätzliche Anreicherung und Hybridisierung notwendig ist.

Methylierte DNA Immunopräzipitation (MeDIP)

Eine der wenigen Methoden, die nicht auf der Verwendung methylierungssensitiver Restriktionsenzyme beruht, ist das von Weber et al. [112] entwickelte MeDIP. Hier wird genomische DNA zunächst fragmentiert und dann methylierte Fragmente durch Immunopräzipitation mit einem Antikörper gegen 5-methyl-Cytosin isoliert. Die Detektion erfolgt zum einen mit einem circa 32.400 überlappende BACs enthaltenden Mikroarray und zum anderen mit einem CpG-Insel Mikroarray. Bei dieser Methode werden die Fragmente in Abhängigkeit von der Dichte methylierter CpGs angereichert. Diese Methode ist daher sehr gut geeignet, genomweit nach hypermethylierten CpG-reichen Sequenzen zu suchen. Es ist jedoch fraglich, ob sie auch bei weniger CpG-dichten Sequenzen anwendbar bleibt und vorteilhaft für die parallele Analyse vieler Proben ist. Weiterhin werden größere Mengen DNA benötigt. Die Autoren arbeiten standardmäßig mit 4 µg für ein MeDIP.

1.2.3. Zusammenfassung – Brustkrebs und DNA-Methylierung

Brustkrebs stellt die häufigste Krebserkrankung bei Frauen dar und auch die häufigste krebsbedingte Todesursache. 90-95% aller Brustkrebserkrankungen sind sporadisch und somit nicht auf einen erblichen Faktor zurückzuführen. Um diesen Krebs zu untersuchen, muss man also nach Veränderungen suchen, die sich während der Lebenszeit der Patientin angesammelt haben. DNA-Methylierung ist eine epigenetische Modifikation, die dieses Kriterium erfüllt und eine wichtige Rolle bei sporadischem Brustkrebs spielt.

In der Krebsforschung gibt es verschiedene Schwerpunkte wie Früherkennung, Diagnose und Klassifizierung des Tumors, Erstellen einer Prognose und die Vorhersage der Wirksamkeit einer Therapie mit Hilfe prädiktiver Marker für die nach neuen molekularen Biomarkern gesucht wird. Die Untersuchung veränderter DNA-Methylierung ist geeignet für alle beschriebenen Fragestellungen spezifische Marker zu finden. Einige DNA-Methylierungsmarkerkandidaten wurden bereits identifiziert, allerdings konnte bis vor kurzem aus Mangel an geeigneten Methoden kaum genomweit nach Unterschieden gesucht werden. Bei der DNA-Methylierungsanalyse kann man grundsätzlich die Analyse des Gesamtanteils methylierter CpGs, die Bestimmung von DNA-Methylierungslevels oder DNA-Methylierungsprofilen unterscheiden. Der Gesamtanteil ist nur von geringer praktischer Bedeutung in der Krebsforschung. Die Analyse von DNA-Methylierungslevels ist besonders für die letztendliche Anwendung z.B. bei der sensitiven Detektion von DNA-Methylierungsmarkern in Körperflüssigkeiten oder der Charakterisierung von Tumorproben aber auch bei der Entwicklung solcher Marker und ihrer genauen Untersuchung von Bedeutung, während DNA-Methylierungsprofile hauptsächlich für die Entdeckung neuer Markerkandidaten genutzt werden aber auch für die Klassifizierung von Tumoren verwendet werden können.

Für die Suche nach neuen DNA-Methylierungsmarkern in Krebsproben, ist eine Methode vorteilhaft, die wenig Probenmaterial benötigt und den Vergleich mehrerer Proben parallel möglich macht. Grundsätzlich besteht jede der gezeigten genomweiten Methoden aus zwei Teilen - zum einen der Anreicherung differentiell methylierter Fragmente und zum anderen der Detektion dieser Fragmente meist mit einem Mikroarray. Methoden, die bei der Anreicherung keinen Amplifikationsschritt enthalten, benötigen große Mengen Ausgangsmaterial und sind daher unvorteilhaft für diese Fragestellung. Ein weiterer wesentlicher Aspekt ist die Abdeckung des Genoms, die durch die entsprechende Methode erzielt werden kann. Die Etablierung und Verbesserung der DMH Methode und die Entwicklung eines dazugehörigen Mikroarrays war daher ein Ziel dieser Arbeit.

1.3. Zielsetzung

Das erste Ziel dieser Arbeit war die Entwicklung und Etablierung einer technisch einfachen Methode zur genomweiten Analyse von DNA-Methylierungsprofilen. Die entwickelte Methode sollte eine quantitative Auswertung der DNA-Methylierung ermöglichen und möglichst geringe Mengen DNA benötigen, da sie auf limitiertem Tumormaterial angewendet werden sollte. Als Grundlage diente die DMH Technologie. Ein weiteres Ziel war die Entwicklung eines für die veränderte Technologie spezifisch angepassten DNA-Mikroarrays und einer entsprechenden Datenauswertung.

Die optimierte DMH-Technologie sollte mit Hilfe eines Modellsystems aus der Brustkrebszelllinie MDA-MB-231 und PBL-DNA hinsichtlich der Reproduzierbarkeit gefundener DNA-Methylierungsunterschiede und ihrer Verifizierbarkeit mittels direkter Bisulfitsequenzierung charakterisiert werden.

Das zweite Ziel dieser Arbeit war es, mit Hilfe der neu entwickelten und charakterisierten Methode neue Methylierungsmarker in Brustkrebs zu identifizieren. Einerseits sollten durch den Vergleich von gesundem Gewebe mit Brusttumoren Tumormarker identifiziert werden, die das Potential haben, in der Früherkennung Anwendung zu finden. Weiterhin sollten Kandidaten für prädiktive Marker für das Ansprechen auf eine Chemotherapie identifiziert werden.

2. Methoden

2.1. Allgemeine molekularbiologische Methoden

2.1.1. DNA Präparation

Die Lyse von gefrorenem Frischgewebe, gefrorenen Zellpellets aus Frischgewebe, Blut oder Zelllinien und die anschließende Extraktion der DNA wurde mit dem QIAamp DNA Mini Kit entsprechend der vom Hersteller empfohlenen Protokolle durchgeführt. Die Elution der DNA erfolgte mit 60 µl auf 55 °C vorgewärmten Puffer AE .

2.1.2. UV Quantifizierung von DNA

Eine UV-spektrophotometrische Bestimmung der DNA-Konzentration und DNA-Qualität wurde bei 260 nm (DNA) und 280 nm (Protein) durchgeführt. UV-spektrophotometrische Untersuchungen wurden entweder mit dem GENios Plattenspektrophotometer (Tecan) oder mit einem BioPhotometer durchgeführt. Resultierend aus molarem Absorptionskoeffizienten (ssDNA bzw. dsDNA) und Schichtdicke ergab sich die Konzentration der DNA in ng/µl durch Multiplikation der OD mit 50 (dsDNA) bzw. 33 (ssDNA). Für nicht konvertierte DNA (doppelsträngig) wurde der Faktor 50 und für Bisulfit-konvertierte DNA (vorwiegend einzelsträngig) der Faktor 33 verwendet. Die Reinheit der DNA hinsichtlich eventuell noch vorhandener Proteine wurde durch den Quotienten der Absorptionen bei 260 nm und 280 nm beschrieben.

2.1.3. Herstellung eines DNA Standards mit 0% DNA-Methylierung

Unmethylierte DNA (im Folgenden bezeichnet als Phi0%) wurde nach dem Prinzip der multiplen Verdrängungsamplifikation [117] erzeugt. Durchgeführt wurde diese genomweite Amplifikation unter Verwendung des GenomiPhi® DNA Amplification Kit. Gemäß der Anweisung des Herstellers wurden 1-10 ng chromosomale DNA in 10 µl Proben-Puffer gelöst, 3 min bei 95°C denaturiert und auf Eis gest ellt. 1 µl phi29 DNA-Polymerase wurde mit 9 µl Enzym-Puffer gelöst und anschließend mit der Probe gemischt. Die Reaktion erfolgte 18 Stunden über Nacht bei 37°C und wurde anschließend 10 Minuten bei 65°C deaktiviert.

2.1.4. Herstellung eines DNA Standards mit 100% DNA-Methylierung

Künstlich methylierte DNA wurde hergestellt, indem genomische DNA, z.B. aus Blut (Promega), bzw. eine durch genomweite Amplifikation hergestellte unmethylierte DNA (Phi0%), durch die Methyltransferase Sss1 methyliert wurde (im Folgenden als Pro100% bzw. Phi100% bezeichnet). Als Donor für die Methylgruppen dient hierbei S-Adenosylmethionin (SAM). Die DNA-Methylierung der DNA fand in einem Volumen von 250 µl folgender Zusammensetzung statt: 10 µg unmethylierte DNA, 1,24 µl SAM, 40 U Sss1 Methylase und 1 x NE-Puffer 2. Dieses Reaktionsgemisch wurde 24 Stunden bei 37 °C im Thermomixer (Eppendorf) inkubiert. Nach 8 Stunden Inkubation wurden weitere 40 U Sss1 Methylase und 1,24 µl SAM zugegeben.

2.2. Herstellung einer Repräsentation methylierter DNA Fragmente (DMH-Amplicon)

Die in dieser Arbeit verwendete Methode zur Analyse von DNA-Methylierungsunterschieden ist eine Optimierung der Differentielle DNA-Methylierungs Hybridisierung (DMH). Diese Methode basiert auf der Anreicherung methylierter Fragmente aus genomischer DNA durch Verwendung methylierungssensitiver Restriktionsenzyme und adapter-vermittelter PCR. Der Prozessverlauf ist in Abbildung 4 dargestellt. Alle Temperaturprogramme wurden mit einem DNA Engine Tetrad® Thermocycler durchgeführt.

2.2.1. Fragmentierung genomischer DNA

Zwischen 500 ng und 1000 ng einer DNA Probe wurden mit 5 units MseI (NEB), Csp6I und Bfal in 30 µl 1x Y⁺/Tango Puffer bei 37°C für 16 h inkubiert. Anschließend wurden die Enzyme für 20 min bei 65°C inaktiviert und die geschnittene DNA mit Hilfe des QiaQuick PCR-Aufreinigungskits entsprechend der Empfehlungen des Herstellers aufgereinigt. Die Elution der DNA erfolgte in 40 µl Wasser. Zur Kontrolle der Komponenten wurde ein Ansatz mit genomischer DNA aus Blut ohne Enzyme mitgeführt. Weiterhin wurde die Aktivität jedes Enzyms einzeln getestet. Dazu wurde λ-DNA für die Kontrolle der Enzyme MseI und Csp6I und PBL DNA (Promega) für Bfal nur mit dem jeweiligen Enzym verdaut. Die Gelelektrophorese dieser Kontrollen muss für die λ-DNA ein enzym-spezifisches Bandenmuster (siehe Abbildung 15 sowie <http://tools.neb.com/NEBcutter2/index.php>) und für die PBL-DNA einen DNA-Schmier zwischen 100bp und 2000bp zeigen. Die Kontrolle ohne Enzyme muss intakte genomische DNA geblieben sein (siehe Abbildung 15).

Zur Kontrolle eventueller DNA-Kontaminationen des gesamten Prozessverlaufs wurde eine Reaktion mit Wasser statt DNA angesetzt und eine Reaktion mit PBL-DNA, bei welcher im

nächsten Schritt keine Ligase eingesetzt wurde. Beide Kontaminationskontrollen wurden bei allen Schritten der DMH-Amplicon Herstellung mitgeführt. Jeweils 3 µl der Kontrollen und der aufgereinigten Proben wurden mittels Gelelektrophorese analysiert. Proben mit DNA zeigten einen Schmier zwischen 100-700 bp, die Kontrolle ohne DNA sollte leer sein (siehe Abbildung 15).

2.2.2. Ligation von Adaptern

34µl der aufgereinigten und fragmentierten DNA wurden in einem Endvolumen von 50 µl 1x Ligase-Puffer mit 400 units T4 DNA Ligase, 1 mM ATP und 10 µM Adapter gemischt. Das Ligationsprodukt wurde nach einer Inkubation bei 16°C für 4 h mit Hilfe von QiaMinElute PCR-Aufreinigungssäulen aufgereinigt und in 30 µl Wasser eluiert. Jeweils 3 µl der aufgereinigten Proben wurden mittels Gelelektrophorese analysiert. Proben mit DNA zeigten einen Schmier zwischen 100-700 bp, die Kontrolle ohne DNA sollte leer sein.

2.2.3. Herstellung des Adapters

Die verwendeten Adapter wurden durch Annealing der beiden Primer H24: 5'-AGGCAACTGTGCTATCCGAGGGAT-3' und H12: 5'-TAATCCCTCGGA-3' hergestellt. Dazu wurden gleiche Volumina der Primer mit einer Konzentration von 200 pmol/µl gemischt und das Gemisch mit folgendem Temperaturprogramm inkubiert: Aufschmelzen für 5 min bei 95°C anschließend schrittweises abkühlen 0,1°C/s bis auf 65°C und dann 134 Zyklen Abkühlen um 0,3°C und halten der Temperatur für 1 min.

2.2.4. DNA-Methylierungs - sensitive Restriktion

Bei der anschließenden DNA-Methylierungs-sensitiven Restriktion wurden 34 µl des Ligationsproduktes in einem Volumen von 48 µl 1x Restriktionspuffer NEB1 zunächst mit jeweils 5 units der Enzyme Bstul, HpaII, HinfI und HpyCH4IV gemischt. Nach einer Inkubation bei 37°C für 2 h und 60°C für 1h wurden erneut 5 units jedes Enzyms zugegeben und für weitere 5 h bei 37°C und 3 h bei 60°C inkubiert. Das Restriktionsprodukt wurde wiederum mit Hilfe des QiaQuick PCR-Aufreinigungskits entsprechend der Empfehlungen des Herstellers aufgereinigt und in 60 µl Wasser eluiert. Zur Kontrolle der Komponenten wurde ein Ansatz mit PBL DNA ohne Enzyme mitgeführt. Weiterhin wurde die Aktivität jedes Enzyms einzeln getestet. Dazu wurde nicht-methylierte Lambda DNA mit allen Enzymen verdaut. Jeweils 8 µl der Kontrollen und der aufgereinigten Proben wurden mittels Gelelektrophorese analysiert. Proben mit DNA zeigten einen Schmier zwischen 100-700 bp, die Kontrolle ohne DNA sollte leer sein (siehe Abbildung 15). Die Komponentenkontrolle sollte sich als intakte genomische DNA zeigen und die λ-DNA sollte ein spezifisches Bandenmuster zeigen (siehe Abbildung 15 und Abbildung 8)

2.2.5. PCR Amplifikation

Die Amplifikation aller ungeschnittenen Fragmente wurde in 100µl 1x PCR-Puffer durchgeführt. Ein PCR-Ansatz enthielt circa 5 ng aufgereinigter geschnittener DNA, 5µl DMSO, 1,9 mM MgSO₄, 175 µM dNTP, 2,5 µM H24 Primer und 5 units DeepVent (exo-) taq-Polymerase. Das PCR-Programm beginnt mit einem initialen Elongations-Schritt von 72°C für 5min. Daran schließen sich 20 Zyklen bestehend aus Denaturierung bei 97°C für 1min und Elongation bei 72°C für 3 min an. Es endet mit einer finalen Elongation von 10 min bei 72°C. Jeweils 3 µl zufällig ausgewählter Proben und aller Kontrollen wurden mittels Gelelektrophorese analysiert. Die Kontrollen ohne Ligase, ohne DNA und die PCR-negativ Kontrolle durften kein Produkt zeigen. Um ausreichend DNA für die anschließende Hybridisierung zu erhalten, wurden 6-12 PCRs benötigt. Jeweils zwei PCRs wurden nacheinander auf eine QiaQuick Aufreinigungssäule aufgetragen und anschließend aufgereinigt. Die Elution erfolgt in einem Volumen von 65 µl EB Puffer. Da für die Hybridisierung 20 µg PCR-Produkt in einem Volumen von 45 µl nötig sind, wurden alle Eluate einer Probe auf einer Microcon Y30 Säule aufkonzentriert und anschließend die DNA-Konzentration mittels UV-Quantifizierung bestimmt.

2.2.6. Fragmentierung und Markierung des PCR-Produktes

Die Fragmentierung und die anschließende Markierung der Amplifikate mit Biotin wurde mit Hilfe des GeneChip® Mapping 10K Xba Assay Kit durchgeführt. Das vom Hersteller empfohlene Protokoll (GeneChip® Mapping Assay Manual, 2003-2004) wurde wie folgt modifiziert. Die Fragmentierung von 20 µg aufgereinigtem DMH-Amplifikat wurde in 1x Fragmentierungs-Puffer mit 0,06 units Fragmentierungsreagenz in einem Gesamtvolumen von 55 µl durchgeführt. Das Reaktionsgemisch wurde für 10 min bei 37°C inkubiert und danach wurde das Enzym bei 95°C für 10 min inaktiviert.

Anschließend wurden 50,6 µl der fragmentierten Probe mit 14 µl TdT Puffer, 2 µl Markierungsreagenz und 3,4 µl TdT versetzt und für 2 h bei 37° inkubiert. Das Enzym wurde wiederum bei 95°C für 15 min inaktiviert.

2.2.7. Hybridisierung, Waschen, Färben und Scanen

Auch die Hybridisierung der fragmentierten und gelabelten Probe wurde mit wenigen Modifizierungen entsprechend des GeneChip® Mapping Assay Manual (2003-2004) durchgeführt. Die Mikroarrays wurden zunächst mit 200 µl Prähybridisierungsmix für 10 min bei 48°C und 60 rpm im Affymetrixhybridisierungsöfen prä-hybridisiert. Währenddessen wurde der Hybridisierungsmix für 10 min bei 95°C im Mastercycler denaturiert, dann für 10 Sekunden auf Eis gekühlt und abschließend 2 min bei 48°C temperiert. 200 µl der denaturierten und temperierten DNA wurden auf einen prähybridisierten Mikroarray

aufgetragen und bei 48°C und 60 rpm für 16-18 h hybridisiert. Waschen, Färben und Scannen der hybridisierten Mikroarrays wurde ohne Modifikation aus dem GeneChip® Mapping Assay Manual (2003-2004) übernommen. Beim Färben handelt es sich um einen Verstärkungsassay in drei Schritten. Im ersten Schritt wird zunächst Streptavidin an das Biotin gekoppelt, dann folgt die Verstärkung durch die Bindung biotinylierter Streptavidinantikörper und die abschließende Färbung mit einem Streptavidin-Phycoerythrin-Konjugat. Die Auswertung der Mikroarray-Hybridisierungen erfolgte mit einem GeneChip® Scanner 3000. Die Intensität des bei 570 nm emittierten Lichts in jeder Probenzelle ist proportional zur Menge der gebundenen Fragmente.

Reagenzien	Prä-Hybridisierung	Hybridisierung
12 x Mes	12,0 µl	11,0 µl
DMSO (100%)	13,0 µl	12,0 µl
Denhardt's Solution (50x)	13,0 µl	12,0 µl
EDTA (0.5M)	3,0 µl	2,8 µl
HSDNA (10mg/µl)	0,0 µl	2,8 µl
Human Cot-1 (1mg/µl)	0,0 µl	2,8 µl
Tween-20 (3%)	1,0 µl	0,9 µl
Control Oligo B2	0,0 µl	1,8 µl
TMACL	140,0 µl	130,0 µl
Water	78,0 µl	0,0 µl
Labeled DNA	0,0 µl	65,0 µl
Total	260,0 µl	241,1 µl

2.2.8. In Silico Analyse des DMH-PCR-Produktes

Die in silico Analyse der beim DMH erzeugten Fragmente und ihrer Eigenschaften (Größe, enthaltene methylierungssensitive Schnittstellen usw.) wurde in Zusammenarbeit mit Tamas Rujan (Epigenomics AG) und Dr. Jörn Lewin (Epigenomics AG) durchgeführt. Als Datenbasis diente die Datenbank Human Genom Ensembl Version NCBI 33. Diese wurde vom Server (www.ensembl.org) im fasta Format heruntergeladen. Die Datei enthält alle verfügbaren contigs des humanen Genoms. Es wurde nicht mit repeat maskierten Sequenzen gearbeitet. Zunächst wurde mit Hilfe der interaktiven Software-Umgebung „R“ [118] ein Programm erstellt, welches alle Fragmente berechnet, die durch Einsatz der ausgewählten nicht-methylierungssensitiven Restriktionsenzyme (*Bfal*, *Csp6I*, *MseI*) entstehen. Anschließend wurden folgende Parameter für diese Fragmente bestimmt, so dass ein anschließendes Suchen, Filtern oder Zählen dieser Parameter möglich wurde: Länge in bp, Anzahl der Schnittstellen jeweils eines der verwendeten methylierungssensitiven Enzyme, Anzahl der CpGs und die Anzahl der bp in Repeatsequenzen. Die Repeatsequenzen wurden mit Hilfe des RepeatMasker (<http://www.repeatmasker.org/>) identifiziert. Zusätzlich wurden die Fragmente im Genom annotiert und bezüglich ihrer Lage wurden ihnen ebenfalls Eigenschaften zugeordnet. Dafür wurden folgende Definitionen verwendet: Nähe zu einem

TSS ist gegeben, wenn in der Umgebung von ± 2000 bp des Fragmentes ein TSS in zu finden ist. Die Überlappung eines Fragmentes mit einem Gen, Exon, Intron oder einer CpG-Insel musste mindestens 50bp betragen, damit das Fragment als überlappend gezählt wurde. Für die Definition einer CpG Insel wurden in dieser Arbeit folgende Kriterien verwendet: Länge > 200 bp, GC-Gehalt > 50%; Quotient von beobachteter zu erwarteter CpG-Dichte > 0,6; nach Gardiner-Garden und Frommer [29]. Diese Daten wurden zum einen für das Oligonukleotid-Mikroarraydesign verwendet und zum anderen für die Analyse der Verteilungen der verschiedenen Parameter aller entstehenden Fragmente oder der auf den Oligonukleotid-Mikroarrays enthaltenen Fragmente.

2.2.9. Mikroarraydesign

Die in dieser Arbeit verwendeten Oligonukleotid-Mikroarrays wurden als kundenspezifische Anfertigung von Affymetrix bezogen. Grundsätzlich sind diese Arrays genauso aufgebaut, wie die in Fachkreisen bekannten GeneChip® Arrays bei denen 25-mer Oligonukleotide, die so genannten Sonden (Probes), mittels Photolithographie direkt auf der Glasoberfläche nebeneinander in den Sonden-Zellen synthetisiert sind. Für jede Ziel-Sequenz (target sequence) befinden sich mehrere Sonden auf dem Array, die ein Sonden-Set (Probeset) bilden. Die Auswahl der Sonden für den Array, wurde durch die Anwendung zur Detektion von Fragmenten des DMH-Amplicons bestimmt. Abweichend vom Stand der Technik, bei dem pro Ziel-Sequenz 11 Sonden mit vollständiger Komplementarität der Basen zum Target (Perfect Match, Detektionssonden) und 11 Sonden mit einem Fehler (Mismatch, fehlerhafte Sonden) verwendet werden, wurden hier ausschließlich PM-Sonden verwendet.

2.2.9.1. Design des Entwicklungs-Mikroarrays

Zunächst wurde ein Mikroarray mit relativ geringer Zahl von Oligonukleotiden designed. Es handelt sich dabei um einen Affymetrix Array auf dem circa 80.000 Oligonukleotide immobilisiert sind. Aus den gemäß Abschnitt 2.2.8 erzeugten Fragmenten wurden alle Fragmente aussortiert, die mehr als 20% Repeats enthielten. Anschließend wurde von Affymetrix ein Design möglicher Sonden nach optimierten Kriterien vorgenommen. Aus diesen Sonden wurden Sondensets aufgrund folgender Kriterien zusammengestellt. Dieser Oligonukleotid-Mikroarray diente der Etablierung und Charakterisierung der gesamten Methode insbesondere aber auch der Analyse folgender Fragestellungen: Ist die Verwendung von ausschließlich Detektionssonden möglich? Welche Anzahl von Detektionssonden pro Sondenset ist notwendig? Wie ist das optimale Größenfenster der analysierten Fragmente? Daher wurden diese Parameter auf dem Mikroarray variiert. Grundsätzlich gibt es zwei Arten Fragmente, die Kontrollfragmente, welche keine methylierungssensitive Schnittstelle enthalten und die Detektionsfragmente mit mindestens einer methylierungssensitiven Schnittstelle. Die Auswahl der Sondensets erfolgte ansonsten

zufällig, d.h. es wurde nicht auf die Nähe von TSS oder CpG-Inseln geachtet. Dadurch sind auf dem Array auch Fragmente, die nicht mit einer CpG Insel überlappen oder in der Nähe eines Genes oder TSS liegen, vorhanden. Dies war erwünscht, um auch diese Sequenzanteile untersuchen zu können. Weitere Details des Designs, insbesondere die Analyse der variierten Parameter, werden in Abschnitt 3.1.3 gezeigt.

2.2.9.2. Design des 50.000 Fragmente Mikroarrays

Auch bei dem 50.000-Fragmente-Mikroarray handelt es sich um einen Affymetrix-Array auf dem rund 500.000 Oligonukleotide immobilisiert werden können. Auch für dieses Design wurden die gemäß Abschnitt 2.2.8 erzeugten Fragmente verwendet und alle Fragmente aussortiert, die mehr als 20% Repeats oder weniger als 6 CpGs enthielten. Anschließend wurde von Affymetrix ein Design möglicher Sonden nach optimierten Kriterien vorgenommen. Aus diesen Sonden wurden Sondensets aufgrund der in Abschnitt 3.1.3 ermittelten Parameter: Fragmentlänge zwischen 126bp-548bp und mindestens 10 Sonden pro Sondenset, ausgewählt. Es wurden wiederum neben den Detektionssondensets auch Kontrollsondensets designed. Die Auswahl der Sondensets bzw. detektierten Fragmente wurde in diesem Fall auch durch die Lage des Fragmentes bestimmt. Es wurden hauptsächlich Fragmente ausgewählt, die mit einer CpG-Insel überlappen oder in der Nähe eines TSS bzw. Genes liegen. Das Ergebnis des Designs ist in Abschnitt 3.1.6 beschrieben.

2.2.10. Datenverarbeitung

Da es sich um einen für die Detektion von DMH-Fragmenten spezifischen Mikroarray handelte, konnte keine Affymetrix Software verwendet werden. Für die Datenverarbeitung haben wir, im besonderen Thomas König, daher mit Hilfe der interaktiven Software-Umgebung R [118] und der freien Software Bioconductor [119] eine spezifische Datenverarbeitung entwickelt. In Abbildung 14 sind die einzelnen Schritte, denen die Rohdaten unterzogen werden, dargestellt. Sowohl bei der log Transformation als auch bei der Quantil-Normalisierung handelt es sich um standardmäßig durchgeführte Normalisierungsmethoden für Affymetrixarrays. Die Basislinienverschiebung ist in Abschnitt 3.1.4 erläutert.

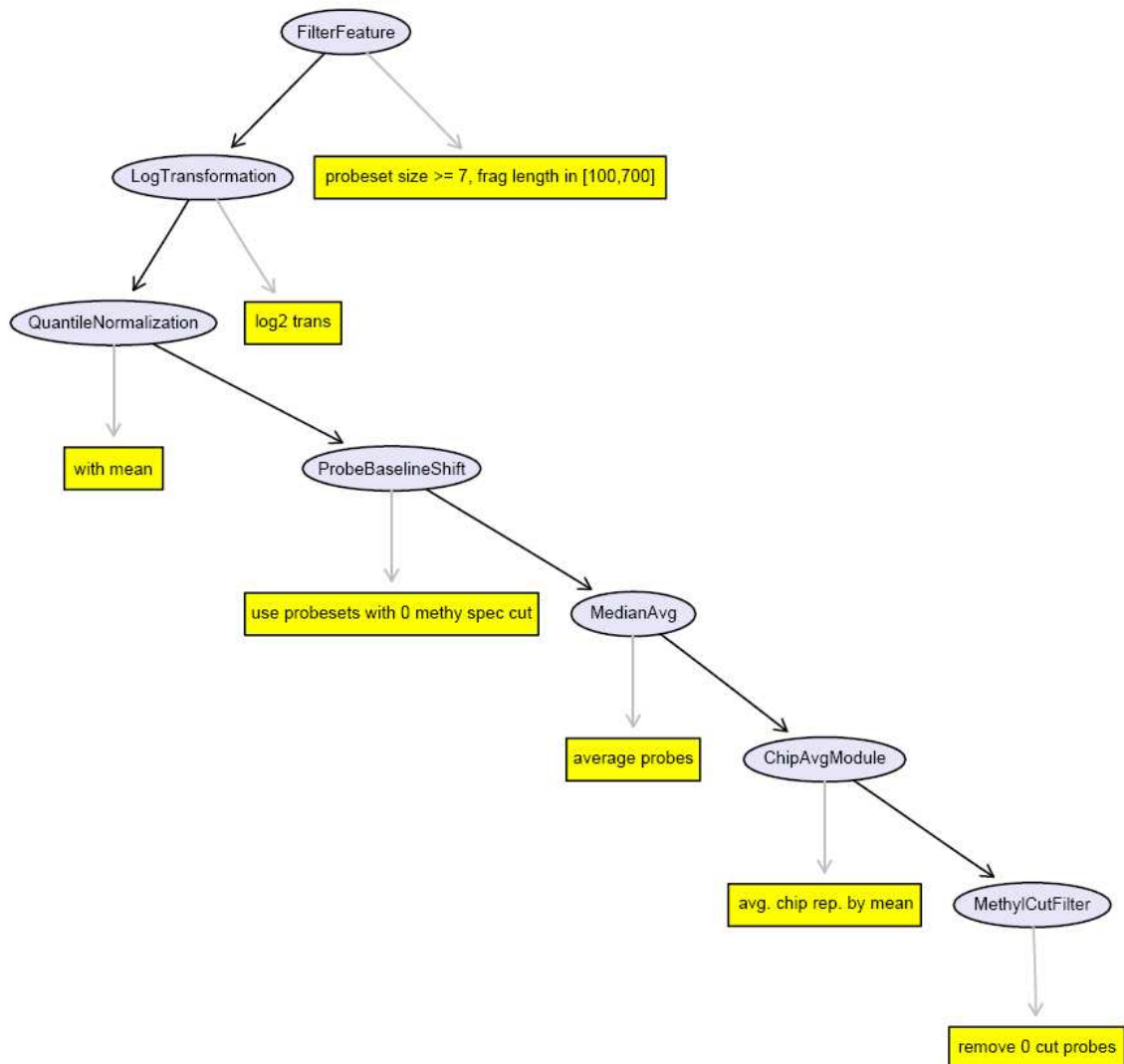


Abbildung 5: Einzelne Schritte der Verarbeitung der Signaldaten eines hybridisierten Mikroarrays.

2.3. Direkte Bisulfitsequenzierung

2.3.1. Herstellung von Bisulfit-DNA

Bei der hier beschriebenen Umwandlung genomischer DNA in bisulfit-konvertierte DNA wurde ein von der Epigenomics AG patentiertes Verfahren (Berlin et al. 2005) verwendet, das auf der erstmals von Frommer beschriebenen Methode (Frommer et al. 1992) basiert.

1 µg DNA wurde in 21 µl Wasser vorgelegt und mit 84 µl Bisulfit-Lösung und 35 µl Dioxan versetzt. Anschließend wurde die DNA nach folgendem Schema inkubiert: 5 min bei 99°C, 25 min bei 60°C, 5 min bei 99°C, 1,25 Stunden bei 60°C, 5 min bei 99°C, 2,55 Stunden bei 60°C. Zur Aufreinigung wurden 140 µl Reaktionsgemisch mit 260 µl Wasser (Fluka) gemischt und auf eine Microcon® YM-30 Aufreinigungssäule aufgetragen und für 15 min mit 14.000 g, bzw. bis keine Flüssigkeit auf der Membran sichtbar war, zentrifugiert. Der Durchlauf wurde verworfen. Im nächsten Schritt wurden 400 µl 0,2 M NaOH auf die Säule aufgetragen und diese wiederum für 12 min bei 14.000 g zentrifugiert und anschließend dreimal gewaschen, indem 400 µl Wasser auf die Säule aufgetragen wurden und diese für 12 min mit 14.000 g zentrifugiert wurde. Der Durchlauf wurde jeweils verworfen. Die aufgereinigte bisulfitierte DNA wurde zum Schluss mit 75 µl 55°C warmen TE-Puffer eluiert, dazu wurde die Säule umgedreht für 5 min mit 1000 g zentrifugiert.

2.3.2. PCR Amplifikation von Bisulfit-DNA

Ein PCR-Ansatz zur Amplifikation einzelner Bisulfit-DNA-Fragmente für die anschließende Sequenzierung enthielt circa 10 ng Template, 800 µM dNTP, jeweils 0,5 µM Primer und 1 unit HotStar Taq Polymerase in 1xPCR Puffer. Das PCR-Programm beginnt mit 15 min bei 95 °C (initiale Aktivierung der Polymerase) und 45 Zyklen bestehend aus Denaturierung für 20 s bei 95 °C, Primer Hybridisierung für 45 s bei 58 °C und Elongation für 30 s bei 72 °C und endet mit einer finalen Elongation von 10 min, 72°C. Das Temperaturprogramm wurde mit einem DNA Engine Tetrad® Thermocycler durchgeführt. 2-5 µl der PCR-Produkte wurden zur Qualitätskontrolle auf ein Gel aufgetragen.

2.3.3. Sequenzierung von Bisulfit-DNA

PCR Produkte wurden mit der Didesoxymethode nach Sanger et al. [120] sequenziert. Für den Verdau der nicht inkorporierten dNTPs und Primer wurden 3,5 µl der PCR Produkte mit 1 µl Wasser und 1 µl ExoSAP-IT einem Gemisch aus Exonuclease I und Shrimp Alkaline Phosphatase versetzt, für 45 min bei 37°C inkubiert und das Enzym zum Schluss für 15 min bei 95°C inaktiviert.

Für die Inkorporation der Fluoreszenz-markierten ddNTPs während der Kettenabbruch-Synthese wurde die BigDye Terminator 3.1 Chemie verwendet. Dazu wurden 5 µl ExoSAP-IT verdautes PCR-Produkt mit 1 µl BigDye Terminator v3.1, 4 µl Sangerpuffer (siehe Anhang), 5 µl 0,08 mM dNTP und 5 µl 2µM Sequenzierprimer gemischt und dieses Gemisch bei folgendem Temperaturprogramm inkubiert: 2 min bei 96 °C, 25 Zyklen mit je 30 s bei 96°C, 15 s bei 55 °C und 4 min bei 60 °C. Beide Temperaturprogramme wurden mit einem DNA Engine Tetrad® Thermocycler durchgeführt.

Die Aufreinigung des gesamten Reaktionsgemisches erfolgte mit Hilfe von DyeEx 96 Aufreinigungsplatten entsprechend den Angaben des Herstellers. Anschließend wurden zu jeder gereinigten Probe noch 10 µl Hi-Di® Formamid gegeben.

Die Kapillargelelektrophorese wurde mit einem 3730 DNA Analyzer durchgeführt. Verwendet wurden dazu 36 cm Kapillaren (3730 DNA Analyzer Capillary Array) und POP-7™ Polymer für 3730/3730xl DNA Analyzers. Dabei wurden folgende Bedingungen gewählt: Ofentemperatur 60 °C, Vorlaufspannung 15 kV, Vorlaufzeit 180 s, Injektionsspannung 1,2 kV, Injektionszeit 15 s, erste Datenaufnahme 250 ms, zweite Datenaufnahme 250 ms, Laufspannung 8,5 kV, Anzahl der Spannungsstufen 30, Spannungsstufenintervall 15 s, Spannungstoleranz 0,6 kV, Stromstabilität 10 µA, Erhöhungsverzögerung 450 s, Datenverzögerung 120 s und Laufzeit 2450 s.

2.3.4. Datenanalyse der Bisulfitsequenzierung

Die Analyse der aus dem ABI-basecaller resultierenden tracefiles wurde mit Hilfe einer zu diesem Zweck von Epigenomics entwickelten Software (ESME) [94] durchgeführt. Diese Software normalisiert zunächst die durch die ABI Auswertung erhaltenen Daten und bringt die Bisulfitsequenz mit der Referenzsequenz, die für das Design verwendet wurde, zur Deckung. Danach werden die durchschnittlichen DNA-Methylierungswerte durch Auswertung des Verhältnisses überlappender C und T Signalpeaks berechnet. Die berechneten Werte werden in eine Gelb-Blau Farbskala übersetzt. Werden mehrere Proben gleichzeitig ausgewertet entsteht eine Matrix in der Reihen den DNA-Methylierungslevel einzelner CpGs darstellen und Spalten Durchschnittswerte einzelner Proben.

3. Ergebnisse

3.1. Optimierung und Etablierung einer Methode für die genomweite Analyse der DNA – Methylierung

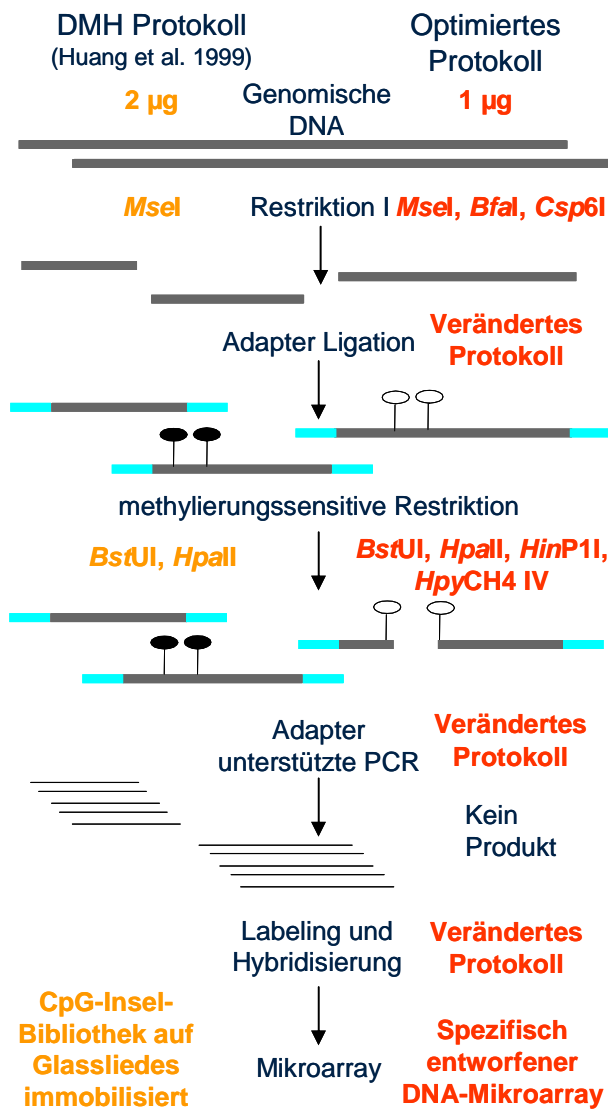


Abbildung 6: Fließschema des DMH:
 1. Genomische DNA wird nicht methylierungssensitiv fragmentiert. 2. An die überhängenden Enden werden Adapter ligiert. 3. Alle unmethylierten DNA-Fragmente werden durch methylierungssensitiven Enzymen geschnitten. 4. PCR Amplifikation methylierte Fragmente bzw. von Fragmenten ohne methylierungssensitive Schnittstelle. 5. Das PCR-Produkt wird gelabelt und Hybridisiert. Schwarze Kreise stehen für methylierte CpGs und weiße Kreise für unmethylierte.

Bei der Suche nach Markern für Früherkennung, Prognose und Vorhersage der Wirksamkeit einer Chemotherapie wird eine genomweite Analyse angestrebt, um möglichst alle möglichen Kandidaten aufzuspüren. Eine der wenigen erprobten potentiell genomweiten Methoden ist die in Abschnitt 1.2.2 beschriebene Methode der Differentiellen Methylierungs Hybridisierung (DMH). Sie basiert auf der Verwendung methylierungssensitiver Enzyme zur Anreicherung methylierter DNA-Fragmente [35]. In Abbildung 6 ist die DMH-Methode dargestellt und gezeigt, welche Veränderungen während der Etablierung und Optimierung vorgenommen wurden. Im Folgenden wird zunächst die Optimierung des Protokolls gezeigt. Ziele hierfür waren: weniger DNA zu benötigen, eine quantitative Auswertung zu ermöglichen und eine möglichst hohe Genomabdeckung zu erreichen. Weiterhin wurde ein Oligonukleotid-Mikroarray spezifisch für DMH-Fragmente entwickelt und etabliert. Abschließend werden Daten der Charakterisierung des gesamt Prozesses anhand eines Modellsystems vorgestellt.

3.1.1. Optimierung und Etablierung des DMH-Protokolls

Im folgenden Abschnitt werden Ergebnisse gezeigt, die bei der Etablierung und Optimierung der DMH-Methode erzielt wurden und einzelne Aspekte des veränderten DMH-Protokolls erläutert, d.h. es werden an dieser Stelle hauptsächlich die Erweiterungen bzw. Verbesserungen der bisherigen Vorgehensweise dargestellt. Die vollständigen Protokolle sind in Abschnitt 2.2 zu finden. Wichtig war, dass das Protokoll möglichst wenig aufwendig ist und eine große Zahl Proben parallel bearbeitet werden kann. Weiterhin ist es wichtig, dass die Gesamtmenge der eingesetzten DNA nicht zu groß ist, da Probenmaterial häufig limitiert ist. Im besten Falle sollte eine Wiederholung der Hybridisierung möglich sein.

3.1.1.1. Nicht-methylierungssensitive Fragmentierung

Für die Fragmentierung der DNA sollten zusätzlich zu *MseI* (TTAA) die zwei Restriktionsenzyme *Bfal* (CTAG) und *Csp6I* (GTAC) verwendet werden, um die Größenverteilung der entstehenden DNA-Fragmente zu verändern (siehe Abschnitt 3.1.2). Um die drei Enzyme gleichzeitig einsetzen zu können, wurde zunächst PBL DNA mit jeweils 5 units der einzelnen Enzyme in NEB2 Puffer und in Y⁺/Tango Puffer für 2h inkubiert.

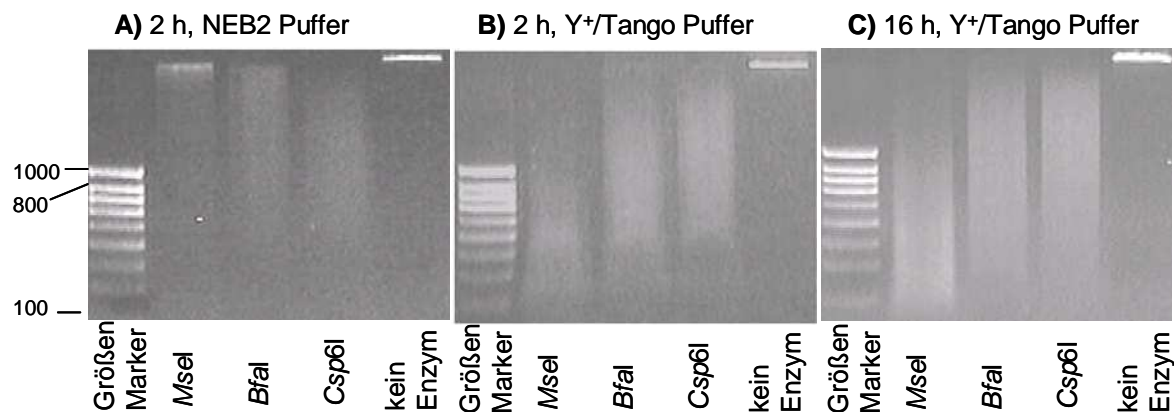


Abbildung 7: Bilder der Gelelektrophorese von durch Restriktionsenzyme fragmentierter genomischer DNA aus Blut: Es wurden in jedem Fall 200 ng DNA auf ein 1,4% Agarosegel aufgetragen. Als Größenmarker dient die SmartLadder. Es wurden drei verschiedene Bedingungen A), B) und C) getestet. In jedem Fall wurde eine Kontrolle mitgeführt bei der kein Enzym eingesetzt wurde und somit hochmolekulare DNA als Ergebnis erwartet wurde.

Das Ergebnis ist in Abbildung 7 dargestellt. Abbildung A zeigt die Ergebnisse bei Verwendung von NEB2 Puffer für 2 h. Besonders im Fall von *MseI*, aber auch für *Bfal*, sind Reste intakter genomischer DNA als hochmolekulare Bande zu erkennen. Bei der Verwendung von Y⁺/Tango Puffer für 2h sind kaum noch Reste unfragmentierter DNA vorhanden, daher wurde dieser Puffer verwendet und bei einer Inkubationszeit von 16h getestet. In Abbildung C ist das Ergebnis gezeigt. Jede DNA ist vollständig fragmentiert und ergibt einen Schmier zwischen 100-1000bp für *MseI* und 100-20.000bp für *Bfal* und *Csp6I*. Der Unterschied in den Größenverteilungen ergibt sich aufgrund der unterschiedlichen Häufigkeit der Erkennungssequenzen der verschiedenen Enzyme.

3.1.1.2. *Ligation*

Für die Ligation wurden von der Gruppe Huang's verschiedene Protokolle publiziert. In Huang et al. [108] wird T4 DNA Ligase für 16h bei 16°C verwendet. Die Ligation wurde mit auf vier Stunden verkürzter Inkubationszeit unter den in Abschnitt 2.2.2 beschriebenen Bedingungen durchgeführt.

3.1.1.3. *Methylierungssensitive Restriktion*

Neben den bereits verwendeten Enzymen *HpaII* und *BstUI* wurden zusätzlich die Enzyme *HpyCH4IV* und *HinP1I* eingeführt, um die Genomabdeckung zu erhöhen (siehe Abschnitt 3.1.2). Die Schnittstellen dieser Enzyme sind wesentlich seltener als die für die nicht-methylierungssensitive Fragmentierung verwendeten. Eine Analyse der Effizienz dieser Enzyme zur Etablierung der Reaktion mittels Gelelektrophorese ist daher nicht mit humaner DNA möglich und es wurde unmethylierte λ -DNA verwendet.

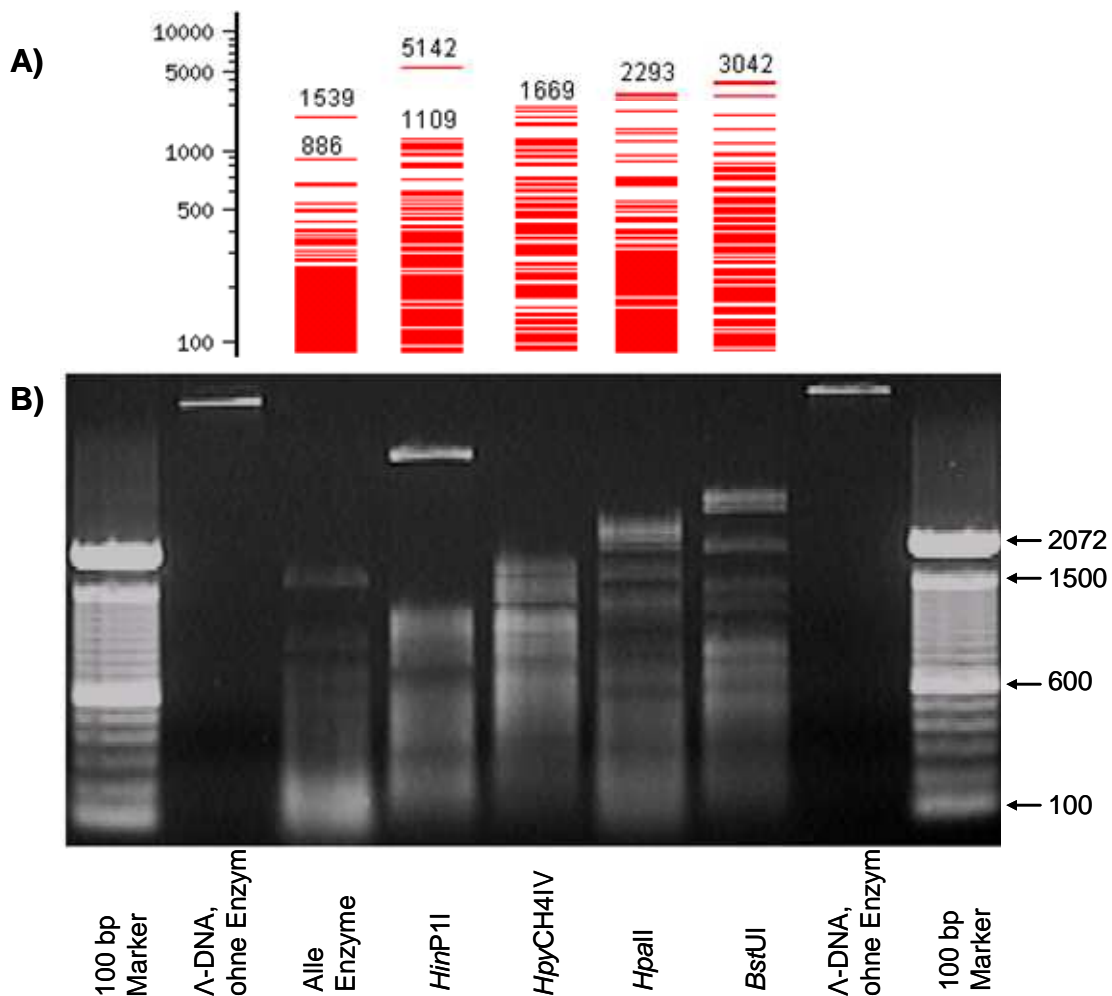


Abbildung 8: **A)** Theoretisches Gelbild der Restriktionen von λ -DNA mit den verschiedenen Enzymen bzw. mit allen vier methylierungssensitiven Enzymen, erstellt mit NEBcutter V2.0 (<http://tools.neb.com/NEBcutter2/index.php>) für ein 1% Agarosegel und eine Lauflänge von 180mm. **B)** 0,8% Agarosegel der Fragmentierungen von λ -DNA mit den einzelnen Enzymen. Es wurden jeweils 200 ng fragmentierter λ -DNA aufgetragen.

Alle vier Enzyme können laut Hersteller in NEB1 Puffer verwendet werden. Um die vollständige Restriktion zu untersuchen, wurde 1µg λ-DNA mit insgesamt 10 units jedes Enzyms hydrolysiert. Die Enzyme wurden in zwei Schritten nach dem in Abschnitt 2.2.4 beschriebenen Schema zugefügt und inkubiert. Das Ergebnis dieser methylierungssensitiven Fragmentierung von λ-DNA ist in Abbildung 8 im Vergleich zu den theoretischen Gelbildern jedes entsprechenden Enzyms dargestellt. Theoretische Gelbilder werden durch eine *in silico* Analyse der λ-DNA Sequenz erstellt. Hierzu wurde das online zugängliche Programm NEBcutter V2.0 (<http://tools.neb.com/NEBcutter2/index.php>) verwendet. Jedes Enzym erzeugt ein anderes spezifisches Bandenmuster, das sich schon anhand der größten Bande identifizieren lässt. Vergleicht man die Bandenmuster der hydrolysierten DNA mit denen aus Reaktionsansätzen denen kein Enzym zugesetzt wurde, d.h. die DNA unfragmentiert vorliegt, ist klar zu erkennen, dass keine Reste unfragmentierter DNA in den Reaktionen mit Enzym zu sehen sind. Weiterhin entspricht in allen Fällen die größte Bande auch der größten Bande des theoretischen Gelbildes. Daher kann von einer vollständigen Hydrolyse der DNA ausgegangen werden. Die Verwendung von λ-DNA wurde aufgrund dieser Ergebnisse auch als Positivkontrolle des methylierungssensitiven Verdaus eingeführt.

3.1.1.4. DMH-Amplicon PCR

In Tabelle 3 ist die Zusammensetzung der optimierten PCR im Vergleich zur ursprünglich publizierten dargestellt. Da es sich bei der PCR um eine Art multiplex-PCR handelt, wurden Primer-, MgSO₄-, und dNTP-Konzentrationen getestet, die sich bei der Epigenomics AG als vorteilhaft erwiesen haben (persönliche Kommunikation Dr. Jürgen Distler). Abbildung 9 zeigt das Ergebnis eines Optimierungsexperimentes für die Gesamtmenge des DNA-Templates.

Tabelle 3: Vergleich der PCR-Zusammensetzung der PCR wie sie in Huang et al. beschrieben ist und der hier beschriebenen optimierten Methode

	Protokoll Huang et al. (1999)	Optimiertes Protokoll
Ausgangsmenge genomischer DNA	2000 ng	1000 ng
PCR		
Volumen	100 µl	100 µl
DMSO	5 %	5 %
MgSO ₄	2 mM	4 mM
Primer	0,4 µM	2,5 µM
dNTPs	200 µM	175 µM
DeepVent exo- DNA Polymerase	2 U	1 U
DNA pro PCR*	400 ng	5 ng
Ausbeute der PCR	20** µg	4-8 µg
maximaler Amplifikationsfaktor	50	1600
maximale Anzahl durchführbarer PCRs*	3	60

* bei 60% Ausbeute nach Fragmentierung; Ligation und methylierungssensitiver Fragmentierung

** geschätzte Ausbeute der PCR

Ergebnisse

Es wurde eine Reihe verschiedener Gesamtmengen von PBL DNA nach der Ligation eingesetzt, die zwischen 0,5 ng und 50 ng lagen, wie aus der Abbildung zu entnehmen ist. Auf höhere Gesamtmengen wurde verzichtet, da die Reduktion benötigter DNA, die Erhöhung der Zahl möglicher Wiederholungen der PCR und die Vergrößerung des Amplifikationsfaktors Ziele der Optimierung waren.

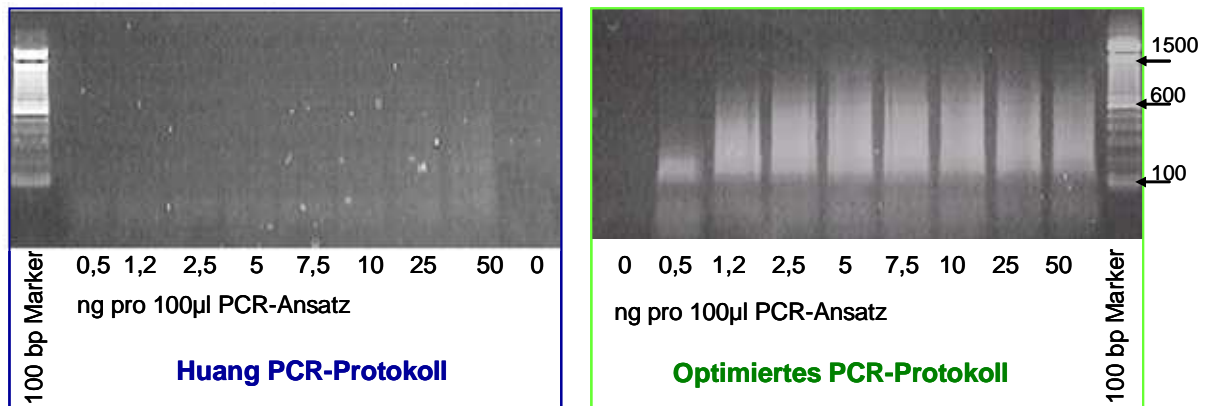


Abbildung 9: 1,4 % Agarosegel von Produkten einer PCR nach dem Huang-Protokoll und nach dem optimierten Protokoll. In die PCR wurden verschiedene Gesamtmengen an Templat-DNA eingesetzt, wie sie an den einzelnen Bahnen vermerkt sind. Es wurden jeweils 10 μ eines 100 μ l PCR-Ansatzes aufgetragen.

Zunächst ist deutlich zu erkennen, dass unter den verwendeten optimierten PCR Bedingungen in jedem Fall ein PCR-Produkt entstanden ist. Die Negativkontrolle zeigt kein PCR-Produkt. Die PCR-Produkte sind intensive DNA-Schmiere in einem Größenbereich von circa 100 bp – 600 bp. Sowohl bei den geringsten als auch bei den höchsten Ausgangsmengen ist eine Abnahme der Intensität und Größe des Schmiere zu erkennen. Als optimale DNA-Menge wurden 5 ng ausgewählt und in allen weiteren Experimenten verwendet.

Bei der Verwendung der PCR-Zusammensetzung, wie sie in Huang et al. [108] angegeben wurde, wurde in diesem Fall kaum PCR-Produkt amplifiziert. Lediglich für sehr hohe DNA-Mengen ist ein sehr schwacher Schmier zuerkennen. Es ist daher vorstellbar, dass bei Einsatz der im Protokoll angegebenen Menge von 400 ng (siehe Tabelle 3) eine gute Amplifikation zu verzeichnen ist. Die Reduktion der Templatmenge hat verschiedene Auswirkungen, die in Tabelle 3 aufgelistet sind. Zunächst ist es möglich, die zur Herstellung eines DMH-Amplicons nötige Ausgangsmenge der DNA von 2000 ng auf 1000 ng zu verringern. Weiterhin steigt der Amplifikationsfaktor der DNA-Menge von 50 auf 1600. Die Größe des Amplifikationsfaktors ist wichtig für die Quantifizierbarkeit, da dadurch der Unterschied zwischen methylierten und nicht methylierten Fragmenten erzeugt wird. Des Weiteren besteht die Möglichkeit, nicht nur 3 PCRs sondern 60 PCRs mit einem Amplicon herzustellen, wodurch Wiederholungen der PCR und nachfolgender Schritte möglich werden.

3.1.1.5. **Fragmentierung, Markierung und Hybridisierung**

Zuletzt wurden Fragmentierung, Markierung mit Biotin und Hybridisierung der PCR-Produkte angepasst. Markierung und Fragmentierung wurden mit dem GeneChip® Mapping 10K Xba Assay Kit durchgeführt und die Hybridisierung erfolgte nach den dort beschriebenen Protokollen. Da das DMH-PCR-Produkt andere Fragmentgrößen aufweist als dieser Kit voraussetzt, wurde eine Anpassung der DNase Menge und Inkubationszeit notwendig.

Nach der Fragmentierung sollten die Fragmentgrößen zwischen 50 bp - 200 bp liegen. Für die Ermittlung der DNase Menge und Inkubationszeit wurden DMH-PCR-Produkte von PBL DNA verwendet. Zunächst wurde das Protokoll des Kits mit 0,24 units DNase und 30min Inkubation bei 37°C getestet (siehe Abbildung 10: A). Wie in der Abbildung zu erkennen ist, wurde die DNA wesentlich stärker fragmentiert als erwünscht, da alle Fragmente deutlich kleiner als 100 bp sind. Weiterhin ist auch anzunehmen, dass ein großer Teil des PCR-Produktes vollständig hydrolysiert wurde, da der DNA-Schmier eine geringe Intensität hat obwohl circa 1000 ng des PCR-Produktes aufgetragen wurden. Daher wurde im nächsten Versuch die DNase-Menge halbiert und Inkubationszeit verringert (siehe Abbildung 10: B). Wie jedoch aus Abbildung 10 B hervorgeht, wurde auch in diesem Fall das PCR-Produkt zu stark fragmentiert. Daher wurde die DNase-Menge auf 0,06 units reduziert und Inkubationszeiten von 20min und 10min getestet. Das Ergebnis ist in Abbildung 10 C) zu sehen. Für beide Inkubationszeiten ist die Größenverteilung zwischen 50-200 bp und bei der Inkubationszeit von 10 min ist die Intensität höher, d.h. es kann von einer besseren Ausbeute ausgegangen werden. Dieses Verhalten wurde in weiteren Experimenten bestätigt, daher wurden diese Bedingungen für die weiteren Experimente verwendet.

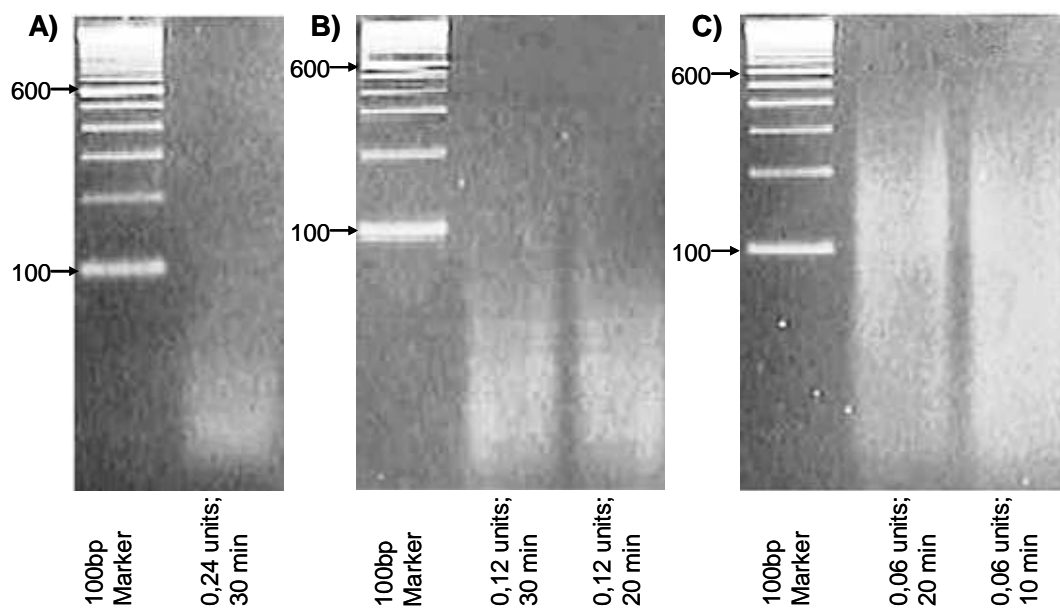


Abbildung 10: 2% Agarosegele von Fragmentierung von DMH-PCR-Produkten mit Hilfe des GeneChip® Mapping 10K Xba Assay Kit. Es wurden jeweils 3 µl des Fragmentierungsansatzes aufgetragen. Die DNase Menge und Inkubationszeit bei 37°C wurden wie angegeben variiert.

3.1.2. Analyse der Komplexität, Größenverteilung und Anzahl informativer Fragmente

Im vorangegangenen Abschnitt wurde die Einführung zusätzlicher Enzyme in die nicht-methylierungssensitive Fragmentierung der DNA und in die methylierungssensitive Restriktion beschrieben. Diese hatte das Ziel, die Größenverteilung und Komplexität des DMH-PCR-Produktes zu verändern und die Genomabdeckung bzw. die Zahl informativer Fragmente (Fragmente mit mindestens einer methylierungssensitiven Schnittstelle) zu erhöhen. Im folgenden Abschnitt werden die Ergebnisse der *in silico* Analyse dieser Parameter gezeigt und erläutert. Die Qualität einer Hybridisierung auf Mikroarrays ist durch die Komplexität der DNA beeinträchtigt. Die Komplexität ist die Anzahl einzelner Basenpaare einer hybridisierten DNA. Ist sie zu hoch, steigt die Möglichkeit unspezifischer Hybridisierungssignale. Weiterhin steigt die benötigte DNA-Gesamtmenge, um ein ausreichendes Signal zu erzeugen, da einzelne DNA-Fragmente umso verdünnter vorliegen je mehr andere Fragmente zur DNA-Gesamtmenge beitragen [121]. Die Komplexität des DMH-Amplicons wird durch die nicht-methylierungssensitive Restriktion reduziert, da hierbei entstehende Fragmente, die kleiner als 100 bp sind, bei der anschließenden Aufreinigung entfernt werden (siehe Produktspezifikation der verwendeten QiaQuick-Aufreinigungssäulen). Die Anzahl informativer Fragmente hängt von den verwendeten methylierungssensitiven Restriktions-enzymen ab und beeinflusst ebenfalls die Komplexität der zu hybridisierenden DNA.

Tabelle 4: Vergleich der Anzahl erzeugter Fragmente bzw. Basen, die nach dem Protokoll von Huang et al. und nach dem optimierten Protokoll entstehen.

Protokoll	Fragmentgröße	Anzahl Fragmente		Anzahl Basen	
		gesamt	informativ*	gesamt	informativ*
DMH-Protokoll nach Huang et al.	≤ 100	1,1E+07	7,8E+04	4,5E+08	5,2E+06
	> 100	8,4E+06	1,3E+06	2,4E+09	7,3E+08
optimiertes DMH	≤ 100	2,2E+07	1,0E+06	9,0E+08	6,0E+07
	> 100	9,8E+06	2,8E+06	2,0E+09	7,8E+08

* Fragmente enthalten mindestens eine methylierungssensitive Schnittstelle

Tabelle 4 zeigt die Ergebnisse der *in silico* Analyse der Anzahl Fragmente bzw. Basenpaare, die durch die ursprüngliche (*MseI*) und die optimierte Enzymkombination nach dem nicht-methylierungssensitiven Fragmentierungsschritt (*MseI*, *BfaI*, *Csp6I*) entstehen. Die Zahlen sind jeweils aufgeteilt in Fragmente die größer oder kleiner als 100bp sind. Des Weiteren ist aufgelistet, wie viele Fragmente informative Fragmente darstellen, d.h. methylierungssensitive Schnittstelle enthalten. Insgesamt entstehen bei der ursprünglichen Methode circa $1,9 \times 10^7$ Fragmente von denen $8,4 \times 10^6$ länger als 100bp sind. Mit der optimierten Methode werden circa $3,2 \times 10^7$ Fragmente erzeugt, davon sind $9,8 \times 10^6$ länger als 100bp.

Um die Auswirkung auf die Komplexität zu beurteilen, muss die Anzahl amplifizierbarer Basen betrachtet werden. Im Fall der ursprünglichen Methode sind das $2,4 \times 10^9$ Basen und der optimierten Methode $2,0 \times 10^9$ Basen. Es können also circa 400 Millionen mehr Basen abgetrennt werden, allerdings ist die Zahl immer noch im Bereich von 10^9 Basen, was deutlich über den von Affymetrix empfohlenen 4×10^7 Basen [121] liegt. Im Folgenden (siehe Abschnitt 3.1.5) kann gezeigt werden, dass trotzdem stabile Signale erzeugt werden und sogar eine quantitative Auswertung möglich ist. Auf eine weitere Optimierung der Komplexität wurde daher verzichtet. Betrachtet man die Zahl informativer Fragmente, zeigt sich, dass durch die optimierte Enzymkombination mehr als doppelt so viele informative Fragmente entstehen. Die Anzahl der Basen in den informativen Fragmenten hat sich jedoch nicht verdoppelt sondern ist von $7,3 \times 10^8$ auf $7,8 \times 10^8$ gestiegen. Dies bedeutet, dass viele der Fragmente der ursprünglichen Methode weiter fragmentiert wurden und dass auch bisher nicht analysierte Sequenzabschnitte erschlossen wurden.

Betrachtet man die Genomabdeckung der beiden Methoden im Vergleich, ergibt sich folgendes Bild. Die durch die optimierte Methode erzeugten Fragmente überlappen mit mindestens 50bp mit 99,7 % aller CpG – Inseln und in der Nähe (± 2000 bp) von 99,2% aller TSS wird mindestens ein Fragment gefunden. Die ursprüngliche Methode zeigt auch eine hohe Abdeckung dieser genomischen Regionen. Es überlappen Fragmente mit 99,6 % aller CpG – Inseln und in der Nähe von 95,0 % aller TSS liegt ein Fragment. Andererseits zeigt sich auch, dass durchaus 4 % mehr TSS abgedeckt sind, d.h. die Methode erzeugt neue Fragmente in Regionen, die vorher nicht analysiert werden konnten.

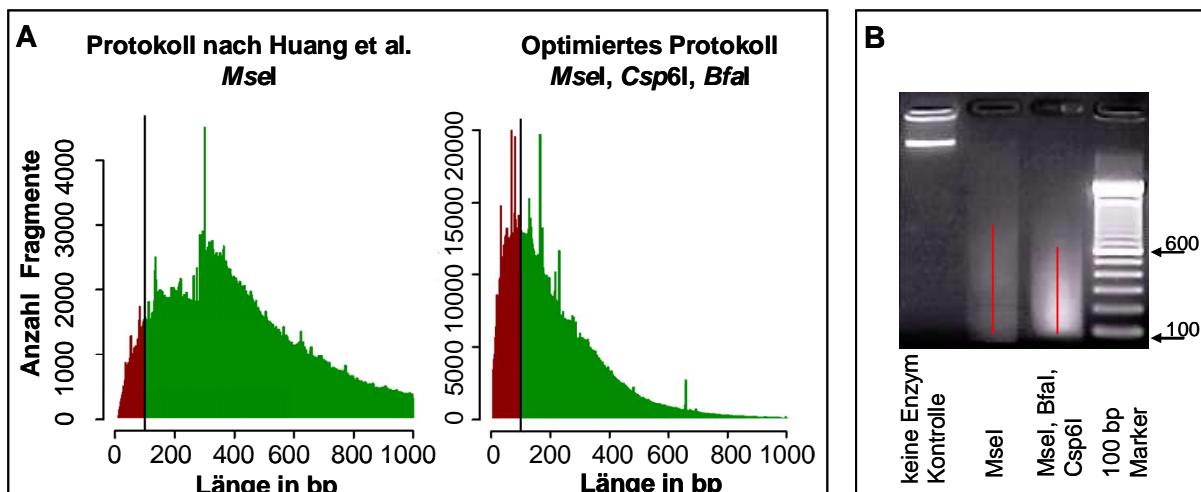


Abbildung 11: Analyse der Fragmentlängenverteilung. **A)** Histogramm der Zahl informativer Fragmente pro Basenpaar, die für das ursprüngliche und das optimierte Protokoll durch *in silico* Fragmentierung des humanen Genoms ermittelt wurden. Bei einer Fragmentlänge von 100bp sind schwarz Grenzen eingezeichnet. Alle Fragmente kleiner 100bp werden durch Aufreinigung abgetrennt und Fragmente größer 100bp sind analysierbar. Die Fragmentgrößenverteilung des optimierten Protokolls zeigt deutlich, dass mehr und kürzere Fragmente entstehen. **B)** 1,4% Agarosegels von PBL DNA, die entsprechend der jeweiligen Protokolle fragmentiert wurde. Die roten Balken markieren die Längebereiche in denen ein einheitlich intensiver DNA-Schmier zu erkennen ist. Auch hier zeigt sich, dass bei der optimierten Methode nur noch bis circa 700bp Fragmentlänge DNA sichtbar ist.

In Abbildung 11 ist die Größenverteilung der entstehenden Fragmente mit methylierungssensitiver Schnittstelle, zum einen als Histogramm der *in silico* erzeugten Fragmente und zum anderen als Gelelektrophorese von PBL-DNA nach der nicht-methylierungssensitiven Fragmentierung, dargestellt. Wie zu erwarten, sind bei der optimierten Methode die Fragmente kleiner, d.h. es werden nicht nur mehr Fragmente abgetrennt sondern es kommt auch zu einer Verkürzung der analysierbaren Fragmente. Dies ist vorteilhaft, da die DMH-Methode bei mehreren Schnittstellen in einem Fragment Komethylierung der Schnittstellen erfordert, damit dieses Fragment angereichert wird. Mit zunehmender Länge steigt die Wahrscheinlichkeit, dass bei einem methylierten Abschnitt eine der im Fragment vorhandenen Schnittstellen nicht methyliert ist und somit DNA-Methylierung nicht gefunden werden kann. Der in dieser Arbeit verwendete Mikroarray enthält z.B. circa 90% Fragmente mit mehr als einer methylierungssensitiven Schnittstelle, d.h. Komethylierung wird erforderlich.

3.1.3. Design des Entwicklungs-Mikroarrays

Der Entwicklungs-Mikroarray diente zuerst der Optimierung des Arraylayouts und der Etablierung einer angepassten Datenauswertung. Dafür wurden verschiedene Layoutparameter variiert und Kontrollsonden eingeführt. Das Design beruht auf dem in den Abschnitten 2.2.8 und 2.2.9 beschriebenen Vorgehen. Optimiert werden sollten die Anzahl der Sonden pro Sondenset und die Länge der Fragmente. In Tabelle 5 ist aufgelistet, wie viele Fragmente und Sonden designed wurden, wie viele davon Kontroll- oder Detektionsfragmente sind und wie viele die nachfolgend bestimmten Designkriterien erfüllen.

Tabelle 5: Layout des Entwicklungsmikroarrays: Anzahl von Fragmenten und Sonden für verschiedene Fragmentarten

	Anzahl Fragmente	Anzahl Sonden
Alle Fragmente (0 - 20 Sonden pro Sondenset; 73bp - 996bp Fragmentlänge)	9044	70502
Fragmente ohne methylierungssensitive Schnittstelle	1135	4045
Kontrollfragmente: $\geq 100\text{bp}$; $\leq 600\text{bp}$; ≥ 7 Sonden	139	1377
Fragmente mit methylierungssensitiver Schnittstelle	7909	66457
Detektionsfragmente: $\geq 100\text{bp}$; $\leq 600\text{bp}$; ≥ 7 Sonden	3780	45368

Die Detektionsfragmente enthalten methylierungssensitive Schnittstellen, d.h. sie können zur Methylierungsanalyse verwendet werden. Kontrollfragmente enthalten keine methylierungssensitiven Schnittstellen und sollten daher immer ein Signal zeigen, welches zur Normalisierung der Mikroarrays verwendet wird (siehe Abschnitt 3.1.4).

Für die Analyse der Fragmentlängen und Anzahl der Sonden pro Sondenset wurden die Daten des in Abschnitt 3.1.5 beschriebenen Vergleichs von PBL-DNA und Brustkrebszelllinien-DNA verwendet. In Abbildung 12 A) ist dargestellt, wie das optimale Fragmentlängenfenster abgeschätzt wurde. Auf dem Mikroarray befinden sich Sonden für circa 9000

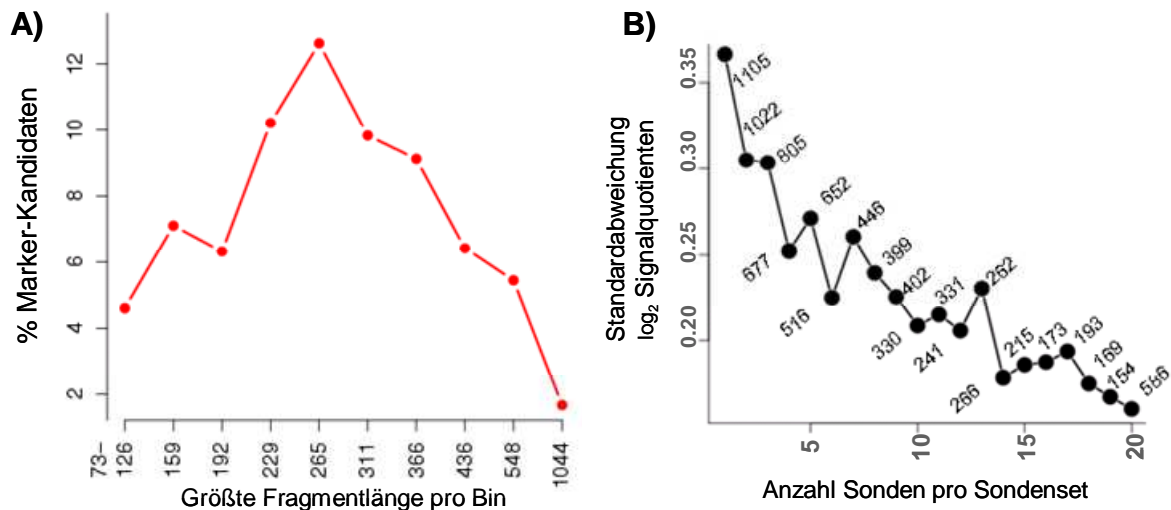


Abbildung 12: Bestimmung des Fragmentlängenfensters und der Anzahl von Sonden pro Sondenset. Es wurden Daten des Vergleichs von PBL-DNA und Brustkrebszelllinien DNA verwendet. Für jede DNA wurden vier Mikroarrays hybridisiert. **A)** Die Fragmente sind in Gruppen (Bins) mit jeweils 900 Fragmenten aufgeteilt. Für jede Bin wurde der Prozentsatz von Marker-Kandidaten in Abhängigkeit von der Fragmentlänge aufgetragen. Marker-Kandidaten sind Fragmente, welche einen logarithmierten Signalquotienten von Brustkrebszelllinie zu PBL-DNA von 0,5 aufwiesen. **B)** Die Standardabweichung der logarithmierten Signalquotienten von Brustkrebszelllinie zu PBL-DNA nimmt mit steigender Anzahl von Sonden pro Sondenset ab.

Fragmente, die in 10 Gruppen (als Bins bezeichnet) mit 900 Fragmenten eingeteilt wurden. Jeder Gruppe wurde ihre durchschnittliche Fragmentlänge zugeordnet. In Abbildung 12 A) wird der Anteil der Fragmente, die eine unterschiedliche Methylierung zwischen den analysierten Proben zeigen, gegen die Fragmentlänge des Bins aufgetragen. Für sehr kleine und sehr große Fragmente ist der Prozentsatz geringer. Zu beachten ist jedoch, dass dies auch mit der Anzahl von Detektionssonden oder der abnehmenden Wahrscheinlichkeit von Komethylierung zusammenhängen kann. Daher wurden von uns für weitere Analysen mit dem Entwicklungs-Mikroarray Fragmente zwischen 100bp bis 600bp verwendet. Für das Design des genomweiten Mikroarrays wurden Fragmentlängen von 126bp-548bp gewählt. Die Abhängigkeit der Varianz einer Hybridisierung von der Anzahl von Sonden pro Sondenset ist in Abbildung 12 B) dargestellt. Mit steigender Anzahl der Sonden nimmt die Standardabweichung der log₂-Signalquotienten und somit die Varianz der Hybridisierungssignale ab. Ab circa 10 Sonden pro Sondenset ist die Abnahme jedoch nur noch gering. Für das Design des genomweiten Mikroarrays wurde daher ≥ 10 Sonden pro Sondenset als optimal angenommen. Um möglichst viele Fragmente auch mit dem Entwicklungsarray noch analysieren zu können, wurden für den weiteren Einsatz des Entwicklungsarrays nur Fragmente mit mindestens 7 Sonden pro Sondenset verwendet. Aufgrund der in Abschnitt 3.1.5 erläuterten Ergebnisse (gute Reproduzierbarkeit, Möglichkeit einer quantitativen Auswertung) kann davon ausgegangen werden, dass die Verwendung von ausschließlich Detektionssonden möglich ist. Somit wurde die Zahl nötiger Sonden pro Fragment von ursprünglich 22 Sonden (11 Detektionssonden und 11 fehlerhaften Sonden) auf circa die Hälfte reduziert.

Wie in Tabelle 5 gezeigt, sind auf dem Entwicklungsmikroarray insgesamt Sondensets für 3780 Fragmente vorhanden, die die beschriebenen Kriterien erfüllen. Aufgrund der in den Abschnitten 2.2.8 und 2.2.9 beschriebenen Analysen ergibt sich, dass 1640 (43%) davon mit einer CpG-Insel überlappen, 2029 (54%) in der Nähe eines TSS liegen und 2555 (68%) mit einem Gen überlappen.

3.1.4. Entwicklung der Datenverarbeitung

Bei der Nutzung von Affymetrixarrays zur Detektion von DMH-Fragmenten kann die Affymetrixsoftware nicht verwendet werden. Daher wurde mit Hilfe der freien

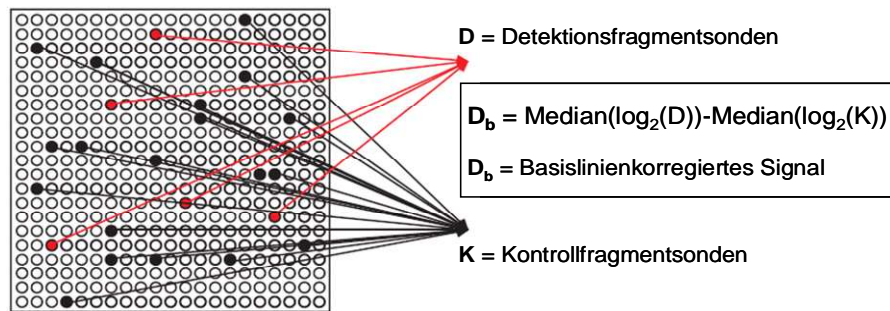


Abbildung 13: Die Basislinienverschiebung

Software Bioconductor [119] eine für die DMH-Methode angepasste Datenanalyse in Zusammenarbeit mit Thomas König (Epigenomics AG) entwickelt.

Die DMH-Methode hat eine Besonderheit, die eine spezielle Normalisierung notwendig macht. Da die Anzahl der amplifizierten und hybridisierten Fragmente aufgrund unterschiedlicher DNA-Methylierung stark variieren kann, liegt ein einzelnes Fragment bei konstanter Gesamtmenge hybridisierter DNA in unterschiedlichen Konzentrationen vor. Das bedeutet, dass jede Probe eine von ihrer DNA-Methylierung abhängige Komplexität hat und daher das gleiche Signal einer Probe mit hoher Komplexität einer höheren DNA-Methylierung als bei einer Probe mit niedriger Komplexität entspricht.

Um die Signale von zwei Proben richtig bewerten und quantifizieren zu können, wurde die Basislinienverschiebung eingeführt. Sie verwendet die Signale der Kontrollfragmente ohne methylierungssensitive Schnittstellen. Die Datenverarbeitung ist in Abschnitt 2.2.10 erläutert. Sie beginnt mit der Auswahl der Fragmente mit einer Länge zwischen 100 bp und 700 bp und mindestens 7 Sonden pro Sondenset. Anschließend werden die Signalintensitäten \log_2 -transformiert und Quantil normalisiert. Darauf folgt die Basislinienverschiebung (Abbildung 13). Dafür wird von den Kontrollsondensets der Median gebildet und von den Detektionssonden-Signalen des jeweiligen Arrays abgezogen. Ist die Komplexität der hybridisierten Probe hoch, wird die mittlere Signalintensität der Kontrollen kleiner sein als bei Proben mit niedriger Komplexität, d.h. bei einer Probe mit hoher Komplexität wird weniger von den Signalen abgezogen und sie werden somit gegenüber den Signalen einer Probe mit geringer Komplexität aufgewertet und umgekehrt.

In Abbildung 14 ist der Effekt der Basislinienverschiebung für die Kontroll-DNAs der Brustkrebsstudien dargestellt (siehe Abschnitt 0). Vergleicht man die Lage der Boxplots zueinander vor der Normalisierung, ist zu erkennen, dass die Boxen der Kontrollen mit 100% Methylierung und mit 0% Methylierung zum Teil völlig überlappen bzw. in einem Fall die Signalintensitäten der 100% methylierten Phi DNA sogar geringer sind als die der 0% methylierten. Nach der Normalisierung ist gut zu erkennen, dass zum einen die Verteilungen

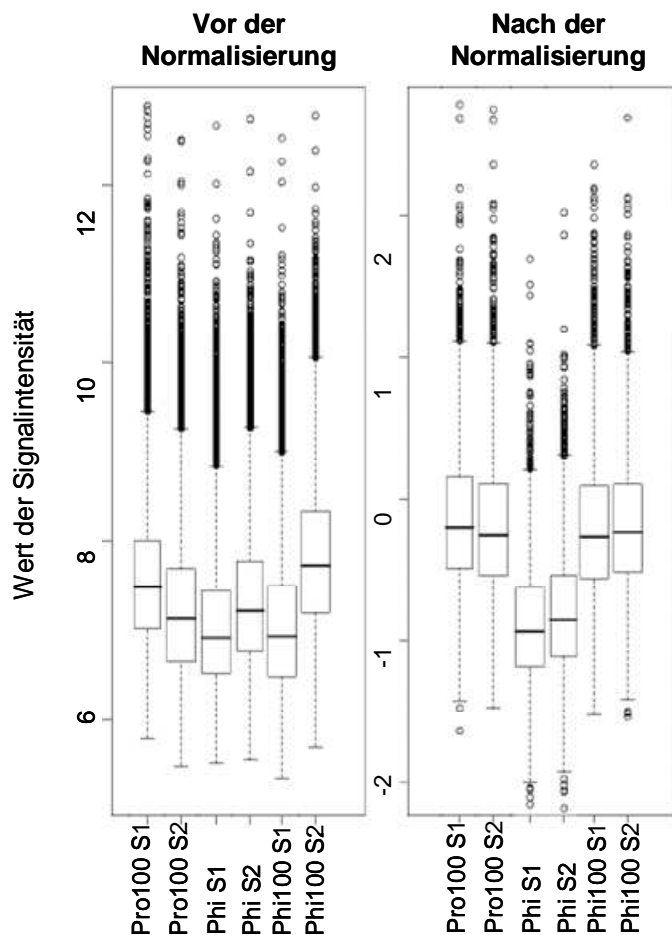


Abbildung 14: Auswirkung der Basislinienverschiebung. Boxplots der Signalintensitäten der einzelnen Sonden von 0%Phi und 100%Phi und 100%Pro Kontroll-DNAs vor und nach der Datenverarbeitung bis zum Schritt der Basislinienverschiebung. Die Box enthält 50% der Daten und der Median aller Daten wird durch die schwarze Linie dargestellt. Die gepunkteten Linien zeigen den Bereich des 1,5 fachen der Boxlänge an. Alle Werte die ausserhalb dieses Bereiches liegen, werden als einzelne Punkte dargestellt und Ausreisser bezeichnet.

von Arrays gleicher DNAs einander angepasst werden und zum anderen die 100% Kontrollen wesentlich höhere Werte aufweisen als die 0% Kontrollen, was für eine quantitative Auswertung essentiell ist. Auffällig ist weiterhin, dass sowohl vor als auch nach der Normalisierung eine gewisse Zahl von hohen Abweichungen (schwarze Punkte) zu erkennen ist, wobei es sich größten Teils um zu hohe Signale handelt. Dies ist möglich, da es sich um eine Darstellung einzelner Sonden handelt. Es ist anzunehmen, dass diese Sonden unspezifische Hybridisierung und somit hohe Signal zeigen. Daher wird mit dem Median der SONDENSETS gearbeitet. Der Median der Signalintensitäten befindet sich in der Mitte der Boxen. Daraus lässt sich in diesem Fall entnehmen, dass die Signale um den Median annähernd gleichmäßig verteilt sind.

3.1.5. Charakterisierung der optimierten DMH-Methode

3.1.5.1. Herstellung von Amplicons aus PBL und Brustkrebszelllinie

Für die Charakterisierung der optimierten Methode wurden PBL-DNA (Promega) und genomische DNA der Brustkrebszelllinie MDA-MB-231 verwendet. Es wurden zeitlich versetzt zwei Amplicons der beiden DNAs hergestellt. In Abbildung 15 werden die Agarosegele der beiden DNAs und der verschiedener Kontrollen für die einzelnen Prozess-Schritte gezeigt. Die Kontrollen und beide Proben zeigen in jedem Schritt die zu erwartenden Ergebnisse, d.h. die Herstellung des Amplicons war erfolgreich.

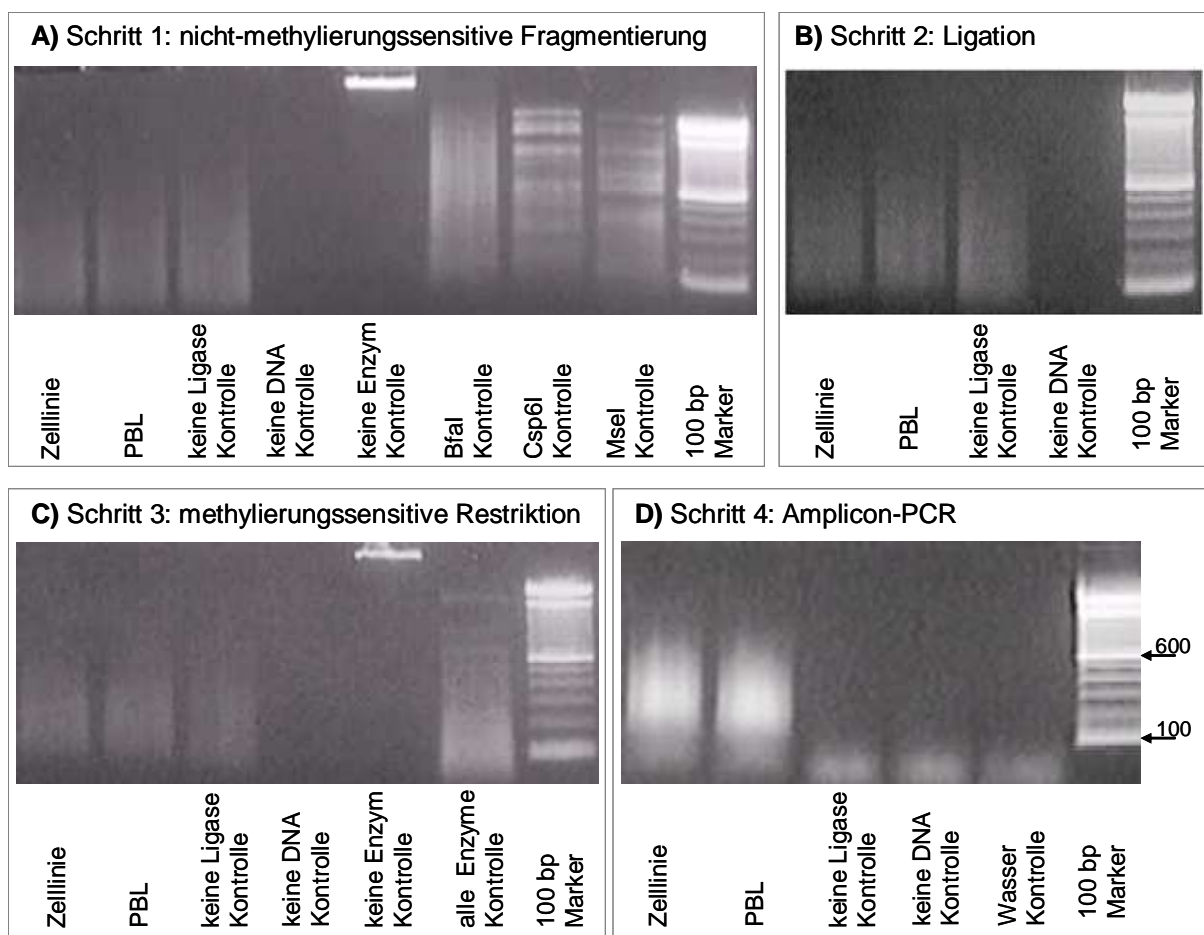


Abbildung 15: 1,4% Agarosegele des DMH-Amplicons 1 von PBL und Brustkrebszelllinien-DNA (als Zelllinie bezeichnet). A) Zelllinien und PBL-DNA sind vollständig fragmentiert, da sie einen DNA-Schmier zwischen 100bp und 700bp zeigen. Die keine DNA-Kontrolle ist negativ. λ -DNA, ohne Enzyme inkubiert, ist weiterhin intakt. Die Kontrolle der Bfal-Aktivität mit PBL-DNA und der MseI- und Csp6I-Aktivität mit λ -DNA sind positiv, da sie ein entsprechendes Bandenmuster oder einen DNA-Schmier zeigen. B) Dieses Gel kontrolliert hauptsächlich, dass die DNA nach Ligation und Aufreinigung noch vorhanden ist. Die Proben mit DNA zeigen den erwarteten Schmier zwischen 100bp und 700bp und die keine DNA Kontrolle ist leer. C) Die Proben mit DNA zeigen den erwarteten Schmier zwischen 100bp und 700bp und die keine DNA Kontrolle ist leer. λ -DNA, ohne Enzyme inkubiert, ist weiterhin intakt und zeigt mit Enzym ein spezifisches Bandenmuster. D) Für Zelllinien DNA und PBL-DNA ist ein PCR-Produkt zwischen 100bp und 700bp zu erkennen. Alle drei negativ Kontrollen sind negativ, d.h. es liegt keine Kontamination vor.

3.1.5.2. Reproduzierbarkeit

Um die Reproduzierbarkeit zu testen, wurden von den beiden DNAs jeweils zwei Amplicons zeitlich versetzt hergestellt und von jedem Amplicon, ausgehend von der PCR, zwei Hybridisierungen durchgeführt, d.h. es wurden für jede DNA vier Mikroarrays hybridisiert. In Abbildung 16 sind die Korrelationsplots der normalisierten Signalintensitäten der Fragmente und die zugehörigen Korrelationskoeffizienten (nach Pearson) der vier Hybridisierungen der Zelllinie MDA-MB-231 zu sehen. Sowohl die Korrelationskoeffizienten zwischen zwei Hybridisierungen der gleichen Amplicon-Präparation als auch zwischen den verschiedenen Amplicons liegen zwischen 0,93 und 0,95. Diese hohen Korrelationen zeigen, dass eine gute Reproduzierbarkeit der Hybridisierungssignale vorlag.

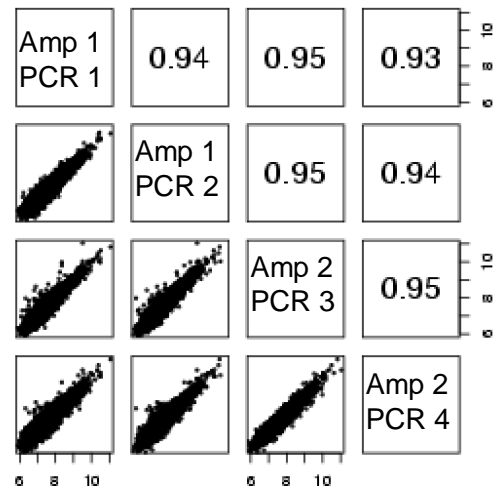


Abbildung 16: Korrelationskoeffizienten nach Pearson und Korrelationsplots der normalisierten Signalintensitäten. Es wurden zwei verschiedenen Amplicon-Präparatione: Amp 1 und Amp2 der Zelllinie MDA-MB-231 verglichen. Von jedem Amplicon wurden zwei Hybridisierungsprobe beginnend mit der PCR hergestellt und hybridisiert: PCR 1-4.

Weiterhin wurde überprüft, ob sich beim Vergleich der Blut-DNA und Brustkrebszelllinien-DNA (MDA-MB-231) die gleichen Unterschiede finden lassen, wenn jeweils die Daten eines Amplicons verwendet wurden. Das Ergebnis ist in Abbildung 17 dargestellt. Abbildung A zeigt die Verteilung der \log_2 -Signalquotienten (engl. \log_2 fold changes) zwischen den analysierten DNAs für Amplicon 1 über die gemittelte Signalintensität. Diese Darstellung wurde verwendet, um den Grenzwert für Markerkandidaten festzulegen und die Qualität der Daten zu prüfen. Der Signalquotient ist die Differenz (PBL – MDA-MB-231) der logarithmierten und gemittelten Signalintensitäten eines Fragmentes. Da die Daten logarithmiert sind, entspricht dies auch dem Quotienten der Signalintensität und ist ein Maß für die Größe des Unterschiedes, der zwischen den beiden DNAs gefunden wurde. Die Daten in Abbildung 17 A zeigen, dass die meisten Fragmente sich in einer Wolke geringer Signalquotienten befinden, d.h. der größte Teil der Fragmente keinen Unterschied aufwies. Abbildung 17 A zeigt außerdem, dass der größte Teil der Signalquotienten bei mittleren Signalintensitäten gefunden wurde und die Zahl der Fragmente gleichmäßig für hohe und niedrige Signalintensitäten abnimmt. Einige Fragmente streuen außerhalb der Wolke und zeigen größere Signalquotienten als der größte Teil der Fragmente. Die Verteilung dieser Fragmente über die Signalintensität zeigte ebenfalls keine Tendenz zu hohen oder niedrigen Signalen, was ein Hinweis auf fälschlich gemessene Unterschiede wäre. Daher wurden Fragmente mit Signalquotienten größer als 0,5 oder kleiner als -0,5 als Markerkandidaten für eine unterschiedliche Methylierung gewertet.

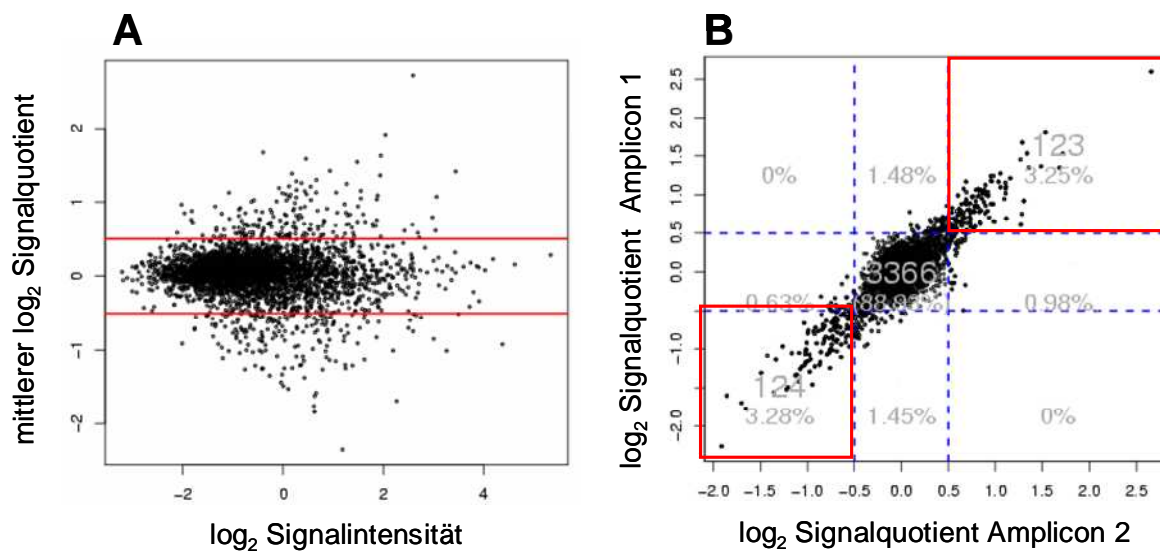


Abbildung 17: **A)** Verteilung der \log_2 Signalquotienten in Abhängigkeit von der mittleren Signalintensität für den Vergleich der Brustkrebszelllinie und PBL bei Amplicon 1. Die Festlegung eines Grenzwertes für Markerkandidaten bei einem Signalquotienten von 0,5 (methyliert in PBL) bzw. -0,5 (methyliert in der Brustkrebszelllinie) ist durch rote Linien gekennzeichnet. **B)** Korrelationsplot der beiden DMH-Amplicons. Die roten Boxen enthalten alle Fragmente die reproduzierbar als Markerkandidaten für Unterschiede zwischen PBL und Zelllinie identifiziert werden konnten.

Abbildung 17 B zeigt die Korrelation der Signalquotienten von Amplicon 1 und 2 mit eingezeichneten Grenzwerten des Signalquotienten von 0,5 bzw. -0,5. Da wiederum die meisten Fragmente keinen Unterschied aufweisen, bilden sie eine Wolke in der Mitte, die um den Nullwert streut. Für die Markerkandidaten zeigt sich eine sehr gute Übereinstimmung positiver und negativer Signalquotienten, d.h. es wurden in beiden Amplicons die gleichen Kandidaten für Methylierung in der Brustkrebszelllinie (negative Signalquotienten: 124 Fragmente) oder Methylierung in PBL (positive Signalquotienten: 123 Fragmente) gefunden. Daneben gibt es auch Kandidaten, die nur in einem der beiden Amplicons gefunden werden. Hierbei handelt es sich größtenteils um Signalquotienten, die nahe am Grenzwert liegen. Dies ist bei Verwendung eines Grenzwertes nicht zu vermeiden, da eine 100% Korrelation der Werte nicht zu erreichen ist. Die linke obere Box und die rechte untere Box sind leer. Hierbei würde es sich um Kandidaten handeln, die in den beiden Amplicons widersprüchliche Ergebnisse zeigten. Insgesamt konnte somit auch für die Markerkandidaten auf eine sehr gute Reproduzierbarkeit geschlossen werden.

3.1.5.3. Überprüfung mittels Bisulfitsequenzierung

Um die gefundenen Methylierungsunterschiede zu überprüfen, wurden 111 Fragmente ausgewählt und durch direkte Bisulfitsequenzierung untersucht. Die Fragmente wurden aus den folgenden drei Gruppen zufällig ausgewählt: i) gute Kandidaten: Signalquotient $> 0,6$ oder $< -0,6$; ii) grenzwertige Kandidaten Signalquotient: $0,4$ bis $0,6$ und $-0,4$ bis $-0,6$ iii) Fragmente ohne Unterschied: Signalquotient $< 0,4$ und $> -0,4$. Diese Vorgehensweise ist notwendig, da die verschiedenen Gruppen sich in der Anzahl der zugehörigen Fragmente stark unterscheiden. Würde man ohne vorherige Gruppierung zufällig Fragmente auswählen, wäre der größte Teil aus der Gruppe der Fragmente ohne Unterschied.

Die Auswertung erfolgte mit Hilfe einer speziellen Software, die für jedes CpG eine mittlere DNA-Methylierung berechnet (siehe Abschnitt 2.3.). Um die Bisulfitsequenzierungsdaten mit den Signalquotienten vergleichen zu können, wurde für jedes Fragment der Median der Methylierung aller CpGs ermittelt und ebenfalls die Differenz (PBL – MDA-MB-231) gebildet.

In Abbildung 18 ist der Vergleich der Methylierungsdifferenzen der Bisulfitsequenzierung mit den gemittelten Signalquotienten der beiden Amplicons dargestellt. Der Korrelationskoeffizient für diesen Vergleich betrug $0,71$. Eine höhere Korrelation wurde nicht erwartet, da die sequenzierten Fragmente nur im idealen Fall genau das DMH Fragment abdecken. In den meisten Fällen gibt es lediglich eine Überschneidung der beiden Fragmente. Weiterhin ist der DMH-Wert abhängig vom CpG mit der geringsten DNA-Methylierung während bei der Bisulfitsequenzierung der Mittelwert aller sequenzierten CpGs gebildet wurde. Aufgrund dieser Unterschiede kann die gefundene

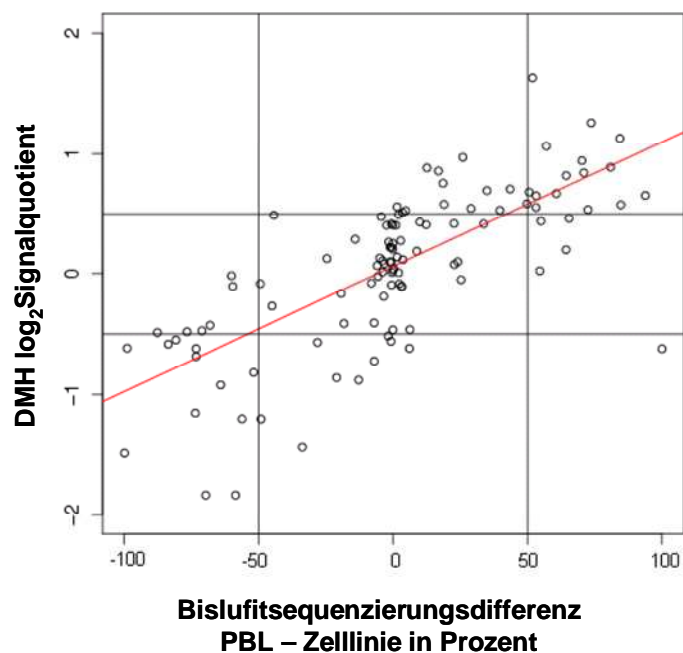


Abbildung 18: Korrelationsplot der DMH - Signalquotienten mit durch direkte Bisulfitsequenzierung ermittelten Methylierungsunterschieden der Brustkrebszelllinie und PBL. Es wurden 111 Fragmente aus drei Gruppen der DMH-Daten ausgewählt: gute Kandidaten, grenzwertige Kandidaten, keine Kandidaten. Die Signalquotienten des DMH wurden mit den Differenzen der mittleren Methylierung aller CpGs der bisulfitsequenzierten Sequenzabschnitte verglichen.

Korrelation als hoch betrachtet werden. Diese gute Korrelation zeigt, dass mit Hilfe der hier beschriebenen verbesserten DMH-Methode DNA-Methylierungsunterschiede gefunden werden können und eine quantitative Auswertung möglich ist, da auch die Größe der Signalquotienten mit den DNA-Methylierungswerten korreliert.

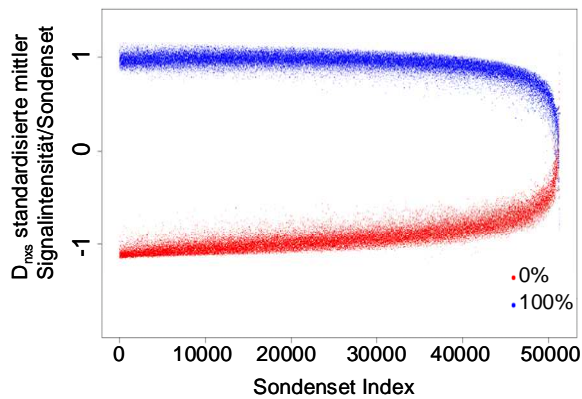
3.1.6. Der 50.000-Fragmente-Mikroarray

Auf der Basis der beschriebenen Ergebnisse für den Entwicklungs-Mikroarray, wurde ein Mikroarray zur Detektion von rund 50.000 Fragmenten designed. Das Ergebnis des in Abschnitt 2.2.9.2 beschriebenen Designs war ein Mikroarray der 491.491 Detektionssonden zur Analyse von 51317 Fragmenten und 9834 Kontrollsonden, die 1000 Fragmente abdecken, enthielt. Beim diesem Design wurden die Fragmente auch nach der Sequenzart ausgesucht, d.h. es wurden Fragmente, die mit einem Promotor assoziiert werden können, da sie in der Nähe eines TSS liegen, ausgewählt und es wurden ebenfalls Fragmente bevorzugt, die mit einer CpG-Insel überlappen sowie CpG arme Regionen ausgeschlossen (siehe auch Abschnitt 2.2.9). In Tabelle 6 ist die erreichte Abdeckung dieser Sequenzarten aufgelistet. Der 50.000-Fragmente-Mikroarray deckt 46% aller Promotorregionen und 47% aller Ensembl annotierten CpG-Inseln ab. Betrachtet man die Anzahl der Fragmente im Vergleich zur Anzahl abgedeckter Sequenzen, wird deutlich, dass in vielen Fällen mehr als ein Fragment für die entsprechende Region vorhanden ist. Weiterhin ist zu bemerken, dass für alle anderen CpG-Inseln oder Promotoren kein Fragment designed werden konnte, da sie eines der Filterkriterien, wie Gehalt repetitiver Sequenzen, CpG-Gehalt oder Länge nicht erfüllt haben. Weiterhin gibt es 22964 Fragmente, die Exons oder Introns von insgesamt 13654 verschiedenen Genen abdecken und 8870 Fragmente liegen in Regionen, die weder mit einem Gen noch mit einem Promotor assoziiert sind. Die in Tabelle 6 aufgelisteten Fragmente bilden Schnittmengen untereinander und mit den restlichen Fragmenten.

Tabelle 6: Fragmente des 50.000-Fragmente-Mikroarray, die mit einem TSS oder einer CpG-Insel assoziiert sind.

	Ensembl annotiert	Abgedeckt durch Mikroarray	Abdeckung in %	Anzahl der Fragmente
± 2000bp TSS	30391	14017	46%	24410
CpG Insel	22395	10522	47%	18293

Um den 50.000-Fragmente-Mikroarray zu charakterisieren und zu testen, ob der Array auch für die quantitative Auswertung geeignet ist, wurden Amplicons von 0% Phi und 100% Phi hergestellt. Es wurden zwei Amplicons pro DNA hergestellt und jeweils zwei Mikroarrays mit Hybridisierungsproben verschiedener PCRs hybridisiert. Das Ergebnis ist in Abbildung 19 dargestellt. Der Plot visualisiert die Fähigkeit eines SONDENSETS, 0% und 100% methylierte DNA zu unterscheiden. Die SONDENSETS werden auf Grund der T-Statistik der Unterscheidung sortiert. Je kleiner der Wert der T-Statistik, desto besser unterscheidet das SONDENSET und desto kleiner ist auch der SONDENSET INDEX, der auf der X-Achse aufgetragen ist. Auf der Y-Achse ist die zwischen -1 und 1 standardisierte mittlere Signalintensität des jeweiligen SONDENSETS für die jeweilige DNA aufgetragen. Die Standardisierung erfolgt über die Standardabweichung der Datenpunkte aller DNAs und sollte daher für 0% = -1 und 100% = 1 sein.



$$D_{nxs} = (D_{xs} - \text{mean}(D_x)) / \text{sd}(D_x)$$

D_x = nicht normalisierte Signalintensität des Sondenset X von allen Proben

D_{xs} = nicht normalisierte Signalintensität des Sondenset X von Probe S für eine der beiden DNAs

D_{nxs} = standardisierte mittlere Signalintensität des Sondenset X von Probe S für eine der beiden DNAs

Abbildung 19: Fähigkeit eines Sondensets 0% und 100% methylierte DNA zu unterscheiden. Auf der X Achse ist der Sondenset Index aufgetragen. Die Sondensets wurden anhand der TStatistik sortiert. Ein kleiner Index bedeutet kleine TStatistik und somit gute Unterscheidung. Auf der Y Achse ist die standardisierte Signalintensität D_{nxs} , berechnet nach der angegebenen Formel, für 0% Phi DNA (rote Punkte) und 100% Phi DNA (blaue Punkte) dargestellt. Erst bei sehr großem Sondenset Index beginnen die Datenpunkte beider DNAs zu streuen und von den Erwartungswerten 1 und -1 stark abzuweichen, so dass keine Unterscheidung mehr möglich ist.

In der Abbildung ist zu erkennen, dass der größte Teil der Sondensets gut zwischen 0% und 100% Methylierung unterscheidet. Erst ab dem circa 45.000 Sondenset beginnen die Werte der Signalintensitäten von den Erwartungswerten 1 und -1 abzuweichen und erst bei Fragmenten mit einem Index > 49.000 beginnen sich die Signalintensitäten zu mischen, was bedeutet, dass bei diesen Sondensets eine Unterscheidung von 100% und 0% nicht möglich ist. Zusammenfassend kann man also davon ausgehen, dass dieser Mikroarray zur reproduzierbaren Analyse der Methylierung von circa 50.000 unterschiedlichen Fragmenten verwendet werden kann.

3.2. Analyse der DNA – Methylierung in Brustkrebs

Im vorangegangenen Kapitel wurde gezeigt, dass eine verbesserte DMH Methode zur genomweiten Analyse von DNA-Methylierungsmustern etabliert werden konnte. Ein Hauptziel der vorliegenden Arbeit war die Suche nach Kandidaten für das Ansprechen von metastasierendem Brustkrebs auf eine Chemotherapie. Im folgenden Kapitel werden die Ergebnisse, die hinsichtlich dieser Fragestellung erzielt werden konnten, gezeigt. Hierzu wurden 28 Tumorproben sowie verschiedene Kontrollen und gesundes Gewebe mit der verbesserten DMH Technologie untersucht. Da dieser Teil der Arbeit parallel zur Entwicklung des 50.000-Fragmente-Mikroarrays durchgeführt wurde, wurde für die Analyse der Entwicklungsarray verwendet.

Es ist allgemein bekannt, dass in Tumoren im Vergleich zu gesundem Gewebe eine Vielzahl veränderter DNA-Methylierungsmuster vorkommt während für die Suche nach Markern für das Ansprechen auf eine Chemotherapie nur wenige Kandidaten zu erwarten sind (siehe auch Abschnitt 1.1.6.3). Da erst in den letzten Jahren Technologien für die genomweite DNA-Methylierungsanalyse entwickelt wurden, liegen jedoch auch für den Vergleich von gesundem Gewebe mit Tumorgewebe noch wenige Datensätze von Methylierungsprofilen vor. Zunächst werden daher die Ergebnisse für den Vergleich der Brusttumore mit gesundem Gewebe und im Anschluss die Ergebnisse für die Suche nach Markerkandidaten für das Ansprechen auf eine Chemotherapie gezeigt. Im Folgenden werden die Ergebnisse der Untersuchung der DNA-Methylierungsmuster mittels Bisulfitsequenzierung dargestellt. Dafür wurden stichprobenartig Kandidaten ausgewählt. Die Ergebnisse der Bisulfitsequenzierung wurden mit den in der DMH-Analyse erhaltenen Daten verglichen. Den Abschluss bilden Bisulfitsequenzierungsergebnisse von weiteren Proben für einen Teil der analysierten Kandidaten, dies diente der Betrachtung der gefundenen Unterschiede mit einer unabhängigen Methode auf unabhängigen Proben.

3.2.1. Verwendete Proben und Gruppeneinteilung

Die Tumorproben für diese Studie wurden von Dr. John Foekens (Erasmus Medical Center, Rotterdam) im Rahmen des EU-Projektes „Epigenetic profiling of breast cancer: prognostic and therapeutic applications“ zur Verfügung gestellt. Das positive Ethikvotum für diese Proben und das Studienprotokoll wurden eingeholt. Es handelte sich um Frischgewebeproben primärer Tumore von Patientinnen, die im Verlauf ihrer Erkrankung Metastasen entwickelt haben, zu deren Behandlung als erste Chemotherapie ein FAC Schema, bestehend aus 5-Fluorouracil, dem Anthracyclin Adriamycin und Cyclophosphamid, eingesetzt wurde. Alle Brustkrebspatientinnen erhielten keine adjuvante Therapie, d.h. auch keine Hormontherapie. Die Beurteilung des Ansprechens der Metastasen auf die Therapie

wurde nach UICC (Union International Contre le Cancer) Standards vorgenommen [122]. Hierbei gibt es folgende Gruppen: Tumore, deren Metastasen auf die Therapie ansprechen und die Metastase schrumpft, bis sie nicht mehr detektierbar ist. Diese Gruppe wird als Gruppe mit vollständiger Remission (engl. complete remission, im Folgenden bezeichnet als CR) bezeichnet. Verkleinert sich die Metastase, bleibt aber detektierbar, handelt es sich um eine unvollständige Remission (engl. partial remission, im Folgenden bezeichnet als PR). Dem gegenüber stehen Tumore, deren Metastasen während der Therapie weiter wachsen. Diese werden als Gruppe mit fortschreitender Erkrankung bezeichnet (engl. progressive disease, im Folgenden bezeichnet als PD). Weiterhin gibt es eine Gruppe von Tumoren, bei denen zunächst weder ein Wachstum noch ein Rückgang der Metastasen festgestellt werden kann. Diese werden als Gruppe der unveränderten Erkrankungen (stable disease, im Folgenden bezeichnet als SD) bezeichnet.

Für die DMH Studie wurden 14 Tumorproben, deren Metastasen auf die Therapie ansprachen (im Folgenden als Ansprechende bezeichnet, engl. Responder), und 14, deren Metastasen nicht auf die Therapie reagierten (im Folgenden als Nicht-Ansprechende bezeichnet, engl. Non-Responder), verwendet (siehe auch im Anhang: Tabelle 13). Für die Gruppe der Ansprechenden wurden PR und CR und für die Gruppe der Nicht-Ansprechenden PD Tumorproben verwendet. Die Zuordnung der Gruppe der SD Tumorproben zur Gruppe der Ansprechenden oder Nicht-Ansprechenden ist fraglich. Häufig werden sie den Ansprechenden zugeordnet, da auch das zeitweise Unterdrücken des Tumorwachstums einen Vorteil für die Patientinnen darstellt und aus molekularbiologischer Sicht angenommen werden kann, dass eine zytostatische Wirkung nachgewiesen ist. Da der Rückgang einer Metastase (PR) bis hin zum vollständigen Verschwinden (CR) jedoch auf eine wesentlich stärkere Reaktion der Metastase schließen lässt, wurde die Verwendung dieser Proben bevorzugt.

3.2.2. Studienaufbau

Die hier vorgestellten Daten wurden in zwei aufeinander folgenden Studien ermittelt. In der ersten Studie (im Folgenden als Tumorstudie bezeichnet) wurden die Tumorproben zusammen mit zwei verschiedenen Produktionslots von Blut-DNA und als DNA-Methylierungsstandards eine Phi0%-DNA (siehe 2.1.3) sowie einer Phi100%-DNA (siehe 2.1.4) untersucht. Die zweite Studie (im Folgenden als Kontrollstudie bezeichnet) diente dazu, Daten für gesundes Brustgewebe zu erhalten, da im Verlaufe der Tumorstudie festgestellt wurde, dass Blut als gesunde Referenz nicht ausreichend ist. Hierfür wurde zwei DNA-Pools gesunden Brustgewebes aus Brustreduktionen von jeweils sieben Frauen, die DNA einer primären Zelllinie (6 Passagen) von Brustepithelzellen (human mammary epithelial cells, HMEC, Cambrex), eine Blut-DNA und als DNA-Methylierungsstandards eine Phi0%-DNA (siehe 2.1.3) und eine Pro100%-DNA (siehe 2.1.4) verwendet.

Tabelle 7: Zahl und Art verwendeter Proben sowie Anzahl hybridisierter Mikroarrays

Gruppe	Bezeichnung	Proben	Mikroarrays
auf die Therapie Ansprechende	CR , vollständige Remission der Metastase (engl.: complete remission)	3	6
	PR , partielle Remission der Metastase (engl.: partial remission)	11	22
auf die Therapie Nicht-Ansprechende	PD , Wachstum der Metastase (engl.: progressive disease)	14	27*
Gesundes Gewebe	Pool B1 Brustgewebe gesunder Frauen 1	1	2
	Pool B2 Brustgewebe gesunder Frauen 2	1	2
	HMEC , Primäre Zelllinie von Brustepithelzellen (engl.: human mammary epithelial cells)	1	2
PBL (Periphere Blut- lymphozyten, Promega**)	Blut 1 (Tumorstudie)	1	2
	Blut 2 (Tumorstudie)	1	2
	Blut 3 (Kontrollstudie)	1	2
Kontrollen ***	Phi0% (Tumorstudie)	1	2
	Phi0% (Kontrollstudie)	1	2
	Pro100% (Kontrollstudie)	1	2

*Verlust einer Probe in Amplikon 1

**Pools von durchschnittlich 5 Individuen

***100% Phi der Tumorstudie wurde nicht verwendet

Es wurden grundsätzlich pro DNA zwei Wiederholungen des DMH-Amplicons hergestellt und daher für jede Probe zwei Arrays hybridisiert. In Tabelle 7 ist der Studienaufbau zusammengefasst. Bis auf eine Probe, die während der Ampliconherstellung scheiterte, wurden für jede Probe zwei Hybridisierungen erfolgreich durchgeführt, d.h. insgesamt 75 Entwicklungs-Mikroarrays mit DMH-Amplicons von 36 verschiedenen Proben hybridisiert.

3.2.3. Datenqualität und Vergleichbarkeit der zwei Studien

Alle Daten wurden, wie in Abschnitt 2.2.10 beschrieben, normalisiert. Der durchschnittliche Korrelationskoeffizient (nach Pearson) der technischen Wiederholungen nach der Normalisierung betrug 0,95. In drei Fällen war die Korrelation kleiner als 0,90 (0,89; 0,89, 0,88). Da es nicht möglich war, einen dieser Mikroarrays aufgrund technischer Kriterien, wie z.B. visuell detektierbarer Verunreinigungen oder anderer Anomalien des Scanbildes, auszuschließen, wurden die Daten aller Mikroarrays für die Auswertung verwendet. Eine wesentliche Fragestellung bei der Auswertung war die Vergleichbarkeit der Daten aus den zu verschiedenen Zeitpunkten durchgeführten Studien. Es ist bekannt, dass dies zu einem systematischen Unterschied der Daten führen kann, der auch nicht durch die Normalisierung ausgeglichen werden kann. Dieser Effekt wird im Folgenden als Batcheffekt bezeichnet.

In Tabelle 8 sind Korrelationskoeffizienten für ausgewählte Proben zusammengestellt. In den Tumorstudien waren 2 Pools und in der Kontrollstudie war ein Pool von Blut-DNA (Promega) enthalten. Der Korrelationskoeffizient der beiden Blutpools, die innerhalb der Tumorstudie eingesetzt wurden, beträgt 98 %, während der Vergleich der Blutprobendaten zwischen den beiden Studien eine Korrelation von 94 % und 95 % ergibt. Da es sich bei allen drei Pools

um verschiedene Produktionslots und somit DNAs verschiedener Individuen handelt, ist nicht auszuschließen, dass ein gewisser Teil dieser Unterschiede auf biologische Unterschiede der Proben zurückzuführen ist. Die Korrelation der beiden unterschiedlichen DNAs innerhalb der Tumorstudie ist jedoch hoch, daher muss man annehmen, dass die geringere Korrelation zwischen den Studien zu einem gewissen Teil durch einen Batcheffekt hervorgerufen wird.

Tabelle 8: Korrelationskoeffizienten der Hybridisierungsdaten ausgewählter Proben nach der Normalisierung

Gruppe	Vergleich	Studie	Korrelation
Gesundes Gewebe	Pool, post/pre	Kontroll-/Kontrollstudie	0,98
	Pool, post/HMEC	Kontroll-/Kontrollstudie	0,93
	Pool, pre/HMEC	Kontroll-/Kontrollstudie	0,94
Blut	Blut 1/2	Tumor-/Tumorstudie	0,98
	Blut 1/3	Tumor-/Kontrollstudie	0,94
	Blut 2/3	Tumor-/Kontrollstudie	0,95
Kontrollen	Phi0%	Tumor-/Kontrollstudie	0,94
	Pro100%/Phi100%	Tumor-/Kontrollstudie	0,85
Gesundes Gewebe/Blut	Pool, post/Blut 1; 2; 3	Tumor-/Kontrollstudie	0,94; 0,94; 0,94
	Pool, pre/Blut 1; 2; 3	Tumor-/Kontrollstudie	0,93; 0,94; 0,94
	HMEC/Blut 1; 2; 3	Tumor-/Kontrollstudie	0,87; 0,88; 0,89

In Tabelle 8 sind weiterhin die Korrelationen der drei Proben gesunden Gewebes dargestellt. Die beiden Pools von Brustgewebe pre- und postmenopausaler Frauen zeigen eine Korrelation von 98% genauso, wie die Daten der beiden Blutproben, die in der Tumorstudie bearbeitet wurden. Die Korrelationen dieser Proben aus der Kontrollstudie mit der ebenfalls in der Kontrollstudie enthaltenen DNA der HMECs (primären Epithelzelllinie) liegen bei 93% bzw. 94%. In diesem Fall repräsentiert dieser Unterschied der Korrelationen den Unterschied der Proben. Während die beiden Pools von Brustgewebe DNAs des gleichen Gewebes enthalten und somit grundsätzlich sehr ähnlich sein sollten, handelt es sich bei der HMEC-DNA um DNA von aufgereinigten und über sechs Passagen amplifizierten Epithelzellen der Brust, die sich vom Gemisch der normalen Gewebeproben (Epithel, Stroma, Fett und Blut) unterscheiden.

Der Unterschied wird noch deutlicher, betrachtet man die Korrelationen, die sich beim Vergleich der gesunden Gewebekontrollen mit den Blut DNAs ergeben. Die gepoolten DNAs der gesunden Gewebe zeigen die gleichen hohen Korrelationen von 94% mit allen drei Blutproben während die Korrelation der HMECs mit den Blutproben sehr gering ist. In beiden Fällen verhalten sich die Korrelationen unabhängig von der Studie aus der die Blutprobe stammte, d.h. der Batcheffekt zeigt hier keine Auswirkung. Die hohe Korrelation der Blut-DNAs mit den DNAs gesunder Gewebe, weist auf einen Anteil von Blut in diesen Proben hin. Zusammenfassend zeigt diese Betrachtung, dass sich der Batcheffekt nicht überall ausprägt und dass die Unterschiede zwischen zwei Geweben (Blut und Epithel) wesentlich größer und somit detektierbar sind. Grundsätzlich kann der Batcheffekt jedoch Auswirkungen auf die Rate falsch-positiver und falsch-negativer Markerandidaten in den anschließenden

Analysen haben. Hierbei soll auf folgende Punkte hingewiesen werden. Zum einen wirkt sich der Effekt ausschließlich auf den Vergleich der gesunden Gewebeproben mit den Tumorproben aus, bei welchem mit relativ großen Unterschieden zwischen den beiden Gruppen zu rechnen ist, d.h. trotzdem Kandidaten gefunden werden sollten. Zum anderen hat er auf den Vergleich von Chemotherapie-Ansprechenden mit Nicht-Ansprechenden keine Auswirkung, da diese Proben alle in der Tumorstudie bearbeitet wurden.

3.2.4. Einteilung in drei DNA-Methylierungsgruppen mit Hilfe der 0% und 100% Kontrollen

Im nächsten Schritt wurden die Daten mit Hilfe der DNA-Methylierungskontrollen normiert. Unvoreilhafterweise wurde die Tumorstudie noch mit Phi100% DNA durchgeführt. Da Phi-DNA einem Amplifikationsschritt unterzogen wurde (siehe Abschnitt 6.1.3), liegen in dieser DNA verschiedene genomische Regionen, wie repetitive Elemente aber auch als Einzelkopie vorkommende DNA Abschnitte, in veränderter Kopienzahl vor [117]. Wird diese DNA als Kontrolle für unmethylierte DNA verwendet (Phi0%), sollte eine veränderte Kopienzahl keinen Einfluss haben, da die entsprechende Sequenz vollständig geschnitten wird. Im Fall der aufmethylierten 100%Phi DNA wirken sich diese Unterschiede jedoch auf das Hybridisierungsergebnis aus, was die geringe Korrelation der Pro100% und Phi100% in Tabelle 8 erklärt. Daher wurde für die Normierung der Daten die Phi100% Kontrolle verworfen und ausschließlich die Daten der Pro100% der Kontrollstudie verwendet.

Zunächst wurde für jedes Fragment die Differenz der Signalintensitäten der DNA-Methylierungskontrollen gebildet ($D_{\text{Pro100\%}} - D_{\text{Phi0\%}}$). Diese Differenz wird als Dynamikumfang (engl.: dynamic range) bezeichnet. Je größer der Dynamikumfang eines Fragmentes ist, desto besser kann man einen prozentualen DNA-Methylierungswert bestimmen. Ist der Wert allerdings sehr klein, wird es immer ungenauer innerhalb dieses Bereichs eine Skala von 0 % bis 100 % aufzuspannen. Abbildung 20 zeigt die Dynamikumfangs aller Fragmente, der Größe nach sortiert, beginnend mit dem kleinsten Wert. Es ist gut zu erkennen, dass der größte Teil der Fragmente Werte zwischen 0,5 und 2,0 aufweist. Der Ausschnitt der ersten 300 Fragmente zeigt, dass die Kurve bis zu einem Wert von 0,3 stark ansteigt und dann in einen flachen Verlauf übergeht. Für die nachfolgenden Analysen wurden alle Fragmente mit einem Wert größer als 0,3 verwendet. Dabei handelt es sich um 3693 der 3780 Fragmente, d.h. circa 2% der Fragmente wurden aufgrund eines kleinen dynamischen Bereichs von der weiteren Analyse ausgeschlossen.

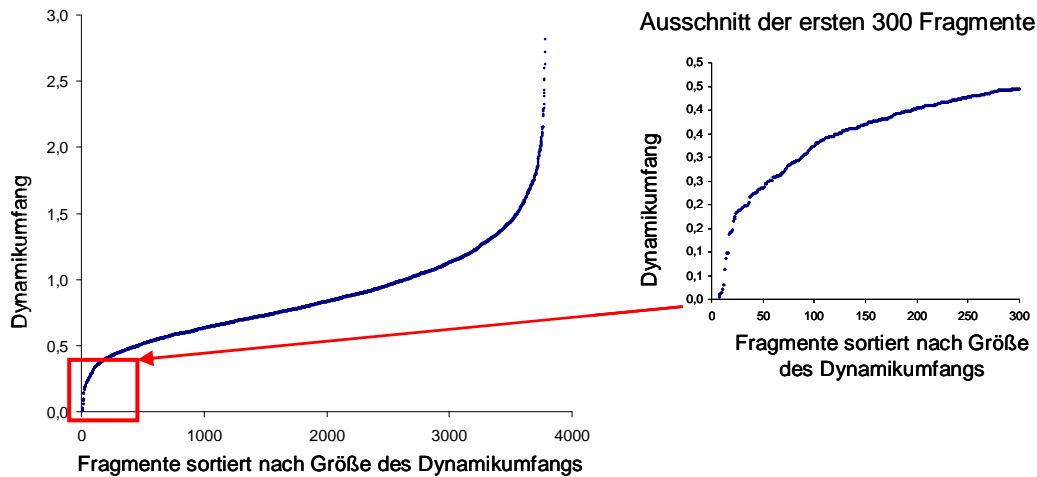


Abbildung 20: Der Dynamikumfang aller Fragmente, der Größe nach sortiert (links) und die Vergrößerung des Bereichs der ersten 300 Fragmente mit den kleinsten Dynamikumfangen (rechts). Der Dynamikumfang ist die Differenz aus den gemittelten, normalisierten Signalintensitäten von Pro100% und Phi0%: $D_{Pro100\%} - D_{Phi0\%}$.

Die 3693 Fragmente wurden anschließend in die drei Gruppen unmethyliert, semimethyliert oder methyliert in gesundem Brustgewebe eingeteilt. Dazu wurden die Daten der zwei Pools gesunden Brustgewebes und die Daten der HMECs verwendet und der durchschnittliche DNA-Methylierungsprozentsatz mit der Formel $(S_{Probe} - S_{Phi0\%}) / (S_{Pro100\%} - S_{Phi0\%})$ berechnet. Fragmente mit weniger als 30 % DNA-Methylierung wurden in die Gruppe der unmethylierten, solche mit mehr als 70 % DNA-Methylierung in die der methylierten und alle mit Werten von 30 % bis 70 % in die der semimethylierten Fragmente eingeordnet.

3.2.5. Verteilung der Methylierung in gesundem Gewebe

Aufgrund der im vorangegangenen Abschnitt beschriebenen Einteilung wurden 481 (13%) Fragmente als unmethyliert, 2043 (55%) Fragmente als semimethyliert und 1169 (32%) Fragmente als methyliert in gesundem Brustgewebe eingeordnet. In Abbildung 21 ist jeweils der Anteil der Fragmente, die mit einer CpG-Insel (Definition siehe Abschnitt 2.2.8) mit mindestens 50 bp überlappen, die in der Nähe (± 2000 bp) eines Transkriptionsstarts liegen oder die mit mindestens 50 bp mit einem Gen überlappen, für jede der drei Gruppen im Vergleich zu den Verhältnissen auf dem Mikroarray dargestellt.

Betrachtet man zunächst die Verteilung der Fragmente auf die drei Gruppen, ist auffällig, dass die unmethylierten Fragmente mit 13 % die kleinste Gruppe, die semimethylierten mit 53% die größte Gruppe und die Gruppe der methylierten mit 32 % die mittlere Gruppe ausmachen. Geht man davon aus, dass CpG-Inseln in gesunden Geweben nicht methyliert sind, würde man mit einer anderen Verteilung rechnen. Der Array enthält jedoch nicht nur CpG-Inseln sondern auch CpG-ärmere Fragmente. Dies und andere Einflüsse, die diese Verteilung begründen, werden in Abschnitt 4.2.2 diskutiert.

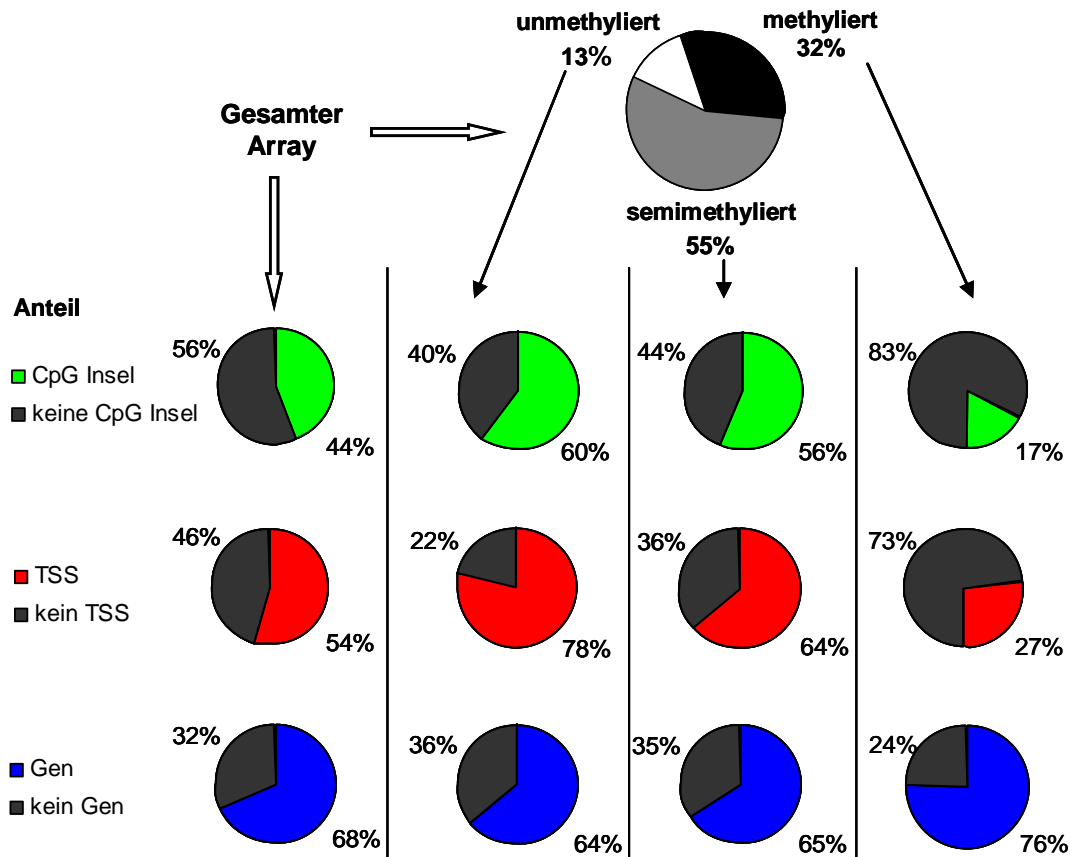


Abbildung 21: Verteilung der Fragmente des Arrays auf die drei Methylierungsgruppen: unmethyliert, semimethyliert und methyliert in gesundem Brustgewebe sowie jeweils der Anteil der Fragmente pro Methylierungsgruppe, die mit einer CpG-Insel oder einem Gen überlappen oder die in der Nähe ($\pm 2000bp$) eines Transkriptionsstart (TSS) lokalisiert sind.

Betrachtet man die Anteile von CpG Inseln, Transkriptionsstarts (TSS) oder Genen an der Zahl der Fragmente für jede Gruppe, wird deutlich, dass sich die drei Gruppen hierin stark unterscheiden. Von den 481 Fragmenten der unmethylierten Gruppe überlappen 60% mit CpG-Inseln, 78% liegen in der Nähe eines Transkriptionsstarts und 64% überlappen mit einem Gen. Die Gruppe der semimethylierten Fragmente zeigt wie erwartet ähnliche, wenn auch etwas kleinere Anteile als die unmethylierten Fragmente. Im Vergleich dazu unterscheidet sich die methylierte Gruppe stark, deren Fragmente nur noch zu 17% mit einer CpG Insel überlappen, zu 27% in der Nähe eines Transkriptionsstarts liegen aber zu 76% mit einem Gen überlappen, d.h. hierbei handelt es sich zum größten Teil um Sequenzen, welche in Exons liegen. Diese Verhältnisse entsprechen der Erwartung, da geringe DNA-Methylierung hauptsächlich in CpG-Inseln, in der Nähe eines Transkriptionsstarts und hohe DNA-Methylierung in Exons und Introns zu finden sein sollten. Daher ist anzunehmen, dass in der jeweiligen Gruppe grundsätzlich die zu erwartenden Fragmente angereichert sind.

3.2.6. Kriterien für die Auswahl von Markerkandidaten

Als Maß zur Unterscheidung von Gruppen wurden auch die Tumorproben in die drei DNA-Methylierungsgruppen eingeteilt und gezählt, wie viele Proben jeweils zu einer Gruppe zuzuordnen sind. Diese Vorgehensweise wurde gewählt, da sie die Proben qualitativ einteilt und so eine gezielte Auswahl anhand der DNA-Methylierung möglich war, ohne ein statistisches Maß zu verwenden, was bei den hier verwendeten Probenzahlen und Fragestellungen als weniger sinnvoll angesehen wurde. D.h. es wurde bei dieser Analyse nicht erwartet, statistisch signifikante Marker zu finden, sondern vielmehr versucht, qualitativ Fragmente zu filtern, die einen bestimmten DNA-Methylierungsunterschied aufwiesen. Da die Einteilung der Gruppen auf der Festsetzung von Grenzwerten beruht, besteht natürlich die Möglichkeit, Proben, deren DNA-Methylierung um diesen Grenzwert schwanken, mal in die eine und mal in die andere Gruppe zu sortieren und daher nur durch Zufall einen Unterschied zu erzeugen.

Um dies zu vermindern, wurde als weiterer Filter der Wert der T-Statistik verwendet. Es wurden nur Fragmente als Kandidaten gewertet für die der Wert der T-Statistik größer als 1 bzw. kleiner als -1 war. Dieser Wert ist relativ klein, d.h. dieser Filter diente dazu, Fragmente zu filtern, bei denen entweder die Messwerte sehr stark schwanken und somit die Verteilungen der beiden Gruppen stark überlappen oder bei denen der Unterschied zwischen den beiden Mittelwerten sehr gering ist, andererseits werden noch viele Fragmente zugelassen, deren Unterschiede auch auf solchen Effekten beruhen können. Auf die Verwendung weiterer statistischer Filter wurde bewusst verzichtet, da eine Auswertung dieser Art für die hier behandelten Fragestellungen und Probenzahlen bzw. Gruppengrößen nicht als sinnvoll erachtet wurde.

3.2.7. Unterschiede zwischen den gesunden Gewebekontrollen und den Tumorproben

Aufgrund der vielfach beschriebenen veränderten DNA-Methylierung von Tumoren, sind für den Vergleich der Proben gesunder Gewebe und der Tumorproben Unterschiede zu erwarten. Zum einen wird beschrieben, dass in Tumoren zuvor unmethylierte CpG Inseln hypermethyliert werden. Auf der anderen Seite wird insbesondere für Brustkrebs eine globale Demethylierung oder Hypomethylierung beschrieben, d.h. viele Sequenzabschnitte, die in gesundem Gewebe methyliert vorliegen, sind im Tumor nicht mehr methyliert. Daher ist hier die getrennte Betrachtung der beiden Gruppen unmethyliert oder methyliert in gesundem Gewebe besonders interessant. Als Kandidaten wurden Fragmente gewertet, bei denen mindestens ein Viertel (7) der 28 Tumorproben nicht in die gleiche DNA-Methylierungsgruppe wie die gesunden Kontrollen eingeordnet wurde. Es handelt sich hierbei um Tumormarker die Kandidaten für Früherkennungsmarker darstellen.

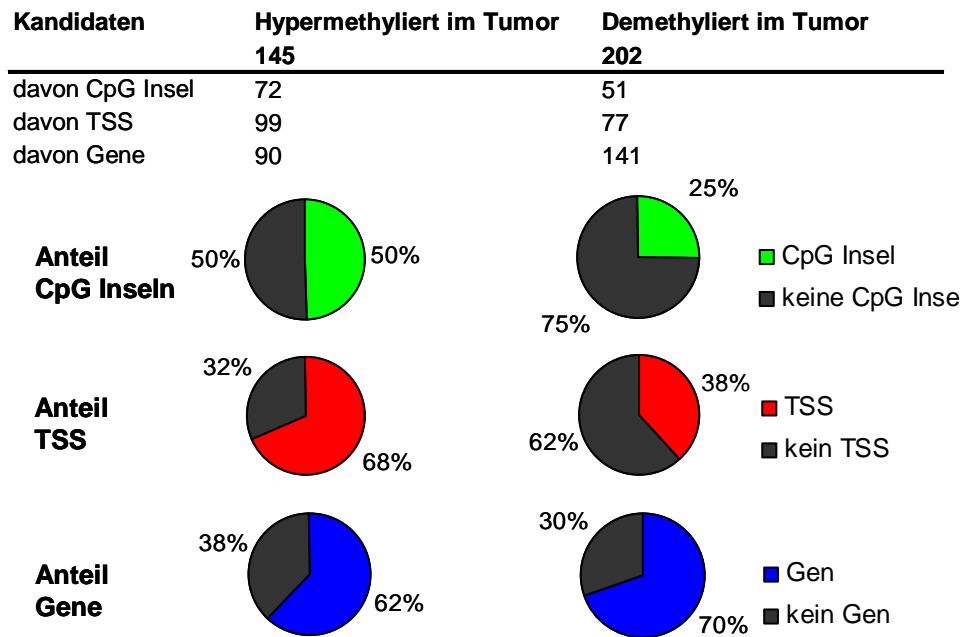


Abbildung 22: Anzahl der Kandidaten für Hypermethylierung im Tumor von in gesundem Brustgewebe unmethylierten Fragmenten und für Demethylierung im Tumor von in gesundem Brustgewebe methylierten Fragmenten. Weiterhin werden die Anteile der Fragmente, die mit einer CpG Insel oder einem Gen überlappen oder die in der Nähe ($\pm 2000\text{bp}$) eines TSS lokalisiert sind für diese beiden Kandidatengruppen dargestellt.

145 (30%) der im gesunden Gewebe unmethylierten Fragmente wurden als Kandidaten für eine Hypermethylierung und 202 (17%) der im gesunden Gewebe methylierten als Kandidaten für eine Demethylierung in den Tumorproben klassifiziert. In Abbildung 22 ist dargestellt, wie sich die Anteile von CpG-Inseln, Transkriptionsstarts und Genen in den beiden Kandidatengruppen in diesem Fall verhalten. Hier lässt sich ein ähnliches Verhalten erkennen, wie bei den zugrunde liegenden Fragmentgruppen in gesundem Brustgewebe. Die Gruppe der Kandidatenfragmente für eine Hypermethylierung im Tumor weist hohe Anteile von CpG-Inseln und Transkriptionsstarts und Genen auf während die Kandidatenfragmente für eine Demethylierung im Tumor weniger Überlappungen mit CpG-Inseln und Transkriptionsstart zeigen aber zu einem großen Teil in Genen liegen bzw. mit einem Gen überlappen.

In Tabelle 9 sind 30 Kandidatenfragmente mit den jeweils höchsten absoluten Werten der T-Statistik und dem Gennamen, wenn das Fragment mit einem Gen assoziiert ist, aufgelistet. Die vollständigen Listen werden im Anhang in Tabelle 15 und Tabelle 16 gezeigt. In den Tabellen wird ebenfalls gezeigt, wie viele der Tumorproben auf die jeweilige DNA-Methylierungsgruppe entfallen. Betrachtet man die Demethylierungs-Kandidaten, zeigt sich, dass bei diesen Kandidaten häufig relativ viele Tumorproben als nicht-methyliert eingeordnet wurden. Die DNA-Methylierung hat sich in diesem Fall also von in gesundem Gewebe methyliert zu unmethyliert in den Tumoren verändert. Betrachtet man die Gruppe der Kandidaten mit Methylierung in den Tumoren, sind die gefundenen Unterschiede kleiner, d.h. hier wurden die meisten Tumorproben als semi-methyliert eingeordnet. Daher sind auch die

Ergebnisse

Werte der T-Statistik bei den Kandidaten für Demethylierung höher als bei den Kandidaten für Methylierung. Da es sich bei den Kandidaten für Demethylierung um Fragmente handelt, welche, wie zu erwarten, hauptsächlich Introns und Exons sind, wird angenommen, dass ein großer Teil der gemessenen Veränderungen korrekt ist, allerdings bei dieser Gruppe mit einer erhöhten Zahl von falsch-positiven zu rechnen ist. Dieser Aspekt wird in Abschnitt 4.2.3 diskutiert.

Tabelle 9: Auswahl von Kandidatenfragmenten für Hyper- und Demethylierung in Brusttumoren.

Aufmethyliert in Brusttumoren					Demethyliert in Brusttumoren						
Fragment ID	T-Statistik	Proben pro Gruppe			Gen Name	Fragment ID	T-Statistik	Proben pro Gruppe			Gen Name
		methyliert	semi-	nicht-				methyliert	semi-	nicht-	
10953	7,0	7	13	8	DTX1	1262	-44,4	0	0	28	-
3944	6,9	12	10	6	-	9395	-18,2	0	0	28	ARHGAP8
4563	6,5	25	3	0	ZNF184	3533	-17,0	0	1	27	-
2376	6,1	26	1	1	C3orf64	10173	-16,3	0	12	16	GFER
2097	5,9	5	10	13	-	1691	-12,3	0	0	28	ADAM17
4233	5,6	8	18	2	-	10088	-10,9	0	0	28	-
4493	5,5	15	9	4	novel gene	9761	-10,7	0	0	28	-
11604	5,4	10	15	3	XPR1	13178	-10,1	15	13	0	-
11671	5,3	24	4	0	PCSK9	9997	-10,1	13	15	0	ADCY6
9713	5,3	4	24	0	ZADH1	5180	-9,8	1	3	24	ARPC4
7524	5,2	8	16	4	BCAT1	8493	-9,8	0	16	12	EBP
5969	4,8	0	25	3	Q9BVA6	6925	-9,4	0	8	20	CACNA2D3
4235	4,7	4	19	5	-	3206	-8,7	0	4	24	NDUFA9
656	4,5	11	12	5	-	8937	-8,7	2	26	0	VENTX2
11509	4,5	12	14	2	ANKS1B	689	-8,5	0	12	16	APP
12787	4,4	14	12	2	-	4349	-8,4	0	3	25	DTX4; OR5A2
2180	4,4	6	16	6	CTNNA3	4659	-8,3	0	14	14	CDCA5; SLC22A11; ZFPL1
10136	4,2	5	13	10	-	1994	-8,2	8	20	0	NDUFV3
10422	4,0	10	13	5	Q96HD6	7041	-7,6	18	10	0	NTRK1
6204	3,9	0	26	2	EPC1	8600	-7,1	3	25	0	novel gene
1674	3,9	8	20	0	C10orf197	13772	-6,9	1	16	11	PCID1
1678	3,9	9	19	0	ADAM18	11299	-6,7	13	13	2	-
4148	3,9	8	20	0	-	13118	-6,5	19	9	0	Q8N4O9; Q674X7
4416	3,8	8	19	1	-	10276	-6,4	1	20	7	USP49
4664	3,6	8	7	13	HOXA2	8038	-6,4	8	19	1	Q9BSM9
2043	3,6	2	13	13	PRDM12	8156	-6,2	9	6	13	-
9857	3,6	6	10	12	CDO1	798	-5,9	8	11	9	GNG4
1665	3,6	5	18	5	-	7251	-5,0	15	13	0	ALKBH3
12406	3,4	1	24	3	RHBG	2382	-4,9	5	12	11	Q8N9G5; ZNF135; ZNF606
2272	3,3	11	12	5	SALL3	13491	-4,7	0	4	24	SNX8

3.2.8. Unterschiede zwischen Chemotherapie-Ansprechenden und Nicht-Ansprechenden

Im folgenden Abschnitt werden die Ergebnisse für den Vergleich der Tumorproben von Patientinnen, die auf die Chemotherapie ansprachen (Ansprechende) mit denen die keine Reaktion auf die verwendete Chemotherapie zeigten (Nicht-Ansprechende), dargestellt. Auch bei diesem Vergleich wurden ausschliesslich solche Fragmente betrachtet die in gesundem Gewebe als unmethyliert oder als methyliert eingeordnet wurden. Als Filter für Unterschiede zwischen den Ansprechenden und Nicht-Ansprechenden Tumoren wurde die Differenz der Probenzahl, die als methyliert eingeordnet wurden und die Differenz der Probenzahlen die als unmethyliert eingeordnet wurden, verwendet (siehe Tabelle 10). Es wurden alle Fragmente als Kandidaten angesehen, bei denen der absolute Wert dieser Differenz ≥ 4 war, d.h. mindestens vier Tumorproben eine andere DNA-Methylierung zeigten als in der Vergleichsgruppe. Als weiteres Kriterium wurde wiederum der Wert der T-Statistik verwendet und alle Kandidaten verworfen, bei denen der absolute Wert ≤ 1 war.

In der Gruppe der im gesunden Gewebe unmethylierten Fragmente wurden 55 Kandidaten und in der Gruppe der im gesunden Gewebe methylierten Fragmente 85 Kandidaten gefunden. Die Anteile von CpG-Inseln, TSS und Genen in den beiden Kandidatengruppen sind in Abbildung 23 dargestellt und zeigen grundsätzlich das gleiche Verhalten wie in den zugrunde liegenden Fragmentgruppen unmethyliert und methyliert in gesundem Gewebe. Allerdings werden circa 10% mehr CpG-Inseln und TSS für die Kandidaten mit Hypermethylierung in den gesunden Geweben gefunden.

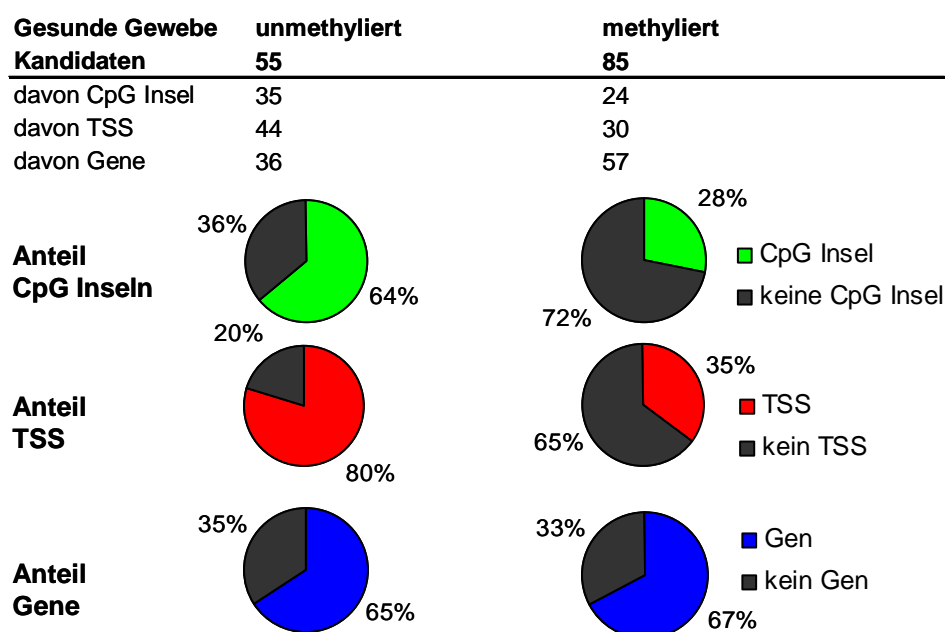


Abbildung 23: Anzahl der Fragmente, die als Kandidaten für das Ansprechen auf eine Chemotherapie gefunden wurden, aufgeteilt in die zwei Gruppen unmethyliert oder methyliert in gesundem Brustgewebe sowie der Anteil der Fragmente, die mit einer CpG-Insel oder einem Gen überlappen oder die in der Nähe ($\pm 2000bp$) eines TSS lokalisiert sind, für die beiden Kandidatengruppen

In Tabelle 10 sind jeweils die Kandidaten mit den 30 besten Werten der T-Statistik, zugehörigen Gennamen, wenn das Fragment mit einem Gen assoziiert war und den Differenzen der Proben pro DNA-Methylierungsgruppe aufgelistet. Die vollständigen Listen sind im Anhang als Tabelle 17 und Tabelle 18 zu finden.

Es ist auffällig, dass die Werte für die T-Statistik alle positiv sind. Da hier keine absoluten Werte angegeben sind, bedeutet dies, dass die Differenz der mittleren DNA-Methylierung von Ansprechenden und Nicht-Ansprechenden immer positiv ist, d.h. es werden hypermethylierte Tumorproben in der Gruppe der Ansprechenden gefunden, wenn das Fragment in den gesunden Gewebekontrollen keine DNA-Methylierung aufweist. In der Gruppe, der im gesunden Gewebe methylierten Fragmente, werden andererseits vermehrt demethylierte Tumorproben in der Gruppe der Nicht-Ansprechenden gefunden. Da dies auch bedeutet, dass in beiden Fällen die DNA-Methylierung in der Gruppe der Ansprechenden größer ist, könnte man hier wiederum einen Einfluss des Batcheffektes vermuten. Dies ist aber ausgeschlossen, da diese Proben nicht in getrennten Studien bearbeitet wurden und auch für die verschiedenen Arbeitsschritte auf eine Durchmischung der Proben zur Verhinderung eines Batcheffektes innerhalb der Studie geachtet wurde.

Wie erwartet sind die gefundenen Unterschiede kleiner als für den Vergleich aller Tumore mit den gesunden Kontrollen, was sich an den kleineren Werten der T-Statistik erkennen lässt. Diese ist allerdings auch dadurch beeinflusst, dass in diesem Fall zwei gleich große Gruppen von jeweils 14 Proben verglichen wurden und nicht 28 Tumorproben mit 3 gesunden Gewebekontrollen. Betrachtet man die Differenzen von Ansprechenden und Nicht-Ansprechenden, ist zuerkennen, dass der größte Teil der Kandidaten (74% bzw. 69%) sich in vier oder fünf Proben unterscheidet. Unterschiede in mehr als sieben Proben wurden kaum gefunden. Ausnahmen sind das Fragment 10156 mit einer Differenz von 9, und das Fragment 1322, welches sich in 11 Proben unterscheidet, sowie in der Gruppe der in den gesunden Kontrollen methylierten Fragmente die beiden Kandidaten 7295 und 12343, die sich beide in 8 Proben unterscheiden. Das Fragment 1322 hat gleichzeitig auch den höchsten Wert der T-Statistik, was es zu einem besonders viel versprechenden Kandidaten macht.

Tabelle 10: Jeweils 30 Kandidaten für das Ansprechen auf eine Chemotherapie bei Sortierung nach dem Wert der T-Statistik von Fragmenten, die in gesundem Gewebe unmethyliert oder methyliert sind.

unmethyliert im gesunden Gewebe					methyliert im gesunden Gewebe				
Fragment ID	T-Statistik	Differenz: Anspr. - Nicht-Anspr.		Gen Name	Fragment ID	T-Statistik	Differenz: Anspr. - Nicht-Anspr.		Gen Name
		methyliert	unmethyliert				methyliert	unmethyliert	
1322	4,4	3	-11	-	3757	3,9	6	-6	NAP1L2
1639	3,1	1	-6	ZNF266	8156	3,2	3	-7	-
5044	2,9	0	-4	CDC2L5	2382	3,1	5	-3	Q8N9G5; ZNF135/606
13370	2,8	0	-7	KPNA2	11069	3,0	5	-6	-
5319	2,7	0	-5	TTBK2	640	3,0	5	-3	-
2659	2,5	0	-5	FGFRL1	7295	2,9	8	-1	Wnt10B
10156	2,4	0	-9	novel gene; ZNF582	1126	2,8	5	-2	CDK5RAP2
13159	2,4	0	-4	SASS6; CCDC76	4575	2,8	4	-5	-
480	2,4	0	-4	HMG3	3712	2,9	0	-7	-
13569	2,3	0	-5	ROPN1L	9427	2,8	2	-4	LENG8
6401	2,3	0	-4	ZBTB2	798	2,8	2	-7	GNG4
3502	2,2	0	-4	SEMA3E	7126	2,8	6	-3	TBX19
1047	2,2	0	-6	COPS4	5968	2,7	7	-4	ST13
2403	2,2	-1	-5	-	10521	2,7	3	-5	USP13
5036	2,2	0	-5	STRA13; LRRC45	9032	2,7	1	-7	RPGR
11968	2,1	0	-6	-	6921	2,7	3	-5	C11orf95
8242	2,1	0	-4	PIGM	5039	2,6	6	0	-
13595	2,1	0	-5	MLL; Q9BRP9	6686	2,6	3	-4	PARP11
7029	2,1	0	-6	STARD4	10053	2,5	5	0	-
7250	2,0	0	-4	ENDOGL1	7656	2,5	5	-2	KIAA1958
12749	2,0	1	-5	-	10380	2,5	6	0	TLE4
4402	2,0	0	-5	-	11219	2,5	4	-2	-
12799	2,0	1	-6	GEMIN7	11003	2,4	1	-7	-
3252	2,0	0	-5	CSTP1	361	2,4	5	-2	ST3GAL5
13196	1,9	0	-6	UTX	5153	2,3	0	-5	-
3615	1,9	1	-6	NGB	1666	2,3	5	0	FXR1
4235	1,9	2	-5	-	12156	2,3	2	-7	C20orf39
5035	1,8	0	-4	ABTB2	3833	2,3	0	-6	C21orf104
5485	1,8	0	-7	USP3	12343	2,2	8	0	SCN10A
4885	1,8	0	-4	NM_198564	8683	2,2	6	-3	RARA

3.2.9. Bisulfitsequenzierung ausgewählter Kandidaten

12 Kandidaten, der in den vorangegangenen Abschnitten beschriebenen Vergleiche, wurden direkt bisulfitsequenziert, d.h. die PCR-Produkte bisulfitierter DNA wurden direkt in die Sequenzierreaktion eingesetzt und nicht vorher subkloniert. Die Auswertung der Sequenzierdaten erfolgte mit Hilfe einer speziellen Software, die für die direkte Bisulfitsequenzierung entwickelt wurde (siehe Abschnitt 2.3.4). Da es für eine spätere Anwendung eines eventuellen Markers besonders vorteilhaft ist, wenn keine DNA-Methylierung im gesunden Gewebe vorliegt, wurden ausschließlich solche Kandidaten ausgewählt. Die Auswahl erfolgte nicht nach einheitlichen Kriterien sondern nach verschiedenen Gesichtspunkten, d.h. es wurden z.B. sowohl Kandidaten mit einem guten Wert der T-Statistik als auch solche, die mit interessanten Genen assoziiert sind oder die in einer CpG-Insel liegen, ausgewählt. Andererseits wurden auch Fragmente mit Werten in Grenzbereichen untersucht, um für spätere Studien Anhaltswerte für die Auswahl von Grenzwerten zu bekommen.

Tabelle 11: Liste ausgewählter Fragmente für die Bisulfitsequenzierung

Fragment ID	Rang T-Statistik Tumormarker	Rang T-Statistik Ansprechen auf Chemotherapie	CpG Insel	TSS	Gen	Gen Name
4563	3	-	-	-	+	ZNF184
4493	7	-	-	+	+	novel trans
12787	16	-	+	-	-	no gene
4664	25	-	-	+	+	HOXA2
9857	27	-	+	+	+	CDO1
7951	66	-	+	+	+	ERCC5
1322	51	1	-	-	-	no gene
3615	137	26	+	+	+	NGB
4235	13	27	+	-	-	no gene
1589	85	39	+	+	-	ZNF145
13165	83	40	-	+	-	Q96HQ8
10422	19	50	+	+	-	Q96HD6 (PAX9)

In Tabelle 11 werden die bisulfitsequenzierten Kandidaten und die dazugehörigen Ränge der T-Statistik für den Vergleich gesunder Gewebe mit allen Tumorproben (Tumormarker) und den Vergleich der Tumorproben untereinander, aufgeteilt nach Chemotherapie-Ansprechenden und Nicht-Ansprechenden (Ansprechen auf Chemotherapie) gezeigt. Weiterhin sind angegeben, ob sich der sequenzierte Kandidat in der Nähe eines TSS befindet oder ob er mit einem Gen oder mit einer CpG-Insel überlappt sowie der zu einem gefundenen TSS oder Gen gehörige Name. Zunächst wurden 24 der DMH-analysierten Tumorproben bisulfit-sequenziert. In den Abbildungen 24 und 25 sind die Ergebnisse für die 24 Tumorproben, die gesunden Gewebekontrollen einschließlich einer PBL und die DNA-Methylierungskontrollen 0%Phi und 100%Pro im Vergleich zu den in der DMH-Analyse erhaltenen DNA-Methylierungswerten dargestellt.

Die Auswahl eines Sequenzabschnittes für die direkte Bisulfitsequenzierung unterliegt verschiedenen Kriterien. Unter anderem ist es notwendig, Primer auszuwählen, die kein CpG enthalten und der sequenzierte Abschnitt sollte eine sequenzierbare Länge von circa 400bp nicht überschreiten. Daher ist es nicht in jedem Fall möglich, eine vollständige Abdeckung eines DMH-Fragmentes zu erreichen. Daher variiert die Anzahl der sequenzierten Schnittstellen eines DMH-Fragmentes und z.T. konnten keine Schnittstelle sondern nur die angrenzenden CpGs sequenziert werden. Die beste Korrelation zwischen Bisulfitsequenzierung und DMH wäre zu erwarten, wenn man jeweils den Wert der Schnittstelle mit der geringsten DNA-Methylierung verwendet, da dieser für die Höhe des Signales beim DMH verantwortlich ist. Daher wurde für die Korrelationen, entweder das Minimum der sequenzierten Schnittstellen oder, wenn keine Schnittstelle mit guter Qualität sequenziert werden konnte, der Mittelwerte aller CpGs eines Fragmentes verwendet. In Tabelle 12 ist aufgelistet, wie viele Schnittstellen für das jeweilige Fragment mit guter Qualität sequenziert wurden, wie viele das entsprechende DMH-Fragment ursprünglich enthielt und welcher Wert für die Korrelation verwendet wurde. Daran ist zu erkennen, wie unterschiedlich die hier gefundenen Sequenzabschnitte sind und wie schwierig daher eine einheitliche Vorgehensweise ist.

Für einige Fragmente ist auffällig, dass Proben DNA-Methylierungswerte von mehr als 200% in der DMH-Analyse aufwiesen. Dies kann bedeuten, dass die Werte der 0% und 100% Kontrollen, die für die Normierung der DMH-Studien verwendet wurden, bei diesem Fragment nicht funktioniert haben, d.h. ein zu hohes oder zu niedriges Hybridisierungssignal zeigten. Handelt es sich bei den Proben, die sehr hohe DNA-Methylierungswerte zeigten, allerdings um Tumormproben, kann es sich auch um eine Amplifikation bei gleichzeitiger DNA-Methylierung der entsprechenden Region handeln, d.h. dieses Fragment liegt möglicherweise in den Tumormproben wirklich mit mehr Kopien vor als in gesundem Gewebe und zeigt daher ein höheres Signal als die 100% methylierte Kontrolle aus PBL.

In Abbildung 24 sind die ersten sechs ausgewählten Kandidaten dargestellt, welche ausschließlich als Tumormarker Kandidaten gefunden wurden. Vergleicht man die Werte zwischen DMH und Bisulfitsequenzierung, zeigen die Fragmente 4563 und 7951 keine, die Fragmente 12787, 4493 und 4664 für einige Proben und das Fragment 9857 weitgehend eine gute Übereinstimmung. In Tabelle 12 ist aufgelistet, wie viele Schnittstellen des jeweiligen DMH-Fragmentes sequenziert und analysiert wurden. Fragment 7951 enthält zwei Schnittstellen und eine wurde sequenziert. Diese Schnittstelle ist für alle Fragmente bis auf zwei Ausnahmen nicht methyliert, d.h. das Fragment hätte in jedem Fall geschnitten werden können. Für das Fragment 4563 wurden vier von fünf Schnittstellen sequenziert. Allerdings sind bei diesem Fragment DNA-Methylierungswerte bis zu 200% gemessen worden, was auf ein Problem bei der DMH-Analyse hinweist. Um eine Amplifikation kann es sich nicht handeln, da das Fragment nicht methyliert ist. D.h. für die Fragmente 7951 und 4563 sind die im DMH-Experiment gefundenen DNA-Methylierungswerte als nicht bestätigt einzustufen.

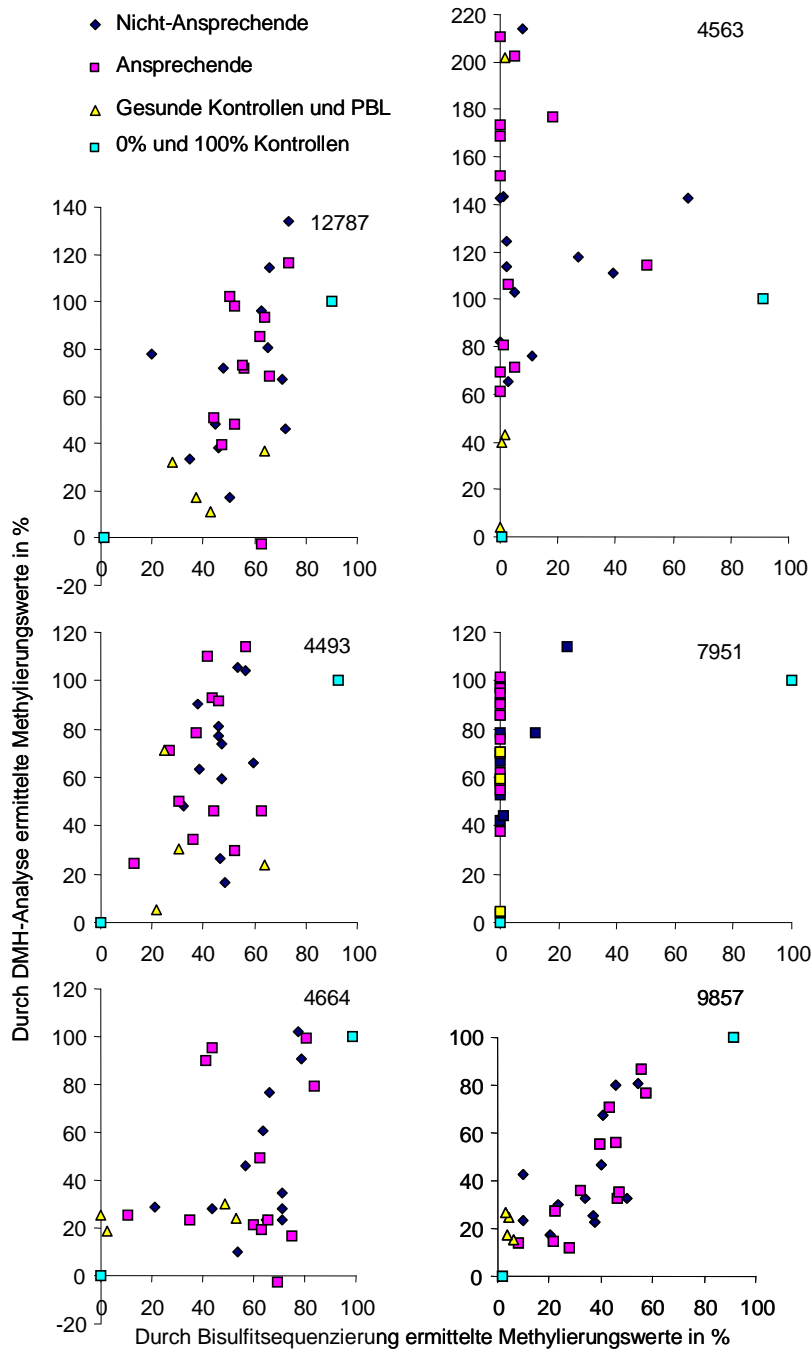


Abbildung 24: DMH-Methylierungswerte im Vergleich zu Bisulfitsequenzierungsmethylierungswerten für sechs Tumormarker-Kandidaten.

Bei den Fragmenten mit mittleren Korrelationen wurden im Fall des Fragmentes 12787 zwei von neun und im Fall der Fragmente 4493 und 4664 keine der sechs bzw. neun Schnittstellen sequenziert. Außerdem zeigen diese Fragmente häufig höhere DNA-Methylierungswerte in der Bisulfitsequenzierung als in der DMH-Analyse, d.h. in diesen Fällen ist es gut möglich, dass bei Sequenzierung der entsprechenden Schnittstellen noch eine Übereinstimmung gefunden werden könnte.

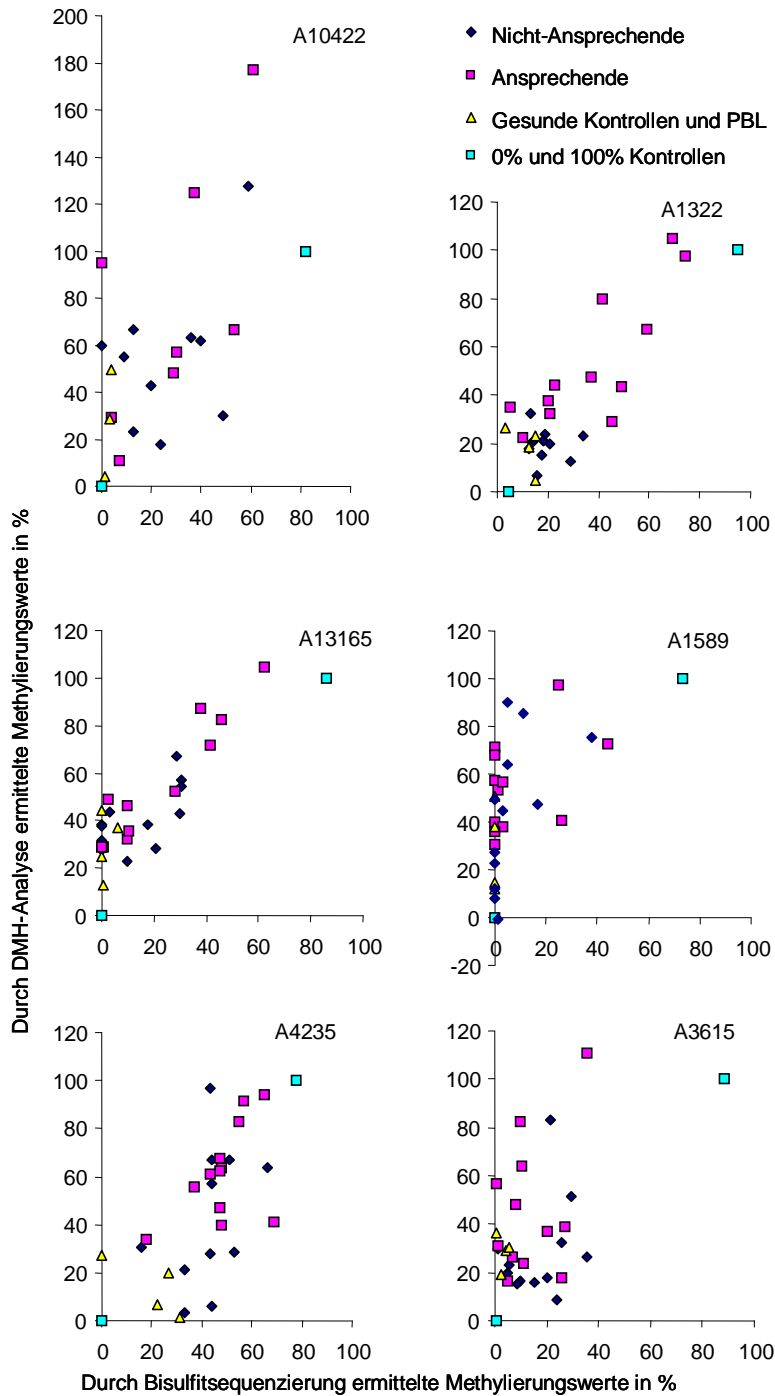


Abbildung 25: DMH-Methylierungswerte im Vergleich zu Bisulfitsequenzierungsmethylierungswerten für sechs Tumormarker-Kandidaten, die ebenfalls das Ansprechen auf die Chemotherapie vorhersagen könnten.

In Abbildung 25 sind die Ergebnisse der Kandidaten für einen Unterschied zwischen Ansprechenden und Nicht-Ansprechenden dargestellt. Diese Fragmente sind ebenfalls Kandidaten für Tumormarker, da auch der Unterschied zwischen den gesunden Kontrollen und den Tumoren groß ist. Die Fragmente 1322, 13165 und 4235 zeigen eine gute Übereinstimmung zwischen den DMH-Methylierungswerten und den durch Bisulfitsequenzierung ermittelten.

Bei den Fragmenten 10422, 1589 und 3651 konnte nur für einen Teil der Proben eine Übereinstimmung gefunden werden. Betrachtet man die Fragmente mit weniger guten Korrelationen, ist wiederum zu beobachten, dass diese Fragmente häufig höhere DNA-Methylierungswerte in der Sequenzierung aufweisen als in der DMH-Analyse gemessen wurden und wenige bzw. keine der Schnittstellen sequenziert wurden. Daher ist es auch hier gut möglich, dass eine Korrelation der beiden Methoden eintreten würde, wenn die entsprechenden Schnittstellen sequenziert würden, zumal in den meisten Fällen die Sequenzierungsergebnisse die richtige Tendenz aufweisen.

Betrachtet man die Auftrennung der Tumorproben im Vergleich zu den gesunden Kontrollen in Richtung der X-Achse, kann man abschätzen, wie gut das jeweilige sequenzierte Fragment als Tumormarker geeignet ist. Besonders gut ist ein Verhalten wie es z.B. das Fragment 9857 in Abbildung 24 zeigt, bei dem die gesunden Kontrollen einheitlich sehr geringe und die Tumorproben hauptsächlich höhere DNA-Methylierungswerte aufweisen. Würde man z.B. einen Schwellenwert von 15% DNA-Methylierung festsetzen, würden lediglich drei der analysierten Tumorproben eine zu geringe DNA-Methylierung haben, um als Tumor erkannt zu werden. Die Fragmente 12787, 4493 und 4664 zeigen auch eine Abtrennung der gesunden Kontrollen von den Tumoren, allerdings weisen die Fragmente 12787 und 4493 in einer und das Fragment 4664 in zwei gesunden Kontrollen höhere DNA-Methylierungswerte auf. Das Fragment 4563 hingegen, zeigt zwar eine schlechte Übereinstimmung der Sequenzierung mit den DMH-Werten, das sequenzierte Fragment wäre aber trotzdem für einige Tumorproben als Tumormarker verwendbar. Ausschließlich das Fragment 7951 zeigt auch in den Tumorproben sehr wenig DNA-Methylierung und kann daher nicht als Kandidat gesehen werden.

Die sechs in Abbildung 25 dargestellten Fragmente sind ebenfalls als Tumormarker-Kandidaten gefunden worden und die sequenzierten Fragmente zeigen ein entsprechendes Verhalten. Betrachtet man die Fragmente in Abbildung 25 hinsichtlich der Auftrennung von Ansprechenden und Nicht-Ansprechenden in Richtung der X-Achse, kann man für die beiden Fragmente 1322 und 13165 eine klare Abtrennung einiger Ansprechenden von den restlichen Proben erkennen. Bei allen anderen Fragmenten kommt es insgesamt zu einer Durchmischung der beiden Gruppen, d.h. es ist keine Auftrennung der beiden Gruppen zu beobachten.

In Tabelle 12 sind die Ergebnisse für alle 12 bisulfitsequenzierten Kandidaten zusammengefasst. Daneben ist aufgelistet, für welche sechs Kandidatenfragmente im Anschluss die weiteren vier noch DMH analysierten Proben und 34 unabhängige Proben (siehe Anhang: Tabelle 14) sequenziert wurden, um festzustellen, ob sich das bisher beobachtete Verhalten bestätigen lässt. Bei der Bewertung der Daten muss man beachten, dass es sich bei der hier verwendeten direkten Bisulfitsequenzierung, nicht um eine kalibrierte Messung handelt, d.h. die bestimmten Methylierungswerte sind nicht abgesichert. Weiterhin wurden die p-values nicht für das Testen mehrere Proben korrigiert und sind daher lediglich als weitere Hinweise

für die Richtigkeit von Markern zu verstehen aber nicht als Bestätigung. Hierfür wären wesentlich umfangreichere Experimente nötig, die insbesondere eine wesentlich größere Probenzahl analysieren. Die Ergebnisse der Sequenzierung werden in Abbildung 26 gezeigt. Alle Fragmente zeigen in einer bestimmten Zahl der Tumorproben eine erhöhte DNA-Methylierung im Vergleich zu den gesunden Proben, d.h. sie könnten als Tumormarker verwendet werden. Das Fragment 4235 zeigt in den gesunden Kontrollen eine DNA-Methylierung von rund 20 % allerdings zeigen die Tumorproben im Vergleich zu den anderen Fragmenten auch höhere DNA-Methylierungswerte von bis zu 90%. Ähnlich verhält es sich beim Fragment 1322. In den gesunden Kontrollen wird eine DNA-Methylierung von circa 10% und in den Tumoren bis zu 80% gemessen. Die anderen vier Fragmente zeigen in den gesunden Geweben annähernd 0% DNA-Methylierung. Während bei den Fragmenten 3615, 1589 und 13165 circa die Hälfte der Tumorproben eine deutlich höhere DNA-Methylierung als die gesunden Gewebe aufwiesen sind es für das Fragment 9857 circa 75%, d.h. Fragment 9857 ist ein besonders guter Tumormarker.

Tabelle 12: Anzahl analysierter Proben und Schnittstellen der bisulfitsequenzierten Fragmente und Tauglichkeit als Tumormarker bzw. Marker für das Ansprechen auf eine Chemotherapie

Fragment ID	Analysierte Proben	Schnittstellen		Verwendet f. Korrelation	Tumormarker Kandidat	Ansprechen auf Chemotherapie Marker Kandidat
		DMH	Bisulfit			
4563	24	5	4	Minimum	Ja	Nein
4493	24	6	0	Mittelwert	Ja	Nein
12787	24	9	2	Minimum	Ja	Nein
4664	24	9	0	Mittelwert	Ja	Nein
9857	62	5	0	Mittelwert	Ja	Nein
7951	24	2	1	Minimum	Nein	Nein
1322	62	3	0	Mittelwert	Ja	Ja
3615	62	8	0	Mittelwert	Ja	Nein
4235	62	7	6	Mittelwert	Ja	Nein
1589	62	3	2	Minimum	Ja	Nein
13165	62	3	0	Mittelwert	Ja	Ja
10422	24	3	2	Minimum	Ja	Nein

Für den Vergleich von Ansprechenden und Nicht-Ansprechenden sind die Ergebnisse (p-values) des T-Tests für den Vergleich von Ansprechenden und Nicht-Ansprechenden angegeben. Für die Berechnung wurden (siehe Tabelle 12) entweder das Minimum der sequenzierten Schnittstellen oder der Mittelwert aller sequenzierten CpGs verwendet (siehe oben). Wie auf Grund der bereits erwähnten Komplexität der Fragestellung zu erwarten, zeigen fünf der Fragmente keinen statistisch relevanten Unterschied. Für das Fragment 1322 konnte jedoch ein signifikanter Unterschied mit einem p-value von 0,03 in der Bisulfitsequenzierung gefunden werden. Die ebenfalls sequenzierten Proben, die als unveränderte Erkrankung nicht einer der beiden Gruppen zugeordnet werden konnten, zeigen keine großen Unterschiede zu den anderen Gruppen. Allerdings bestätigen sie grundsätzlich die gefundenen Unterschiede von Tumor- und gesundem Gewebe und somit ist dieses Fragment auch sehr gut als Tumormarke zu verwenden.

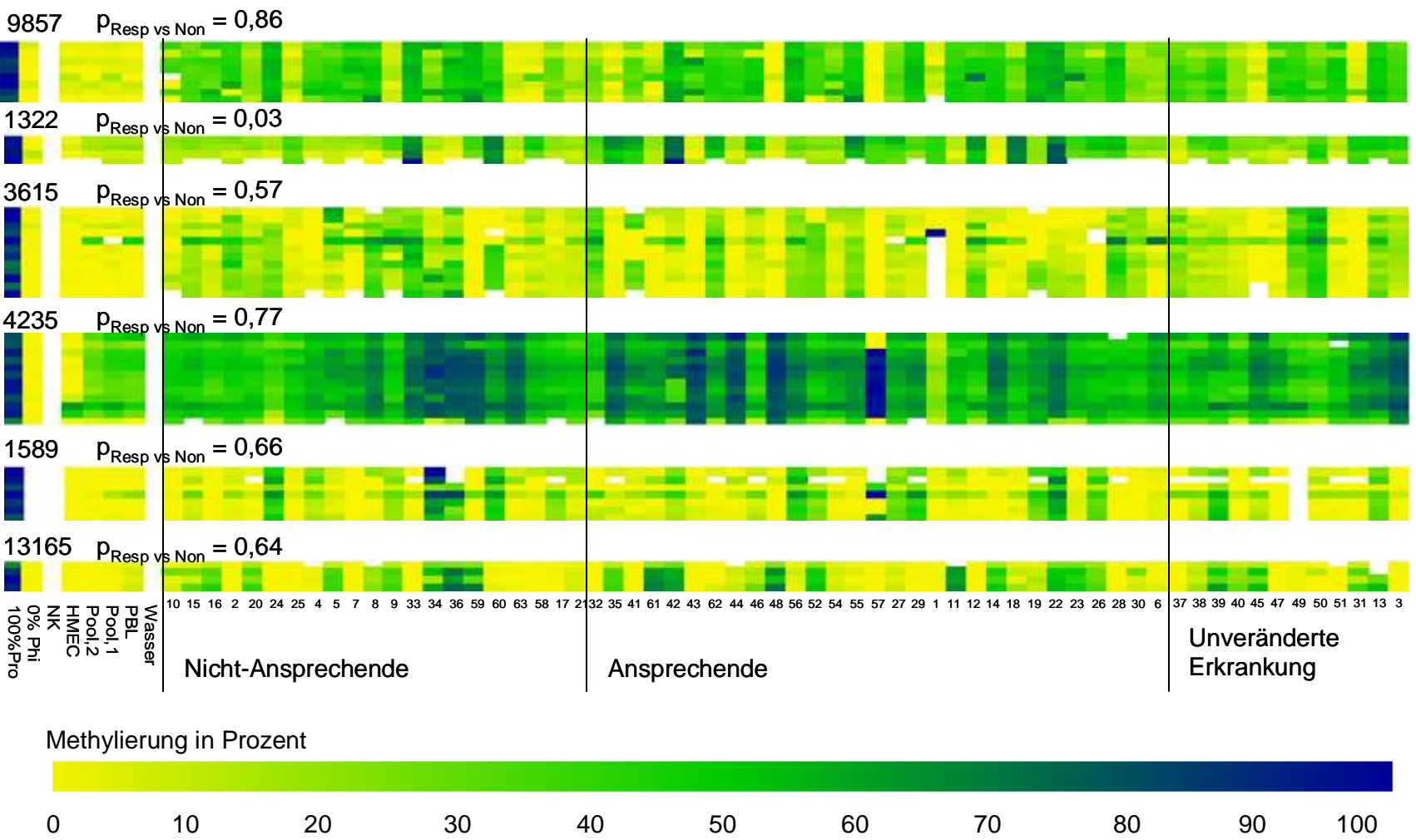


Abbildung 26: Methylierung von sechs Marker-Kandidaten in 62 Tumoren und gesunden Kontrollen. Jede Zeile innerhalb eines Fragements stellt ein CpG und jede Spalte eine analysierte Probe dar (für die Probennummern siehe Tabelle 13 und Tabelle 14). Die Methylierungswerte sind durch eine Farbskala von gelb (0%) zu blau (100%) kodiert. Weiterhin sind die p-values für den Vergleich von Ansprechenden und Nicht-Ansprechenden für jedes Fragment aufgelistet.

4. Diskussion

4.1. Genomweite Analyse von DNA-Methylierungsunterschieden

Um neue Kandidaten für Früherkennung, Krebsdiagnostik, prädiktive und prognostische Marker möglichst effizient zu finden, ist eine genomweite Methode erstrebenswert. In dieser Arbeit wurde die bereits existierende DMH Methode ausgewählt und grundlegend verbessert. Dies wurde durch Optimierung der verwendeten Enzyme, durch Veränderungen in den einzelnen Arbeitsschritten und durch die Einführung eines spezifisch entworfenen Oligonukleotid-Mikroarrays erreicht. Anschließend konnte gezeigt werden, dass es nach diesen Verbesserungen möglich ist, nicht nur nach DNA-Methylierungsunterschieden zu suchen sondern auch eine quantitative Auswertung der detektierten Signale vorzunehmen. Durch die Etablierung und Optimierung des Designs der Oligonukleotid-Mikroarrays konnte ein Mikroarray zur Detektion von circa 50.000 verschiedenen Fragmenten entwickelt werden.

4.1.1. Vergleich der DMH-Methode mit anderen Technologien

In Abschnitt 1.2.2 sind die bisher verwendeten Methoden für die genomweite Analyse von DNA-Methylierungsunterschieden beschrieben. Ein wesentlicher Aspekt bei der Auswahl einer geeigneten Methode, war die zu untersuchende Fragestellung. Eine komplexe Fragestellung, wie die Suche nach Markern für das Ansprechen auf eine Chemotherapie, erforderte eine Methode, die den Vergleich vieler Proben möglich macht und die parallele Analyse einer Vielzahl von Sequenzabschnitten also eine möglichst hohe Abdeckung des Genoms gewährleistet. Um dieses Kriterium zu erfüllen, bieten sich besonders Methoden an, bei denen Mikroarrays für die Detektion verwendet werden. In 1.2.2.2 sind die bisher entwickelten Methoden die Mikroarrays verwenden beschrieben. Soll, wie in dieser Arbeit, Tumormaterial untersucht werden, ist die DNA-Menge häufig ein limitierender Faktor, da nur eine begrenzte Menge Gewebe vorliegt. Eine Anforderung an die Methode war daher, dass nur wenig DNA benötigt wird. Aus diesem Grund waren die Methoden Tompa et al. mit 20-30 μg DNA, Lippman et al. mit mindestens 10 μg DNA und MeDIP mit 4 μg DNA nicht geeignet. Von den verbleibenden fünf Methoden arbeiten alle mit methylierungssensitiven Enzymen und einer adapter-unterstützten PCR. Lediglich die Methode von Ching et al. verwendet eine andere Anreicherungs-methode. Da also das grundlegende Prinzip aller Methoden das gleiche ist, wurde die DMH – Methode gewählt, da diese bereits in mehreren Studien gute Daten erzeugte [50, 51, 107, 108, 123, 124] und durch die Epigenomics AG einlizensiert wurde. Einige der neueren Methoden waren zu Beginn dieser Arbeit noch nicht publiziert. Sie bieten jedoch auch keine bemerkenswerten Vorteile gegenüber der DMH – Technologie oder der hier vorgestellten verbesserten Variante.

4.1.2. Oligonukleotid-Mikroarrays

Eine wesentliche Verbesserung erfährt die Methode durch die Einführung eines Oligonukleotid-Mikroarrays von Affymetrix, der speziell für diese Anwendung entwickelt wurde. Die meisten der bereits beschriebenen Methoden verwenden Glasslides auf die entweder ganze Klone, BACs oder auch Oligonukleotide gespottet sind. Darauf werden meist zwei Proben, mit den Farbstoffen Cy3 bzw. Cy5 markiert, hybridisiert. Dies kann zunächst für den direkten Vergleich zweier Proben, bei dem keine Normalisierung der Daten nötig ist, verwendet werden. Der Vergleich vieler Proben parallel wird möglich, wenn eine Referenzprobe, die der Normalisierung dient, als Vergleich hybridisiert wird. Dies bedeutet jedoch einen erhöhten Aufwand. Weiterhin sollte bei diesen Experimenten ein Kreuztest der beiden Farbstoffe durchgeführt werden, da sie sich nicht gleich verhalten. Dies ist alles, sowohl die Laborarbeit als auch die Kosten betreffend, sehr aufwendig.

Bei dem ursprünglich für die DMH Methode verwendeten Array werden PCR Produkte einer CpG-Insel-Bibliothek direkt auf Glasslides gespottet, d.h. es ist zunächst notwendig, die gesamte Bibliothek mittels PCR zu amplifizieren. Das Spotting ist auch mit einem dafür verwendeten Roboter ein zeitaufwendiger Prozess, der viel Laborarbeit benötigt. Die Qualität der Mikroarrays, die hauptsächlich durch die Homogenität der gespoteten DNA-Menge und die Form der Spots gekennzeichnet ist, kann mitunter stark schwanken. Ein anderes Qualitätsmerkmal ist die Vorbereitung und Beschichtung der Glasslides. Hierbei kann es zu Verunreinigungen der Oberfläche kommen. Daneben sind die so hergestellten Arrays nur begrenzt lagerfähig (Persönliche Erfahrung, Daten nicht gezeigt). Ein weiteres Problem ist die Pflege der Bibliothek. Es kann zu Kreuzkontaminationen kommen, d.h. dass nicht immer das Insert eines einzelnen Klones gespottet wird. Genauso ist es möglich, dass ein Klon zu stark verdünnt wird und nicht mehr wächst. Die Bibliothek stellt eine Auswahl von Klonen dar, die über eine methylierte DNA bindende Säule hergestellt wurde [116], d.h. es gibt redundante Klone. Neben diesen technischen Nachteilen, ist die Bibliothek auch inhaltlich limitiert, da sie nur CpG-Inseln enthält.

Dem gegenüber bieten Oligonukleotid-Mikroarrays von Affymetrix einen hohen qualitativen Standard. Auf die Arrays werden nur einzelne Proben hybridisiert, welche insbesondere durch die in dieser Arbeit entwickelte Normalisierung untereinander und auch zwischen verschiedenen Experimenten verglichen werden können. In dieser Arbeit wurde ein relativ kleines Arrayformat verwendet, um die Designkriterien zu optimieren, welches jedoch bereits die Analyse von 3780 Fragmenten möglich machte. Die abgeleiteten Designkriterien machten die Entwicklung eines Arrays für die Analyse von 50 000 Fragmenten möglich, welcher bereits erfolgreich eingesetzt wurde [125].

4.1.3. Verbesserung des Protokolls und der Datenauswertung

Bei der Etablierung der DMH-Methode wurden verschiedene Veränderungen im Protokoll vorgenommen. Ein besonders wichtiges Ergebnis war die Senkung der benötigten DNA-Menge von 2000 ng. Mit den beschriebenen Veränderungen des Protokolls ist es möglich, mit 1000 ng DNA zu beginnen und trotzdem mindestens zwei Hybridisierungsproben aus einer Probe herzustellen. Mittlerweile wurden auch Experimente mit 500 ng DNA erfolgreich durchgeführt (Daten nicht gezeigt). Die Senkung der DNA-Menge beruht hauptsächlich auf der Veränderung des PCR Protokolls. Die PCR ist von großer Bedeutung, da es an dieser Stelle zur eigentlichen Anreicherung der methylierter Fragmente kommt. Während nicht-methylierte Fragmente geschnitten werden und nur linear amplifiziert werden können, sind methylierte Fragmente intakt und sollten daher wesentlich stärker amplifiziert werden. Es ist allerdings anzunehmen, dass auch eine erhöhte Effektivität der Ligation eine Rolle spielte. Diese Vermutung liegt nah, da bei der ursprünglichen Methode mit circa 400 ng DNA circa 80-mal mehr DNA in eine PCR eingesetzt werden (siehe Tabelle 3). Betrachtet man dazu die Tatsache, dass bei der Methode circa 8 Millionen Fragmente pro Genom mit dem Adapter verknüpft worden sein sollten, wäre nach einer effizienten Ligation mit einem hohen Überschuss an Zielsequenzen für den Primer zu rechnen. Weiterhin ist zu vermuten, dass bei der ursprünglichen PCR der Amplifikationsfaktor nicht größer als 50 ist, d.h. der zu erwartende Dynamikumfang ist wesentlich kleiner als nach den hier beschriebenen Veränderungen, die einen Amplifikationsfaktor von 1600 ermöglichen (siehe Tabelle 3). Womit das verbesserte Protokoll nicht nur zur Senkung der benötigten Ausgangsmenge sondern auch zur Möglichkeit einer quantitativen Auswertung entscheidend beiträgt.

4.1.4. Verbesserung der Datenauswertung

Ein anderer wesentlicher Schritt für eine quantitative Auswertung war der Ausgleich des durch die methylierungssensitive Restriktion entstehenden Komplexitätsunterschiedes durch die Basislinienverschiebung. Hierfür wurden Kontrollfragmente, die keine Schnittstelle der methylierungssensitiven Enzyme enthalten, verwendet. Auf deren Basis kann nun dem Hybridisierungssignal einer einzelnen Probe, eine ihrer Komplexität entsprechende Basislinie zugeordnet werden. Bei den meisten Hybridisierungen komplexer genomischer DNA, wie z.B. für die SNP-Analyse, die Analyse von Kopienzahlen etc. geht man davon aus, dass die Gesamtverteilung der Signalintensitäten verschiedener Proben gleich sein sollte und verwendet Normalisierungsmethoden, die dieses herstellen. Die Basislinienverschiebung kann auch das Gegenteil bewirken, nämlich dass ähnliche Verteilungen gegeneinander verschoben werden und so Unterschiede entstehen, die vorher nicht sichtbar waren. Daher ist anzunehmen, dass die Möglichkeit einer quantitativen Auswertung essentiell von der Einführung dieser Normalisierung abhängig ist.

4.1.5. Einfluss genomischer Veränderungen

Eine Eigenschaft der Hypermethylierung in Krebs ist, dass sie in Regionen gefunden wird, die häufig auch Veränderungen in den Kopienzahlen der Gene zeigen. Dies führte zur Erweiterung von Knudsons Two-Hit Hypothese, d.h. in diesem Fall wird z.B. ein Allel durch eine Deletion und das andere durch DNA-Methylierung ausgeschaltet. Damit ein Gen überexprimiert wird, kann es amplifiziert werden muss aber gleichzeitig auch demethyliert vorliegen. Um genomische Veränderungen von DNA-Methylierungsunterschieden abzugrenzen, wäre eine CGH-Analyse der verwendeten Proben ein gutes Mittel. Bei der hier beschriebenen Methode könnte das mit dem gleichen Array durchgeführt werden, indem man die Proben vor dem Schritt der methylierungssensitiven Restriktion aufteilt und ebenfalls eine Hybridisierungsprobe von dem ungeschnittenen Teil herstellt und hybridisiert. Der Vergleich der nicht-methylierungssensitiv geschnittenen Proben sollte Kopienzahlunterschiede zeigen. Grundsätzlich besteht aber, auch wenn ein Kopienzahlunterschied gefunden wird, immer die Möglichkeit, dass gleichzeitig ein DNA-Methylierungsunterschied vorliegt. Gefundene DNA-Methylierungsunterschiede lassen sich daher nur durch eine unabhängige DNA-Methylierungsanalyse, wie der Bisulfidsequenzierung ausschließen oder bestätigen.

Eine andere Quelle für fälschlich detektierte DNA-Methylierungsunterschiede sind SNPs in den entsprechenden Schnittstellen. Sind mehrere Schnittstellen in einem Fragment enthalten, sinkt allerdings die Wahrscheinlichkeit, dass sich ein gefundener Unterschied durch SNPs erklären lässt. Hauptsächlich von Interesse ist dieser Einfluss beim Vergleich der Tumorproben untereinander. SNPs sind individuelle Variationen, d.h. sie könnten für einzelne Unterschiede zwischen den Ansprechenden und Nicht-Ansprechenden verantwortlich sein, da hier die Zahl der Proben, die den Unterschied aufweisen, nicht sehr groß ist. Unterscheiden sich viele Proben, sinkt die Wahrscheinlichkeit, dass der Unterschied bei allen auf einem SNP beruht. Geht man davon aus, dass es sich bei dem SNP um eine individuelle Mutation handelt, die nicht erst im Tumorgewebe als Mutation entstanden ist, sollte der gefundene Unterschied groß sein, d.h. alle Unterschiede, die wesentlich kleiner als eine Differenz von 100% sind, sind vermutlich nicht auf SNPs zurückzuführen. Genomische Veränderungen sind trotzdem als Erklärung für gefundene Unterschiede nicht auszuschließen und können daher für die fälschliche Detektion von DNA-Methylierungsunterschieden verantwortlich sein.

4.1.6. Datenqualität und quantitative Auswertung

Um die Qualität der neuen Methode zu evaluieren wurden Daten für PBL und die Brustkrebszelllinie MDA-MB-231 erhoben. Durch den Vergleich konnte zunächst gezeigt werden, dass die Methode eine hohe Reproduzierbarkeit von Markerkandidaten aufweist (siehe Abschnitt 3.1.5.2). Um festzustellen wie gut die Daten DNA-Methylierungswerte darstellen wurde eine Auswahl von Fragmenten durch Bisulfitsequenzierung als unabhängige Methode analysiert. Grundsätzlich wäre es am besten zufällig gewählt Fragmente zu wählen, da aber bei einer zufälligen Auswahl zum größten Teil Fragmente, die keine Unterschied aufweisen, ausgewählt würden, mussten zunächst Gruppen definiert werden aus denen dann zufällig Fragmente ausgesucht wurden. Dadurch wurde das gesamte Spektrum möglicher DNA-Methylierungsunterschiede abgedeckt. Beim Vergleich der Daten wurde eine hohe Korrelation von 71% zwischen den DMH-Signalunterschieden und den Unterschieden der Methylierungswerte der Bisulfitsequenzierung gefunden (siehe Abschnitt 3.1.5.3). Dieser Wert wurde als hoch eingestuft, da die bisulfitsequenzierten Fragmente in den wenigsten Fällen das gesamte DMH Fragment abdecken. Aus dieser guten Korrelation wurde geschlussfolgert, dass die Daten DNA-Methylierungslevel darstellen und somit eine Übersetzung der Signalintensitäten in einen DNA-Methylierungsprozentsatz möglich sein sollte. Dies wurde dann in der folgenden Anwendung der Methode zur Analyse von Brustkrebsproben erfolgreich umgesetzt. Die Notwendigkeit und Vorteile einer quantitativen Einordnung der Signale wird dabei besonders deutlich, da es große funktionale Unterschiede zwischen Sequenzen, die im Krebs hypermethyliert, semimethyliert oder demethyliert im Vergleich zu gesundem Gewebe vorliegen, gibt. Des Weiteren ist diese Information sehr wichtig für die Auswahl von Kandidaten für die weitere Analyse und Entwicklung klinisch anwendbarer Marker. Da der Fall von sehr kleiner DNA-Methylierung in gesunden Geweben und hoher DNA-Methylierung in den Tumoren für die Anwendung vorteilhaft ist. Die Analyse der Datenqualität hat ergeben, dass die Methode nicht nur gut Reproduzierbare Daten liefert sondern auch eine quantitative Auswertung möglich ist.

4.1.7. Auswirkungen der veränderten Enzymkombinationen

Bei der Veränderung der verwendeten Enzymkombination gab es zwei Ziele. Zum einen die Verringerung der Komplexität der zu hybridisierenden Amplicons und zum anderen die Erhöhung der Zahl informativer Fragmente. Zunächst muss jedoch auch gesagt werden, dass über diese Zahlen bis dahin nichts bekannt war, d.h. in erster Linie wurden hier erstmalig Komplexität und Genomabdeckung analysiert.

Betrachtet man also die beiden Methoden zunächst auf der Ebene der Fragmentzahlen, wird deutlich, dass in beiden Fällen nur circa ein Zehntel aller Fragmente eine methylierungssensitive Schnittstelle enthält, d.h. der größte Teil der Fragmente nicht informativ ist. Der größte Teil dieser Fragmente ist allerdings kleiner als 100 bp und kann daher durch die Aufreinigung abgetrennt werden. Betrachtet man die Fragmentzahlen nur für Fragmente > 100 bp wird deutlich, dass bei circa gleicher Gesamtzahl der Fragmente bei der Verwendung der veränderten Enzymkombinationen doppelt so viele informative Fragmente entstehen. Dies beruht allerdings z.T. darauf, dass die Fragmente kürzer geworden sind, d.h. aus einem informativen Fragment sind durch die hinzugekommenen Enzyme zwei oder mehr informative Fragmente geworden. Andererseits zeigt z.B. die Erhöhung der abgedeckten TSS um 5%, dass neue Fragmente hinzugekommen sind. Weiterhin stellen diese 5% circa 1200 TSS dar, die jetzt zusätzlich untersucht werden können.

Die Komplexität des Amplicons ist in beiden Fällen sehr hoch. Dies liegt daran, dass die kleinen Fragmente abgetrennt werden. In den meisten Methoden zur Herstellung genomischer Repräsentationen wird die Komplexität reduziert indem nur Fragmente bis zu einer bestimmten Größe amplifiziert werden. Auf diese Weise ist es wesentlich leichter, eine große Zahl von Basen abzutrennen. Bei der DMH Methode kann daher davon ausgegangen werden, dass auf diese Weise keine Reduktion der Komplexität auf ein Hundertstel erreichbar ist. Möglich wäre z.B. eine Größenfraktionierung, was allerdings einen wesentlich erhöhten Aufwand bedeuten würde und den Verlust von informativen Fragmenten.

Da die ursprüngliche Methode bereits trotz zu hoher Komplexität erfolgreich hybridisiert wurde, wurde auch in dieser Arbeit mit der hohen Komplexität gearbeitet und konnte im weiteren Verlauf gezeigt werden, dass stabile und aussagekräftige Hybridisierungssignale erzeugt werden können. Die Verwendung der veränderten Enzymkombination hat einen wesentlichen Vorteil, nämlich die Verkürzung der Fragmente und damit die bessere Amplifizierbarkeit und die verringerte Notwendigkeit von Komethylierung. Ein weiterer Vorteil wird in Abschnitt 4.1.5 beschrieben. Zusammenfassend kann also gesagt werden, dass durch die veränderten Enzymkombinationen die Komplexität in geringem Maße reduziert wurde und die bereits relativ hohe Genomabdeckung verbessert wurde. Der wesentliche Vorteil liegt aber in der Verkürzung der Fragmentlängen.

4.2. Methylierungsmarker in Brustkrebs

Bei der Behandlung von Krebs gibt es verschiedene Fragestellungen zu deren Beantwortung neue molekulare Marker gesucht werden und die Analyse der DNA-Methylierung geeignet ist, entsprechende Markerkandidaten zu finden. Zum einen werden Tumormarker für die Früherkennung gesucht. Diese sollten gezielt und einfach, d.h. mit geringerem Aufwand als einer Mammografie und besserer Spezifität, Brustkrebs in frühen Stadien oder schon als Vorstufen erkennen. Ebenso wird eine bessere Klassifizierung von Tumoren angestrebt, was eng mit der Suche nach prognostischen und prädiktiven Markern zusammenhängt. Während es bei der Prognose um die Bewertung der Aggressivität und somit der Wahrscheinlichkeit der Wiederkehr des Tumors oder einer Metastasenbildung geht, soll ein prädiktiver Marker eine Vorhersage der Wirksamkeit einer Behandlung, wie z.B. einer adjuvanten oder palliativen Chemotherapie, möglich machen. Ein weiteres Anwendungsgebiet von DNA-Methylierungsunterschieden ist die Verwendung als therapeutisches Target. Hierbei gibt es zwei Möglichkeiten. Zum einen kann die Veränderung der DNA-Methylierung einzelner Gene bzw. das Aufdecken neuer Mechanismen von Interesse sein zum anderen werden Demethylierungsagenzien heute für die Behandlung von Krebs getestet.

Es gibt verschiedene Veränderungen im DNA-Methylierungsmuster von Krebs insbesondere Brustkrebs. Eine ist die Hypermethylierung von ursprünglich unmethylierten CpG Inseln, die zur Inaktivierung von Genen, wie z.B. Tumorsuppressorgenen, führen kann, eine weitere die Demethylierung von CpG Inseln, die zur Aktivierung von Genen, wie z.B. Onkogenen, führt. Es gibt aber auch die globale Hypomethylierung des Genoms. Dabei handelt es sich um die Demethylierung der in gesundem Gewebe grundsätzlich methyliert vorliegenden Sequenzabschnitte wie Introns, Exons oder nicht-kodierende DNA-Abschnitte. Die Demethylierung dieser DNA-Sequenzen führt zur Instabilität des Genoms, da es sich hierbei um repetitive Elemente wie z.B. Transposons handelt und angenommen wird, dass DNA-Methylierung für die Unterdrückung homologer Rekombination verantwortlich ist. Eine weitere Veränderung ist der Verlust der elterlichen Imprints, d.h. der spezifischen DNA-Methylierung eines der beiden elterlichen Allele das LOI (englisch: loss of imprinting).

Aufgrund der Möglichkeit, die DNA-Methylierung mit Hilfe von 0% und 100% methylierten Kontrollen quantitativ auszuwerten, konnten die verschiedenen Veränderungen des DNA-Methylierungsmusters in dieser Arbeit getrennt betrachtet werden. Die mit Hilfe des verwendeten Mikroarrays analysierten Fragmente wurden zunächst hinsichtlich der Verteilung der DNA-Methylierung in gesundem Brustgewebe analysiert. Dies bildete die Grundlage für eine Einteilung der Fragmente in die drei DNA-Methylierungsgruppen methyliert, semi- und unmethyliert und bildete die Grundlage für eine Beurteilung, welche Art der DNA-Methylierungsveränderung vorlag. Anschließend wurden die Tumorproben mit gesundem Gewebe verglichen und, aufgeteilt nach Hypermethylierung und Demethylierung der entsprechenden Fragmente, betrachtet. Unterschiede, die bei diesem Vergleich

gefunden wurden, können als Kandidaten für Tumormarker und somit Früherkennungsmarker angesehen werden. Anschließend wurden die Tumorproben, aufgeteilt nach Ansprechenden und Nicht-Ansprechenden, untereinander verglichen und Kandidaten gefunden, die sowohl prädiktive als auch prognostische Markerkandidaten sein können.

4.2.1. Die verwendeten Patientinnen

Die in dieser Arbeit verwendeten Tumorproben sind primäre Tumore von Brustkrebspatientinnen die im Laufe der Erkrankung Metastasen bekommen haben. Diese wurden mit einem so genannten FAC Schema behandelt, welches aus den drei Wirkstoffen 5-Fluorouracil, dem Anthracyclin Adriamycin und Cyclophosphamid, die unterschiedliche Wirkmechanismen haben, besteht. Die Hauptfragestellung, die mit diesen Proben analysiert werden sollte, war die Suche nach Markerkandidaten für das Ansprechen der Metastasen auf die Chemotherapie. Zunächst würde man hierfür DNA der Metastasen verwenden wollen. Da Metastasen bei Brustkrebs nicht entfernt werden, liegt solches Probenmaterial aber im Allgemeinen nicht vor. Daher wurden die primären Tumore verwendet, welche somit auch das Probenmaterial darstellen, das später für die Beurteilung zur Verfügung stände. Ebenfalls konnte bereits gezeigt werden, dass das Expressionsmuster von primären Tumoren in der Metastase erhalten bleibt [23], d.h. es ist zu vermuten, dass dies auch für Methylierungsmuster gilt.

Die verwendeten Proben sind durchschnittlich 30 Jahre alt. Zum einen ist das Probenmaterial so alt, weil der Verlauf einer Brustkrebserkrankung heute einen langen Zeitraum umfassen kann. Proben werden erst dann aussagekräftig, wenn Patientinnen über den gesamten Verlauf ihrer Erkrankung beobachtet wurden und entweder als geheilt eingestuft werden konnten oder die Patientin verstarb. Die Proben stammen von Patientinnen, die keine adjuvante Chemotherapie erhalten haben, d.h. nachdem der primäre Tumor operativ entfernt wurde, ist zunächst keine weitere Behandlung vorgenommen worden, die die Rückkehr des Tumors oder die Ausbildung von Metastasen verhindern sollte. Daher ist es besonders gut möglich, von den primären Tumoren auf die Metastasen zu schließen und es kann ausgeschlossen werden, dass gefundene Unterschiede auch auf der unterschiedlichen Sensitivität gegenüber der adjuvanten Therapie beruhen. Heute ist es üblich, Frauen mit Brustkrebs immer mit einer adjuvanten Hormontherapie zu behandeln und häufig auch mit einer adjuvanten Chemotherapie, da sich gezeigt hat, dass diese in vielen Fällen sehr vorteilhaft ist und die Rückkehr des Tumors bzw. die Entwicklung von Metastasen verhindert. Daher ist die Zahl der Proben, die für diese Fragestellung verwendet werden können, beschränkt.

Auch wenn die in dieser Arbeit verwendeten Proben anders behandelt wurden als es heute üblich ist, sind die Daten von großer Relevanz. Sie bieten zum einen die Möglichkeit, gezielt

das Ansprechen gegenüber der verwendeten Chemotherapie zu untersuchen. Zum anderen sind Unterschiede, die bei dem Vergleich der Ansprechenden und Nicht-Ansprechenden der palliativen Therapie gefunden werden, möglicherweise auch übertragbar auf eine adjuvante Therapie. Weiterhin ist es nicht ausgeschlossen, dass ein prädiktiver Marker auch eine prognostische Komponente besitzt, d.h. dass er etwas über die Aggressivität des Tumors aussagt. Für die Metastasenbildung mit oder ohne Therapie kann es verschiedene Gründe geben. Der Tumor kann nicht vollständig entfernt worden sein. Es ist möglich, dass bereits Zellen gewandert waren, die die Lymphknoten befallen hatten oder Mikrometastasen gebildet haben. Diese konnten aber noch nicht gefunden und ihr Wachstum durch die adjuvante Therapie nicht unterbunden werden. Ein Tumor wird dann als aggressiv eingestuft, wenn er Metastasen ausbildet. Kam es jedoch rechtzeitig zur vollständigen Entfernung eines Tumors, kann der Verlauf der Erkrankung auch bei einem aggressiven Tumor ohne Metastasenbildung verlaufen. Andererseits wird ein weniger aggressiver Tumor der nicht vollständig entfernt wurde zunächst ein Rezidiv bilden und im Allgemeinen wird dann im weiteren Verlauf auch die Bildung von Metastasen beobachtet. Dies zeigt das Dilemma bei der Klassifizierung von Tumorproben in aggressiv und nicht aggressiv. Man würde alle hier verwendeten Proben als aggressiv einstufen, da Metastasen gebildet wurden. Andererseits ist es durchaus möglich, dass auch unter diesen Proben weniger aggressive Tumore waren, die nicht vollständig entfernt wurden. Dann beruht die Reaktion der Metastasen nicht nur auf der Ausbildung einer Resistenz gegen die Therapie sondern auch auf der Aggressivität des Tumors. Daher ist anzunehmen, dass ein prädiktiver Marker immer auch eine prognostische Komponente hat.

In Tabelle 13 sind die Charakteristika der einzelnen Tumorproben aufgelistet. 11 der 14 Ansprechenden und 12 der 14 Nicht-Ansprechenden waren invasive-duktales Karzinome (IDC) oder ein Gemisch aus duktalem Karzinom in situ (DCIS) und IDCs. In der Gruppe der Ansprechenden gab es weiterhin eine Probe bei der keine histologischen Angaben zur Verfügung standen und ein invasives-lobuläres Karzinom (ILC) und in der Gruppe der Nicht-Ansprechenden weitere zwei ILC. Es konnte bereits gezeigt werden, dass die verschiedenen Gewebeanteile von Brusttumoren unterschiedliche DNA-Methylierungsmuster aufweisen [105]. Weiterhin werden auch zum Zeitpunkt, wenn der Tumor beginnt größer und invasiv zu werden, d.h. in das umliegende Gewebe einzuwachsen, Veränderungen in den umliegenden Geweben induziert, z.B. durch die Vaskularisierung des Tumors oder durch Veränderungen des Stroma [17, 25, 126]. Es muss also davon ausgegangen werden, dass die Proben in dieser Hinsicht Unterschiede aufwiesen.

In Tabelle 13 sind weitere Eigenschaften, die bereits für die Prognose verwendet werden, wie der Hormonrezeptorstatus, der Lymphknotenstatus, Tumorgröße, Alter der Patientin und ob die Patientinnen pre- oder postmenopausal waren, aufgelistet. In Ansprechenden und Nicht-Ansprechenden sind diese Eigenschaften annähernd gleich verteilt, d.h. ein Einfluss sollte gering sein.

4.2.2. Die Methylierung in gesundem Brustgewebe

Um Beurteilen zu können, wie sich das DNA-Methylierungsmuster in einem Tumor verändert hat, ist es hilfreich, die Verhältnisse in den gesunden Zellen zu kennen. Dies klingt trivial, ist aber erschwert, da es sich sowohl bei Tumoren als auch bei gesunden Gewebeproben immer um Mischgewebe handelt und neben den dadurch erzeugten Unterschieden auch weitere individuelle Unterschiede zu erwarten sind.

Wie in Abschnitt 1.1.2 beschrieben, entwickelt sich Brustkrebs hauptsächlich aus Epithelzellen der milchabführenden Gänge (Duktuli) oder am zweithäufigsten aus Epithelzellen der Drüsenläppchen (Lobuli). Daneben gibt es noch seltene Formen, die in dieser Studie nicht betrachtet wurden. Dies macht die erste Schwierigkeit bei der Auswahl des geeigneten gesunden Gewebes deutlich. Beide Arten des Mammakarzinoms entstehen aus Epithelzellen, allerdings gibt es durchaus Unterschiede zwischen diesen Epithelzellen und dem Aufbau der Duktuli und Lobuli (siehe Abschnitt 1.1.2).

Unbehandelte Frischgewebeproben stammen meist aus Brustreduktionen und sind ein Gemisch aus verschiedenen Gewebearten, die in der Brust vorkommen, d.h. neben den entsprechenden Epithelzellen gibt es Myoepithel, Muskulatur, Fett- und Bindegewebe sowie Stroma und Blutgefäße, welche in unterschiedlichen Anteilen vorliegen können. Auch zwischen den verschiedenen gesunden Gewebearten gibt es DNA-Methylierungsunterschiede [127, 105]. Eine weitere Quelle für Variationen sind individuelle Unterschiede zwischen verschiedenen Patientinnen. Wie bereits erwähnt, verändert sich das DNA-Methylierungsmuster mit dem Alter [36, 37] und konnte gezeigt werden, dass es durch die Ernährung beeinflusst werden kann [32]. Weiterhin kann davon ausgegangen werden, dass auch andere Umwelteinflüsse individuelle Unterschiede hervorrufen bzw. durch natürliche Variation Unterschiede zwischen verschiedenen Individuen bestehen [128]. Aus den beschriebenen Gründen wurden in dieser Studie 2 Pools von gesundem Brustgewebe verschiedener Individuen verwendet. Zusätzlich wurden auch HMECs (human mammary epithelial cells, Cambrex) verwendet. Hierbei handelt es sich um Brustepithelzellen, die aufgereinigt und in Zellkultur über sechs Passagen vermehrt wurden.

In Abschnitt 3.2.5 sind die Ergebnisse für die genomweite Analyse der gesunden Gewebe dargestellt. Da in dieser Studie in erster Linie nach prädiktiven Markern gesucht wurde, ist Blut hier nicht mit betrachtet worden, auch wenn es als Kontrolle in der DMH-Analyse mitgeführt wurde. Die Werte der Blutanalyse sind in Tabelle 15 und Tabelle 16 mit aufgelistet, da hier nach Markern für die Früherkennung gesucht wurde.

Mit dem verwendeten Entwicklungs-Mikroarray wurden circa 3780 unterschiedliche Fragmente untersucht. Davon sind 13% als unmethyliert, 55% als semimethyliert und 32% als hypomethyliert eingestuft worden. Von den 3780 Fragmenten überlappen 1619 (44%) mit mit einer CpG-Insel, 2524 (68%) mit einem Gen und 1991 (54%) liegen in der Nähe (\pm 2000bp) eines Transkriptionsstarts (TSS). Von den CpG-Inseln fallen 278 (17%) in die

Gruppe der unmethylierten, 1137 (70%) in die Gruppe der semimethylierten und 204 (13%) in die Gruppe der methylierten Fragmente (siehe Abschnitt 3.2.5). Daraus folgt, dass der größte Teil der CpG-Inseln in die Gruppe der semimethylierten Fragmente fällt. Grundsätzlich kann davon ausgegangen werden, dass CpG-Inseln in gesundem Gewebe unmethyliert vorliegen. Ein gewisser Prozentsatz ist allerdings hypermethyliert. Yamada et al identifizierte 31 (20%) von 149 analysierten CpG-Inseln als hypermethyliert in Blut. Andere finden wesentlich geringere Werte von circa 6% [127, 104], allerdings wurde in diesen Arbeiten nach CpG-Inseln gesucht, die in unterschiedlichsten Gewebearten immer hypermethyliert vorliegen. In der Arbeit von Eckhardt et al. [127] wurden 873 Gene in 42 unterschiedlichen Geweben untersucht. 17% der analysierten Gene wiesen DNA-Methylierungsunterschiede in ihren 5'-UTRs auf, dies weicht stark von dem hier gefunden hohen Anteil von 70% an semimethylierten Fragmenten ab.

Eine Hauptursache für die hohe Zahl von Fragmenten, die eine DNA-Methylierung aufwiesen, ist, dass 50 bp Überlappung als Kriterium für die Überlappung verwendet wurden. Dadurch kann ein großer Teil des Fragmentes nicht CpG-Insel sein wodurch die gemessene DNA-Methylierung eigentlich nicht einer CpG-Insel zugeordnet werden sollte. Eine weitere Ursache ist die Auswahl der CpG-Inseln nach den Kriterien von Gardener und Frommer. Wie bereits erwähnt, existieren verschiedene Definitionen für CpG-Inseln und wurden striktere Kriterien als die hier verwendeten vorgeschlagen [127, 129]. Ein weiterer Grund ist die gemeinsame Verwendung der HMECs und der Pools gesunden Brustgewebes, d.h. die Daten sind eine Mischung der DNA-Methylierung verschiedener Gewebe und Individuen. Ist ein Fragment in einer der verwendeten Proben oder Gewebearten methyliert, kann dies zu einer mittleren DNA-Methylierung führen, die zwischen die Grenzwerte für die semimethylierten Fragmente fällt. Die gemeinsame Verwendung der beiden Brustgewebepools und der HMECs wurde trotzdem beibehalten, da für spätere Anwendungen eines Markers eine geringe DNA-Methylierung in allen den Tumor umgebenden Geweben von großer Bedeutung ist. Weiteren Einfluss haben die willkürliche Festsetzung der Grenzen für die DNA-Methylierungsgruppen und der bereits erwähnte Batcheffekt. Die Gruppe der unmethylierten Fragmente enthält alle Fragmente, die DNA-Methylierungswerte zwischen 0% und 30% aufweisen. Zunächst ist eine DNA-Methylierung von 30% bereits relativ hoch, d.h. entspricht einem zu erwartenden Wert für eine Mischmethylierung. Es ist jedoch möglich, dass z.B. der bereits erwähnte Batcheffekt oder allgemeine Signalschwankungen in einigen Fällen zu einem erhöhten Messwert beigetragen haben. Außerdem handelt es sich bei den HMECs um eine primäre Zelllinie, d.h. die Zellen sind nicht immortalisiert und wurden vorteilhafter Weise nur 6 Passagen unterzogen. Daher kann angenommen werden, dass durch die Vermehrung der Epithelzellen in Zellkultur keine großen Unterschiede eingeführt wurden. Trotzdem ist nicht auszuschließen, dass das DNA-Methylierungsmuster dieser Zellen Unterschiede gegenüber Epithelzellen in vivo aufweist, d.h. die HMECs können z.T. große Unterschiede zu den gesunden Proben aufweisen.

Um detaillierten Aufschluss über die Verhältnisse in gesundem Brustgewebe zu erhalten, müssten mehr Daten von Einzelindividuen und den verschiedenen Gewebearten erhoben werden. Weiterhin wäre die Analyse weiterer gesunder Gewebe und Zellmodelle vorstellbar. Für die hier erwünschte Möglichkeit, nach Unterschieden in den Tumorproben zu suchen, wurde der vorliegende Datensatz jedoch als ausreichend betrachtet.

4.2.3. Methylierungsunterschiede in Brustkrebs und deren Verwendung als Früherkennungsmarker

Es ist erwiesen, dass die Entwicklung von Krebs nicht nur von genetischen sondern auch wesentlich von epigenetischen Veränderungen beeinflusst wird [130, 131]. Daher haben DNA-Methylierungsunterschiede ein großes Potential, als Tumormarker in der Früherkennung angewendet zu werden. Weiterhin wird immer deutlicher, dass epigenetische Veränderungen sehr früh, d.h. lange bevor der Tumor invasiv wird, in den Zellen auftreten [130, 131]. Dies macht DNA-Methylierung für die Früherkennung besonders geeignet. Die Möglichkeit, Krebszellen zu detektieren, kann nicht nur in der Früherkennung sondern auch für die Überwachung des Wiederauftretens oder der Wirksamkeit einer Therapie verwendet werden.

Die Anwendung eines Markers in der klinischen Praxis erfordert einfache und billige Assays für die Untersuchung der DNA-Methylierung. Es ist daher vorteilhaft, Marker, die auf einem Mikroarray identifiziert wurden, in eine PCR-basierte Analyse eines oder weniger Einzelmarker zu übertragen. Häufig wird ein Assay angestrebt, welcher eine Realtime-PCR verwendet, da dadurch die Quantifizierung der DNA-Methylierung möglich wird.

Bei der Anwendung der DNA-Methylierung als Tumormarker in der Früherkennung ist häufig die Analyse der DNA in Körperflüssigkeiten das Ziel. Im Falle von Brustkrebs handelt es sich um Serum oder Plasma. Für die Klassifizierung von Tumoren, die Erstellung einer Prognose oder die Vorhersage der Reaktion auf eine Chemotherapie kann Tumormaterial verwendet werden. Besonders bei der Analyse von Serum oder Plasma, aber auch von Tumormaterial, muss die DNA-Methylierung der Tumor-DNA vor dem Hintergrund anderer DNA aus gesunden Zellen mit hochsensitiven Detektionsmethoden ermittelt werden. Dies ist einfacher, wenn die Hintergrund-DNA keine DNA-Methylierung aufweist bzw. ist es wesentlich schwerer, die DNA-Methylierung von Tumor-DNA zu messen, wenn im Hintergrund bereits eine DNA-Methylierung vorhanden ist, zumal sowohl in Serum und Plasma als auch in Tumorproben der Anteil der Tumor-DNA stark schwankt. Die Analyse der DNA-Methylierung in gesundem Brustgewebe und auch in Blut ist daher wesentlich für die erfolgreiche Entwicklung eines Markers.

In dieser Arbeit konnten den Hybridisierungssignalen durch die Verwendung externer Standards DNA-Methylierungsprozentwerte zugeordnet werden. Durch die Analyse von gesundem Gewebe wurden die analysierten Proben in die drei Gruppen methyliert, semi-

und unmethyliert eingeteilt. Für die weiteren Betrachtungen wurden die Fragmentgruppen methyliert und unmethyliert getrennt betrachtet, da dadurch eine genaue Einteilung der DNA-Methylierungsveränderung in Hypermethylierung und Demethylierung möglich war.

Es wurden 145 (30%) Fragmente gefunden, die eine Hypermethylierung gegenüber den unmethylierten gesunden Geweben zeigten. Auf der anderen Seite wurden 202 (17%) Fragmente gefunden, die eine Demethylierung in den Tumorproben zeigten. Die beiden Gruppen unterschieden sich, wie erwartet, stark in ihrer Zusammensetzung bezüglich der CpG-Inseln, TSS oder Lage innerhalb eines Genes. Auffällig hierbei war zum einen, dass nicht nur CpG-Inseln eine Hypermethylierung zeigten, sondern 50% der Fragmente nicht die hier verwendeten Kriterien (200 bp lang, GC-Gehalt > 50% und der Quotient von beobachteter zu erwarteter CpG-Dichte > 0.6) für die Zuordnung zu einer CpG-Insel erfüllen. Wie bereits erwähnt, gibt es verschiedene Definitionen für CpG-Inseln. Es ist also durchaus möglich, dass auch Regionen, die nicht als solche klassifiziert wurden, dazu zählen bzw. regulatorische Funktionen haben.

Andererseits wurden in der Gruppe der im Krebs demethylierten Fragmente 25% (51) CpG-Inseln gefunden. Hierbei kann es sich um Fragmente handeln, die nur zu einem geringen Anteil mit einer CpG-Insel überlappen, d.h. die Demethylierung betrifft, wie erwartet, nicht-regulatorische Sequenzabschnitte. Dann kann es sich auch um CpG Inseln handeln, die in repetitiven Regionen liegen, obwohl deren Zahl aufgrund der Auswahlkriterien für die Fragmente gering sein sollte. Einen weiteren Anteil werden aber auch Fragmente haben, bei denen die Demethylierung einer CpG-Insel gefunden wurde, die z.B. die Aktivierung des entsprechenden Genes zur Folge haben kann.

Die gefundenen Veränderungen in der Gruppe der in den Tumoren demethylierten Fragmente sind wesentlich größer als in der Gruppe mit Aufmethylierung (siehe Abschnitt 3.2.7). Das kann entweder bedeuten, dass es sich hierbei um Veränderungen handelt, die in allen Tumoren auftreten und in den verschiedenen Zellen und Geweben der Tumorproben gleichermaßen ausgeprägt sind oder dass dies ein Messartefakt ist, da ein solches Verhalten für Tumorproben, die auch als Mischgewebe anzusehen sind, nicht zu erwarten ist. Man würde vielmehr erwarten, dass die Fragmente eine DNA-Methylierung zwischen 30% und 70% aufweisen und daher in die Gruppe für Semimethylierung eingeordnet werden, wie es bei den Kandidaten für Hypermethylierung in Tumorproben der Fall ist. Eine technische Ursache hierfür kann der Batcheffekt sein, der zu einer falschen Einordnung einiger Fragmente für die gesunde Gewebe geführt hat. Da es sich bei den Fragmenten grundsätzlich um Fragmente handelt, die hohe DNA-Methylierungswerte aufweisen, kommen als weiterer Grund Signalschwankungen, die für hohe Signale auf Mikroarrays stärker ausfallen, in Frage. Andererseits wird die Demethylierung als globales Ereignis in Brustkrebs beschrieben, d.h. es ist das gesamte Genom betroffen. Dies würde bedeuten, dass in dieser Gruppe wesentlich mehr Kandidaten zu erwarten sind und somit auch Kandidaten mit größeren Unterschieden gefunden werden können.

Fast alle bekannten differentiell methylierten Gene in Brustkrebs und auch allgemein in Krebs wurden durch eine Analyse von ausgewählten Kandidaten gefunden. Eine genomweite Suche wurde bisher nur in wenigen Fällen durchgeführt [105]. Der Grund dafür war der Mangel an geeigneten Methoden. Obwohl in den letzten Jahren verschiedene neue genomweite Methoden entwickelt wurden (siehe Absatz 1.2.2), sind damit bisher kaum Tumorproben, insbesondere nicht in großer Zahl, analysiert worden. Bisher publizierte Daten gibt es aus der Gruppe von T. Huang, mit DMH erzeugt für Brust- [123]; Eierstock- [50] und Darmkrebs [124], sowie in der Arbeit von Costello et al. [101], in der verschiedene Tumore, darunter auch 14 Proben von Brusttumoren, mit RLGS analysiert wurden. Geringere Zahlen von Tumorproben wurden analysiert von Yuan et al. [113], hier wurden vier Leukämie Proben von Patienten, die mit einem Demethylierungsagens behandelt wurden analysiert sowie vier Wilmstumore im Vergleich zu gesundem Nierengewebe. Interessant ist auch der Datensatz von Hu et al [105]. Hier wurden Epithel- und Myoepithelzellen sowie stromale Fibroblasten von gesundem Brustgewebe von einem duktalem Karzinom in situ und von invasivem Brustkrebs isoliert und verglichen. Zusammenfassend kann man sagen, dass in keinem Fall gleichzeitig eine hohe Zahl von Proben analysiert wurde und die Sequenzen der analysierten Fragmente bekannt waren. Eine Auswertung, wie sie in dieser Arbeit vorgenommen wurde, war daher nicht möglich.

Der DMH Datensatz für Brustkrebs [123] umfasst 17 Tumorproben. Die Tumorproben wurden mit Cy5 gelabelt und zusammen mit gesunden Gewebeproben, die mit Cy3 gelabelt waren, hybridisiert. Diese Tumorproben zeigten in insgesamt 496 (15 bis 207 pro Tumorprobe) der analysierten CpG-Inseln einen Cy5/Cy3 Quotienten größer als 1,5, welcher von der Gruppe als Grenzwert für eine Hypermethylierung im Tumor angenommen wurde. Durchschnittlich wurden auf diese Weise 83 (1,1%) Klone identifiziert die eine Hypermethylierung in einer Tumorprobe aufwiesen. In dieser Studie wurden auch hypomethylierte CpG-Inseln gefunden, d.h. Fragmente die ein Cy5/Cy3 Quotienten kleiner als 0,5 aufwiesen. Eine genaue Zahl wird allerdings nicht angegeben und angenommen, dass es sich dabei ausschließlich um repetitive Elemente handelt. Daneben gab es noch zwei weitere Gruppen von Fragmenten, nämlich solche, die gleiche Signale sowohl für die gesunden als auch die kranken Samples zeigen und somit einen Cy5/Cy3 Quotienten nahe 1 zeigen und solche Fragmente deren Signal in beiden Fällen zu gering ist, um detektiert zu werden, da sowohl in den Tumor- als auch in den gesunden Proben keine DNA-Methylierung vorliegt. Über die Anzahl dieser Fragmente werden keine Angaben gemacht. Da keine Kontrollen für 0% oder 100% verwendet wurden, gibt es keine Aussage darüber, ob bestimmte Spots eventuell keinen ausreichenden Dynamikumfang aufweisen bzw. ein ausreichendes Signal erzeugen.

Ein Vergleich der hier vorgestellten Daten mit denen von Yan et al. ist also aus verschiedenen Gründen schwer möglich. Zunächst ist die Selektion der Fragmente, die als hypermethyliert gewertet wurden, unterschiedlich. Weiterhin handelt es sich bei den 7776

analysierten Fragmenten um angereicherte CpG-Inseln und der in dieser Arbeit verwendete Array enthält auch weniger CpG-reiche Regionen. Zum Zeitpunkt der Publikation der Daten waren lediglich 25% der CpG-Insel-Bibliothek sequenziert. Eine genaue Aufschlüsselung der Fragmente wurde daher nicht vorgenommen. Weiterhin hat der in dieser Arbeit erzeugte Datensatz den Vorteil, durch die Verwendung der externen DNA-Methylierungsstandards, quantitativ ausgewertet zu sein. Es ist außerdem möglich, Kandidaten zu selektieren, die besonders gut für die spätere Verwendung als sensitive Marker für die Analyse von Serum oder Plasma geeignet sind, da auf eine geringe DNA-Methylierung in den gesunden Gewebeproben selektiert werden kann. Dies ist insbesondere daher wichtig, da gezeigt wurde, dass auch in gesundem Gewebe DNA-Methylierung von CpG Inseln zu finden ist [129, 127].

In Abschnitt 3.2.9 sind die Ergebnisse der Bisulfitsequenzierung einiger Kandidaten gezeigt. Die Sequenzierung von insgesamt 62 Tumorproben zeigte, das Fragment 9857 aber auch 3615 sind gute Kandidaten für eine Weiterentwicklung als Früherkennungsmarker, da die gesunden Kontrollen kaum DNA-Methylierung aufweisen. Das Fragment 9857 liegt in einer CpG-Insel am Anfang des Genes *Cystein Dioxygenase, Type I (CDO1)*. Dieses Gen ist bisher nicht mit Krebs in Verbindung gebracht worden. Es dient hauptsächlich dem Abbau von Cystein in der Zelle, wurde aber zumindest auch mit der Antwort auf Entzündungen in Verbindung gebracht. Auch Fragment 3615 liegt in einer CpG-Insel am Anfang des Genes *Neuroglobin (NGB)*. Die Funktion dieses Genes ist ungeklärt. Es wird allerdings angenommen, dass es Zellen vor oxidativ induziertem Zelltod schützt.

Insgesamt können 112 der 145 hypermethylierten Fragmente Gene zugeordnet werden. Für einige davon wurde bereits Hypermethylierung in Krebszellen beschrieben, z.B. für *LAMC2* in Brustkrebs [132], in Prostatakrebs [56] und in Blasenkrebs [57] sowie für die Gene *DVL1* [133], *LSAMP* [134], *TOP1* [135], *DKK2* [136], *NRP2* [137] und *Wnt10B* [138]. Unter den gefundenen Genen sind einige Gene, die für ihre Beteiligung an der Entstehung von Krebs bekannt sind, wie z.B. *beta-Catenin (CTNNB1)*, das eine wichtige Rolle in der Zelladhäsion einnimmt aber auch als Transkriptionsfaktor wirkt, der Wachstumsfaktor Pleiotropin, der Zinkfinger und Transkriptionsfaktor *ZBTB16*, welcher ebenfalls an der Regulation des Zellwachstums beteiligt ist, der Zinkfinger und Transkriptionsfaktor *PIAS3* sowie die mitogen-aktivierte Kinase *MAP3K8*. Daneben wurde auch das an der Wachstumsinhibierung beteiligte retinoplastomabindende Protein *RBBP8* gefunden oder an der Apoptose beteiligte Gene, wie *ATG5* und *NCR1* und das DNA Reparatur Gen *ERCC5*. Neben den bekannten Genen wurden auch Gene gefunden, deren Beteiligung an der Tumorentstehung noch weniger bekannt ist und für die noch keine veränderte DNA-Methylierung in Krebszellen beschrieben wurde, die aber relevante Funktionen haben, wie z.B. die Zinkfinger *ZADH1*, *ZNF227*, *ZNF25/ZNF680* die Homeoboxgene *HOXA2*, *NKX2-3* und *EMX1* oder das Catenin *CTNNA3* etc.

Da es sich bei den hypomethylierten Sequenzabschnitten wesentlich seltener um Fragmente

in der Nähe eines TSS handelt, ist die Beurteilung der funktionellen Relevanz der gefundenen Gene wesentlich schwieriger. Es wurden zwei Gene gefunden für die bereits Demethylierung in Tumoren beschrieben wurde. Dabei handelt es sich um das Gen *BRDT* [139], ein so genanntes Cancer/Germline Gen, welches dadurch charakterisiert ist, dass es nur in männlichen Keimzellen oder Krebszellen expremiert wird. Das zweite Gen ist *DDX53* [140, 141] alias *CAGE*, was für englisch: cancer associated gene steht, da dieses Gen ausschließlich in Krebszellen expremiert wird [142]. Einige Fragmente liegen auch in der Nähe eines TSS interessanter Gene, wie z.B. *ADAM17*, einem Gen mit Metalloprotease Aktivität, für das z.B. gezeigt wurde, dass die Expression dieses Genes von gesundem Brustgewebe zu metastasierendem Brustkrebs kontinuierlich ansteigt [143] oder *HOXB4*, einem Homeobox Gen, und *DUSP7*, einer mitogen aktivierten Kinase, *SEPT7*, einem Gen das am Zellzyklus beteiligt ist, und *CITED2*, einem Transkriptionsfaktor.

Zusammenfassend kann gesagt werden, dass eine Reihe von Tumormarker-Kandidaten gefunden werden konnte. Diese müssen zunächst noch weiter validiert werden, d.h. die DNA-Methylierungsmuster müssen mit einer unabhängigen Methode und auf weiteren Proben bestätigt werden. Bei den gefundenen Kandidaten handelt es sich unter anderem um bekannte mit Krebs assoziierte Gene, aber auch um Gene, deren Rolle bei der Entwicklung von insbesondere Brustkrebs noch nicht bekannt ist, oder auch Sequenzabschnitte, die nicht mit einem Gen assoziiert sind. Die hier gefundenen Methylierungsmuster könnten daher auch ein Ansatzpunkt für funktionale Studien sein.

4.2.4. Markerkandidaten zur Vorhersage des Ansprechens auf eine Chemotherapie

Hauptziel der Arbeit war es, Kandidaten für die Entwicklung von prädiktiven Markern für das Ansprechen von Brustkrebspatientinnen mit Metastasenbildung auf eine FAC-Chemotherapie zu finden. Bei diesem Vergleich wurden keine großen Unterschiede erwartet, da diese Fragestellung durch verschiedene Faktoren erschwert wird. Zunächst ist davon auszugehen, dass die Resistenz von Tumoren gegen Chemotherapien, insbesondere bei einer wie in dieser Studie verwendeten Polychemotherapie, auf unterschiedlichen Mechanismen beruhen kann. Weiterhin stammen die verwendeten Proben von primären Tumoren, d.h. es ist nicht gesichert, dass sich bei allen Proben das Verhalten der Metastase schon im primären Tumor erkennen lässt. Die Metastasen auf denen die Untersuchung beruht, traten an verschiedenen Organen auf, was ebenfalls für unterschiedliche Mechanismen beim Ansprechen auf die Therapie oder die Entwicklung einer Resistenz verantwortlich sein kann. Darüber hinaus wird angenommen, dass ein gutes Ansprechen einer Patientin auf eine Therapie auch von anderen Kriterien, wie z.B. ihrer allgemeinen gesundheitlichen Verfassung, welche z.B. die Belastbarkeit der Leber oder die Stärke des Immunsystems beinhaltet, beeinflusst wird.

In Kapitel 1.1.6.3 wird gezeigt, dass es möglich ist, Marker zu identifizieren, deren DNA-Methylierungsstatus als prädiktiver Marker verwendet werden kann. Andererseits zeigt die Zahl der bereits identifizierten Kandidaten, dass auf diesem Gebiet noch ein großer Forschungsbedarf besteht, insbesondere wenn man bedenkt, wie viele unterschiedliche Krebs- und Therapiearten existieren. Für die in dieser Arbeit untersuchte palliative Anwendung einer FAC Chemotherapie, als erster Therapie nach dem Auftreten von Metastasen bei Brustkrebspatientinnen, lagen bisher keine Daten vor. Außerdem handelt es sich hierbei um eine Kombinationstherapie bei der eine Resistenz der Patientin ausgeprägt wird, wenn sie gegen alle drei der verwendeten Medikamente eine so genannte Kreuzresistenz aufgebaut hat. Die Kreuzresistenz von Tumoren führt letztendlich zum Scheitern der Chemotherapien bei 90% aller Krebspatienten die Metastasen ausgeprägt haben [144].

Die Ausbildung einer Chemotherapieresistenz ist eine komplexe Fragestellung, da verschiedene Faktoren Einfluss nehmen, wie z.B. ein verstärkter Transport des Wirkstoffes aus der Zelle heraus, die Veränderung der Aktivierung und Deaktivierung des Wirkstoffes, Veränderung in der Stoffwechselwege bzw. Zellfunktionen in die der Wirkstoff eingreift oder das Umgehen der Apoptose. Um eine Kreuzresistenz gegen verschiedene Wirkstoffe auszuprägen, kommen besonders der verstärkte Transport, das Umgehen der Apoptose und, da die meisten Wirkstoffe die DNA angreifen auch Defekte in der DNA-Reparaturmaschinerie in Frage.

Für einen verstärkten Stofftransport aus der Zelle heraus ist häufig die Überexpression von Transporterproteinen mit einer ATP bindenden Kasette (ATP binding cassette; ABC) verantwortlich. Der bekannteste ABC-Transporter ist das P-Glykoprotein, das durch das *Multi Drug Resistance Gen MDR1* kodiert wird [145]. ABC-Transporter sind in der Lage, z.B. Taxane, Mitosehemmer oder Anthracycline und auch 5-Fluorouracil aus der Zelle auszuschleusen [146]. Da Apoptose und DNA-Reparatur hingegen komplexe Mechanismen in der Zelle darstellen, ist die Eingrenzung der Gene, die hier Einfluss nehmen, wesentlich schwieriger. Sehr bekannte Gene für den Einfluss auf die Apoptose und ihre Rolle in der Krebsentwicklung sind *p53*, dessen Rolle für die Chemotherapieresistenz allerdings gegensätzlich ist, oder die bereits erwähnte *Caspase 8*. Bei der DNA-Reparatur sind Gene der Nukleotid Exzisions Reparatur, wie *ERCC1*, oder der Mismatch Reparatur, wie das bereits erwähnte *hMLH1*, von Bedeutung [146].

Ein anderer Einflussfaktor ist die bereits erwähnte prognostische Komponente, d.h. die Aggressivität der Metastasen. Hierbei ist einerseits möglich, dass gerade solche Metastasen, die besonders schnell wachsen, auf Zytostatika ansprechen, da Zytostatika nur bei sich teilenden Zellen wirken, andererseits werden langsam wachsende Metastasen möglicherweise als stabile Erkrankung eingeordnet, was häufig auch als Ansprechen auf die Therapie gewertet wird obwohl eigentlich eine Resistenz vorliegt.

Wie bereits beschrieben, weisen prinzipiell mehr Ansprechende ein entsprechendes hypermethyliertes Fragment als Nicht-Ansprechende auf, d.h. bei diesen Kandidaten könnten die Expression der Gene entweder eine Resistenz unterstützen oder zu einer negativen Prognose beitragen, was dann wiederum wie eine Resistenz der Metastase aussieht. Im Folgenden kann gezeigt werden, dass einige der hier gefundenen Kandidaten genau diese Kriterien aufweisen.

Überexpression des Genes *KPNA2* wurde bereits als Marker für eine schlechte Prognose, d.h. eine kürzere Zeit bis zum Auftreten von Metastasen und auch eine kürzeres Gesamtüberleben in Brustkrebs gefunden [147]. Eine möglicherweise durch die Hypermethylierung gekennzeichnete Unterdrückung der Expression dieses Genes könnte also zu einer besseren Prognose führen, und dies ist in der Gruppe der Ansprechenden häufiger der Fall. Das Gen *SEMA3E* ist ebenfalls häufig überexprimiert in Brustkrebszellen und es konnte gezeigt werden, dass dieses Gen direkt am invasiven Wachstum, der Zellwanderung und der Bildung von Lungenmetastasen beteiligt ist [148].

Weitere in Brustkrebs überexprimierte Gene die zur Aggressivität des Tumors beitragen, sind *STRA13*, das eine wichtige Rolle beim Übergang von Brusttumoren zum invasiven Wachstum spielt [149], oder das Gen *MAP3K8* (auch bekannt als *TLP-2* für tumor progression locus), das bekannt ist, normale Zellen in Tumorzellen zu transformieren [150].

In dieser Arbeit wird auch das Cytochrom *P450 1B1 (CYP1B1)* gefunden. Überexpression von *CYP1B1* ist relevant für die Entwicklung von verschiedenen Tumoren insbesondere hormonabhängigen wie Brustkrebs. Es ist am Östrogenmetabolismus beteiligt und in den Metabolismus verschiedener Chemotherapiewirkstoffe involviert [151]. Weiterhin wurde bereits gezeigt, dass dieses Gen durch DNA-Methylierung des Promotors reguliert sein kann [79]. Weiterhin finden Widschwendter et al., dass Patientinnen, die mit Tamoxifen behandelt wurden, und DNA-Methylierung dieses Genes aufweisen, längere Überlebenszeiten aufweisen, da *CYP1B1* die Wirksamkeit von Tamoxifen verringert. Im Gegensatz dazu weisen Patientinnen, die nicht mit Tamoxifen behandelt wurden, bei DNA-Methylierung des Genes eine schlechtere Prognose auf.

Auch in der Gruppe der in den Tumoren demethylierten Kandidaten gibt es Gene deren Funktion zu einer Resistenz beitragen könnte. So wurden z.B. die ABC Transporter *ABCF2* und *ABCG8* gefunden. Da hier die Nicht-Ansprechenden vermehrt eine Demethylierung zeigen, würde die Aktivierung dieser Transporter also zur Resistenz der entsprechenden Tumore beitragen. Während *ABCG8* bisher nicht mit Tumoren in Verbindung gebracht wurde, konnte bereits gezeigt werden, dass *ABCF2* eine Rolle in Brustkrebs spielt [152].

Weiterhin wurde das Gen *PRKCQ* gefunden, welches die Expression von *MDR1* aktiviert [153] und somit zu einer Chemotherapieresistenz beitragen kann. Bei diesen Kandidaten liegen die analysierten Fragmente allerdings nicht in der Nähe des TSS. Es ist also in diesem Fall unklar, welche Funktionalität die veränderte DNA-Methylierung hat. Hierzu sind weiterführende Untersuchungen notwendig.

Durch die in dieser Arbeit durchgeführte Bisulfitsequenzierung bisher weniger Kandidaten konnte das Fragment 1322 als besonders viel versprechend bestätigt werden. Interessanterweise liegt dieses Fragment in einer Region in der bisher keine Gene identifiziert wurden. Seine funktionale Relevanz ist also ungeklärt. Dies zeigt andererseits, dass man sich bei der Suche nach neuen Markern nicht auf Promotorregionen bekannter Gene beschränken sollte. Zusammenfassend kann gesagt werden, dass unter den gefundenen Kandidaten Gene mit relevanter Funktion enthalten waren und ein Kandidat gefunden wurde, dessen differentielle DNA-Methylierung in einer größeren Probenzahl bestätigt werden konnte.

4.3. Zusammenfassung

Brustkrebs ist die häufigste Krebserkrankung bei Frauen in Deutschland. Sowohl für die Früherkennung als auch für die Prognose des Krankheitsverlaufes oder die Vorhersage des Ansprechens auf eine Chemotherapie werden daher neue Biomarker gesucht. DNA-Methylierung ist eine epigenetische Veränderung, die sehr früh und häufig bei der Entstehung von Krebs beobachtet wurde. Daher sind Methylierungsunterschiede für alle genannten Fragestellungen relevant. Hauptziel dieser Arbeit war es, neue DNA-Methylierungsbiomarker für das Ansprechen von Brusttumoren auf eine Chemotherapie zu finden. Um dieses zu erreichen wurde eine Technologie benötigt, die es möglich machte die DNA-Methylierung einer Vielzahl von Sequenzabschnitten möglichst quantitativ zu untersuchen.

Zu diesem Zweck wurde die Methode der Differentiellen Methylierungs Hybridisierung (DMH) optimiert. Zunächst wurde durch veränderte Enzymkombinationen der Anteil analysierbarer CpGs erhöht und die Komplexität reduziert. Weiterhin wurde eine Verkürzung der analysierbaren Fragmente erreicht, weshalb Komethylierung nur noch in kürzeren Abschnitten notwendig war. Veränderungen des Protokolls führten zur Senkung der benötigten DNA Ausgangsmenge von 2000 ng auf 1000 ng. Ein wesentlicher Schritt war die Einführung eines für die entstehenden DMH-Fragmente spezifisch designten Mikroarrays. Es wurde zunächst ein Entwicklungs-Mikroarray verwendet, um Kriterien für die Auswahl der Fragmente zu optimieren sowie die verbesserte Methode zu charakterisieren. Besonders wichtig für die quantitative Datenauswertung war die Einführung von Kontrollen für die Normalisierung von Komplexitätsunterschieden zwischen den hybridisierten Proben.

Die Leistungsfähigkeit der veränderten Methode wurde mit Hilfe von Daten der Brustkrebszelllinie MDA-MB-231 im Vergleich zu DNA aus Blut analysiert. Die Reproduzierbarkeit des gesamten Prozesses war mit einem durchschnittlichen Korrelationskoeffizienten von 94% sehr hoch. Die Analyse einer Auswahl von 111 Fragmenten durch direktes Bisulfitsequenzieren zeigte für den Vergleich der Daten beider Methoden ebenfalls eine hohe Korrelation von 71%. Dadurch konnte gezeigt werden, dass die verbesserte Methode eine quantitative genomweite Methylierungsanalyse erlaubt. Auf

der Basis der Daten des Entwicklungs-Mikroarrays wurde ein Mikroarray entwickelt, der die Analyse von 50.000 Fragmenten parallel möglich machte. Mit diesem Microarray wurde die Methylierungsanalyse von 46% aller Promotorregionen und 47% aller Ensembl annotierten CpG-Inseln möglich.

Für die Suche nach DNA-Methylierungsbiomarkern für das Ansprechen auf eine Chemotherapie wurden 28 Proben primärer Tumore verwendet und mittels der verbesserten DMH-Methode auf dem Entwicklungs-Mikroarray analysiert. Der Entwicklungs-Mikroarray ermöglicht die parallel Analyse von 3780 unterschiedlichen Fragmenten. Aufgrund des Verhaltens der Metastasen gegenüber einer Anthracyclin basierten Chemotherapie wurden diese Proben aufgeteilt in 14 auf die Therapie Ansprechende und 14 auf die Therapie Nicht-Ansprechende Patientinnen. Für den Vergleich der Tumore mit gesundem Gewebe wurden zwei Pools gesunder Gewebeproben und eine primären Zelllinie humaner Brustepithelzellen (HMEC) untersucht. Durch die Verwendung von 0% und 100% methylierten DNA-Kontrollen war es möglich die Daten in Methylierungswerte zu normieren. Anschließend wurden die Fragmente in die drei Gruppen unmethyliert, semimethyliert und methyliert im gesunden Gewebe eingeteilt. Die Gruppeneinteilung wurde verwendet, um sich bei den nachfolgenden Vergleichen auf die beiden Gruppen unmethyliert und methyliert in gesundem Gewebe zu fokussieren. Der Vergleich der Tumorproben mit gesundem Gewebe ergab 145 Kandidaten für Aufmethylierung und 202 für Demethylierung in den Tumoren. Diese Tumormarker-Kandidaten können als Grundlage zur Entwicklung von Früherkennungsmarkern verwendet werden. Beim Vergleich von Ansprechenden und Nicht-Ansprechenden wurden für die im gesunden Gewebe unmethylierten Fragmente 55 Kandidaten und für die methylierten 85 Kandidaten gefunden.

Für die Weiterentwicklung dieser Kandidaten zu Früherkennungsmarkern oder Markern für das Ansprechen auf eine Chemotherapie wäre zunächst die Bestätigung der Methylierungsunterschiede mit einer unabhängigen Methode und in unabhängigen Proben notwendig. Ein erster Schritt in diese Richtung war die Analyse von 12 Tumormarker-Kandidaten durch direkte Bisulfitsequenzierung. Die Kandidaten waren nach unterschiedlichen Kriterien ausgewählt, d.h. es wurde nicht auf die Größe der Unterschiede innerhalb der Kandidatengruppe geachtet. 11 der 12 Fragmente zeigten auch in der Bisulfitsequenzierung, dass sie als Tumormarker geeignet sind. Für sechs dieser Fragmente wurde dieses Verhalten im Anschluss auf weiteren 34 Proben bestätigt. Sechs der ausgewählten Tumormarkerkandidaten waren ebenfalls Kandidaten das Ansprechen auf die Chemotherapie vorherzusagen. Für zwei Fragmente konnte dies auch durch die Bisulfitsequenzierungsdaten bestätigt werden. Besonders zu erwähnen ist das Fragment 1322 welches ein p-value von 0,03 beim Vergleich der insgesamt sequenzierten 29 Ansprechenden mit 21 Nicht-Ansprechenden aufwies, d.h. dieses Fragment zeigt in den hier untersuchten Proben einen signifikanten Unterschied zwischen den beiden Probengruppen und ist daher ein viel versprechender Kandidat für einen prädiktiven Marker.

5. Anhang

5.1. Firmen

Kurzform	Name	Adresse
ABI	Applied Biosystems,	Californien, USA
Affymetrix	Affymetrix, Inc., Californien, USA	Californien, USA
Amersham	Amersham Biosciences	Piscataway, NJ, USA
AXXORA	AXXORA	Lörrach, Deutschland
Bio-Rad	Bio-Rad Laboratories	Californien, USA
Biozym	Biozym Scientific	Oldendorf, Deutschland
Cambrex	Cambrex	East Rutherford, NJ, USA
Eppendorf	Eppendorf AG	Hamburg, Deutschland
Eurogentec	Eurogentec	Köln, Deutschland
Fermentas	MBI Fermentas GmbH	St.Leon-Rot, Deutschland
Fluka	Fluka Chemie GmbH	Buchs, Schweiz
GE Healthcare	GE Healthcare	Chalfont St. Giles, Großbritannien
GFL	GFL GmbH	Burgwedel, Deutschland
IKA	IKA	Wilmington, USA
Invitrogen	Invitrogen GmbH	Karlsruhe, Deutschland
Merck	Merck KGaA	Darmstadt, Deutschland
Millipore	Millipore	Billerica, MA, USA
NanoDrop	NanoDrop Technologies	Delaware, USA
NEB	New England Biolabs	Ipswich, MA, USA
OHAUS	OHAUS Corporation	New Jersey, USA
PHASE	PHASE GmbH	Lübeck, Deutschland
Pierce	Pierce sold as Thermo Scientific	Rockford, IL, USA
Promega	Promega GmbH	Mannheim, Deutschland
Qiagen	Qiagen GmbH	Hilden, Deutschland
Roche	Roche Diagnostics	Penzberg, Deutschland
Roth	Carl Roth GmbH & Co.	Karlsruhe, Deutschland
Sartorius	Sartorius AG	Göttingen, Deutschland
Sigma	Sigma-Aldrich	Deisenhofen, Deutschland
Tecan	Tecan AG	Männedorf, Schweiz
Perbio	Perbio Science sold as Thermo Scientific	Rockford, IL, USA

5.2. Geräte

Gerät	Provider
36 cm Kapillaren (3730 DNA Analyzer Capillary Array)	ABI
3730 DNA Analyzer (Sequenzierung)	ABI
Analysen Waage, Explorer	OHAUS
Biophotometer	Eppendorf
CCD Gel Cam S/W	PHASE
Centrifuge 5810 R und 5417 R	Eppendorf
Centrifuge MiniSpin	Eppendorf
Engine Tetrad [®] Thermocycler	Bio-Rad
Gelelektrophorese-Kammern	Biozym
GeneChip [®] Fluidics Station 450	Affymetrix
GeneChip [®] Hybridization Oven 640	Affymetrix
GeneChip [®] Scanner 3000	Affymetrix
GENios Plattenspektrophotometer	Tecan
Magnetrührer RH basic	IKA
Minishaker MS2	IKA
NanoDrop [®] ND-1000 Spektrophotometer	NanoDrop
Referenz-Pipetten	Eppendorf
Spannungsgeber (Gelelektrophorese)	Biozym
Thermocycler Mastercycler [®] gradient	Eppendorf
Thermocycler Mastercycler [®] gradient	Eppendorf
Thermomixer 5355 Comfort	Eppendorf
Waage BL 600	Sartorius
Waterbath 1002	GFL

5.3. Chemikalien

Chemikalien	Provider
Agarose	Eurogentec
Bromphenolblau	Roth
Dioxan	Merck
Ethidiumbromid	Merck
Glycerin	Fluka
Hydrochinon	Merck
6-Hydroxy-2,5,7,8-Tetramethyl- chromane-2-Carboxyl-Säure	Fluka
Natriumsulfit	Merck
Natriumbisulfit	Merck
Tris, Trishydroxymethylaminomethane	Roth
Wasser, steril, gefiltert 0,2 µm, DEPC behandelt	Fluka
POP-7 [™] Polymer für 3730/3730xl DNA Analyzers	ABI
Sulfolan	Merck
50XDenhardt's	Eppendorf
0.5M EDTA	Invitrogen
3% Tween	Pierce
20xSSPE	Cambrex
Alle anderen Chemikalien wurden von Sigma bezogen.	

5.4. Puffer und Lösungen

Puffer und Lösungen

Gelelektrophorese

TBE-Puffer (pH 8,0-8,3)	90 mM Tris, 90 mM Borsäure, 2,5 mM EDTA
TE-Puffer (pH 8,0-8,3)	10 mM Tris, 1 mM EDTA
Ladepuffer	3 mM Bromphenolblau, 23 % Glycerin

Bisulfitreaktion

Bisulfit-Lösung	0,4708 g/ml Natriumbisulfite; 0,1128 g/ml Natriumsulfite in Wasser, pH 5,5
Dioxan mit Radikalfänger	98,6 mg 6-Hydroxy-2,5,7,8-Tetramethylchromane-2-Carboxyl-Säure in 2,5 ml 1,4-Dioxan
Sangerpuffer	320 mM Tris-HCl (pH 9), 6 mM MgCl ₂ und 10 % [v/v] Sulfolan

DMH-Amplifikat Herstellung

ATP	0,275 g ATP disodium salt in 40ml Wasser, pH 7,5 (nur mit NaOH einstellen)
-----	--

Affymetrix Hybridisierung und Waschen

MES 12x	1,22 M MES; 0,89 M [Na ⁺] pH 6,5-6,7 0,02 M EDTA; 0,2 M NaH ₂ PO ₄ x H ₂ O; 3,6 M NaCl; pH 7,0 mit NaOH
SSPE (20 x)	7,0 mit NaOH
Waschpuffer A	6 x SSPE; 0,01% Tween 20
Waschpuffer B	0,6 x SSPE; 0,01% Tween 20

5.5. Enzyme, Kits und molekularbiologische Reagenzien

Enzyme, Kits und molekularbiologische Reagenzien	Provider
Anti-Streptavidin Antibody (0,5mg/ml)	AXXORA
Basenleiter, 100bp DNA ladder	Invitrogen
<i>Bfa</i> I, Type II restriction enzyme	NEB
BigDye Terminator 3.1 Chemie	ABI
<i>Bst</i> U I, Type II restriction enzyme	NEB
Control Oligo B2	Affymetrix
<i>Csp6</i> I, Type II restriction enzyme	Fermentas
Deep Vent _r (exo ⁻) DNA Polymerase	NEB
DNA, humane genomische DNA aus peripherem Blut	Promega
DNA, Lambda genomic DNA	NEB
dNTPs (dATP, dCTP, dGTP, dTTP)	Fermentas
DyeEx 96 Aufreinigungsplatten	Qiagen
ExoSAP-IT,	GE Healthcare
GeneChip Mapping 10K Xba Assay Kit	Affymetrix
GenomiPhi - Whole Genome Amplification Kit	Amersham
Hering Sperm (10mg/ml)	Promega
<i>Hin</i> P1 I, Type II restriction enzyme	NEB
HotStarTaq DNA Polymerase	Qiagen
<i>Hpa</i> II, Type II restriction enzyme	NEB
<i>Hpy</i> CH4 IV, Type II restriction enzyme	NEB
Human Cot-1 (1mg/ul)	Roche
Microcon Centrifugal Filter Units YM30	Millipore
<i>Mse</i> I, Type II restriction enzyme	NEB
QIAamp DNA Mini Kit	Qiagen
QIAminElute PCR Purification Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
SAM, S-Adenosyl-Methionin	NEB
SssI (CpG) Methylase	NEB
Streptavidin (1mg/ml)	Perbio
Streptavidin Phycoerythrin (1mg/ml)	Invitrogen
T4 DNA Ligase	NEB

5.6. Bisulfitprimer - Brustkrebsstudie

Fragment ID	Gen Name	Forward primer	Reverse Primer
4563	ZNF184	TTTTAGAAGTTTTAGGTTTTTTTATAG	CCACAACCAAAAACCATACAA
4493	novel trans	GGAGGGTGTGTTGTTTTATT	AATACCTCTCATCTCTCTATTCC
12787	no gene	AAGTTTTAAAGATATTATTGGTTTTTT	CCTCCCTCTCTACCTTTACTTAAT
4664	HOXA2	TTTTTTTTTTTTTTATTTAGGGGTATT	AACAAACCATTCCCAACCTA
9857	CDO1	GGTAAAGAGTTGGTGTAGGATG	CCCCCAACTCCTTAACTC
7951	ERCC5	TTTAGGGGTTTTGGAAGTTGT	ATCCCTACACCCCAAATCC
1322	no gene	TAGGTTGAATTTTTGAAGGTGT	TAAACTACAAATATCTACCAAACTC
3615	NGB	GTTTTGTTGGTTGGAGGG	AATAAAAACACCCCAACTATACTTC
4235	no gene	GATTTGATTAGGAATGTTAGAGTT	TCAACTATAATCCTCTCAAATCAA
1589	ZNF145	GAGAGAGAAAAGTTTTTTTTAGGA	ATAAACCAATACCAAAAAATCC
13165	Q96HQ8	GGGATGGTGTTTTTGGATATT	CTACTCTCCTAAAAAACCTAAAAACTA
10422	Q96HD6	GGTTTTATTAGGTAGTTAGAATGGTAG	AAAACTAATCCCACCTCCC

5.7. Fragmentsequenzen – Brustkrebsstudie

Fragment ID	genomic sequence
4563	ccctagaagcttcaggctcctccacaggcgagtgctctaccacctatcaggtaaagacgtgcaactgacactcagcgcgaggaggcc cgctgagaagggtttacctgtccgaagagaactgcgcttgcattgctgctgctgctgctg
4493	ggagggtgctgctgccccaccatgagaaacctggaagcgggtagggtgaggtgctgaggtcgggtgactggtgctgatagaaaa cgcgagaggaacaggagagatgagaggcatt
12787	aagtccaaagacatcactggtcctctttggaccggagtctccggagaaaccgctccggcgtcctgacacactgagatagaagttc acagaaatagggaaggtataggcgagggctccagggcagggcgtcctagggcacctggcgcaggggaggaccggaaccccag gcaaacgcagagtcggagatggcattccaggcgaagatcaagcaaggcagagaggagg
4664	cttctctcctcatccaggggtactcggcggtgcagggcgccggcgaccgggctcccgcgggctgccgagggtcggctggg cggcgccagcgcgctggcgagggtgactccgggtcaggctgggaatggtctgct
9857	ggcaaagagctggtgcaggatgaggatcagatcagccagggtccgtggctcagcacttcggtctgttccatctctggtgggagctggctgc cgcgcgctcactgctgggctgctggaggagctgagcgcagcaaggagctggggg
7951	tccaggggctctggaagctgctggagtctccgggcggcaggtcagccccgaagcgtggaagggaagatcctggtgttggtatcctta acgccgcttgggacttgggtgcagggat
1322	tgagcactacaatatctaccagggtcttattttaatcgtaccgtatgcatgaaacgccagcatcggatttggcgtaatcaagacaattc acacctcaaagaattcagcctg
3615	ggtggagacacccagctgtgcttccggggaccccgcttgccgctgcgccctgcgcccgagccgagccgcccaccccgcc gacgctgggctggcgctttaaagccgccacggccccgcccctccaggccagcaaggc
4235	tcaactgtggtcctcagatcaggcctcctggaccgatttccagctctcggggcgtctacgcgctggcgagccgcccctcaggag gccaccgcccctctggggcgtaccattggaagcctggaacgcgggggctgaggtcgtgggctctggccattcctggtcagg tc
1589	gagagagaaaagtttcccaggactcgaccttggcctccagcaatcgcgacagctaaaaacgggtgctcgtctcgacaatagatcccc cggacctctggcacctggtcac
13165	gggatggtgcctctggacatttccgacagcaggatcgttgggtgaggttagcagcgcagccgggttagcggagcagcctcaggt ctcctaggagagcag
10422	ggggctgatcccacctcccagggctccggtcgtcgtcgtcgtcggaccgagaaacgggagcatttccgagctcggggccgaccc aaccgctgccattctagctacctagtgagcc

5.8. Proben

Tabelle 13: Proben der Tumorstudie

Name	Responsgruppe	Respon- s- Typ	Hormon- status	Östrogen	Pro- gesteron	Tumor- grösse: T	Lymph- knoten- status: N	Fern- Metastasen: M	Histologie	Stadium	Analyse in	Alter	Menopause
1	Responder	CR	neg	10	1	T2	>3	M0	n.a.	IIb	Tumorstudie	49	pre
12	Responder	CR	pos	512	23	T2	0	M0	IDC	IIa	Tumorstudie	60	post
30	Responder	CR	neg	6	4	T1	0	M0	IDC	I	Tumorstudie	37	pre
11	Responder	PR	neg	0	0	T2	0	M0	ILC	IIa	Tumorstudie	n.a.	n.a.
14	Responder	PR	neg	5	0	T1	0	M0	IDC	I	Tumorstudie	40	pre
18	Responder	PR	neg	0	0	T2	0	M0	IDC	IIa	Tumorstudie	25	pre
19	Responder	PR	pos	598	500	T2	>3	M0	IDC	IIb	Tumorstudie	33	pre
22	Responder	PR	pos	68	179	T2	0	M0	DCIS+IDC	IIa	Tumorstudie	48	post
23	Responder	PR	pos	65	1	T1	0	M0	IDC	I	Tumorstudie	66	post
26	Responder	PR	pos	70	46	T1	>3	M0	IDC	IIa	Tumorstudie	41	pre
27	Responder	PR	neg	0	0	T3	0	M0	IDC	IIb	Tumorstudie	48	pre
28	Responder	PR	neg	7	14	T2	0	M0	IDC	IIa	Tumorstudie	51	pre
29	Responder	PR	neg	6	0	T2	0	M0	DCIS+IDC	IIa	Tumorstudie	67	post
6	Responder	PR	neg	0	0	T2	0	M0	DCIS+IDC	IIa	Tumorstudie	39	pre
10	Non-Responder	PD	neg	0	0	T2	>3	M0	IDC	IIb	Tumorstudie	52	post
15	Non-Responder	PD	pos	52	70	T1	1-3	M0	ILC	IIa	Tumorstudie	66	post
16	Non-Responder	PD	pos	84	927	T1	>3	M0	IDC	IIa	Tumorstudie	53	post
17	Non-Responder	PD	n.a.	n.a.	n.a.	T1	1-3	M0	IDC	IIa	Tumorstudie	44	pre
2	Non-Responder	PD	neg	0	0	T2	0	M0	IDC	IIa	Tumorstudie	45	pre
20	Non-Responder	PD	neg	5	0	T3	0	M0	IDC	IIb	Tumorstudie	43	pre
21	Non-Responder	PD	neg	0	0	T1	0	M0	DCIS+IDC	I	Tumorstudie	53	post
24	Non-Responder	PD	pos	186	2	T1	0	M0	IDC	I	Tumorstudie	43	pre
25	Non-Responder	PD	neg	7	0	T2	0	M0	IDC	IIa	Tumorstudie	60	post
4	Non-Responder	PD	neg	0	0	T2	0	M0	IDC	IIa	Tumorstudie	53	pre
5	Non-Responder	PD	pos	204	34	T2	0	M0	ILC	IIa	Tumorstudie	57	post
7	Non-Responder	PD	pos	13	80	T1	0	M2=SCCI	IDC	IV	Tumorstudie	46	pre
8	Non-Responder	PD	pos	194	46	T2	0	M0	IDC	IIa	Tumorstudie	51	post
9	Non-Responder	PD	neg	0	0	T2	0	M0	IDC	IIa	Tumorstudie	36	pre

Nr.	ID	Responsgruppe	Respons Typ	Hormon- status	Östrogen	Pro- gesteron	Tumor- grösse: T	Lymph- knoten- status: N	Fern- Metastasen: M	Histologie	Stadium	Analyse in	Menopa- use	Alter
32	1142	Responder	PR	neg	0	0	T2	>3	M0	DCIS+IDC	IIb	Bisulfitseq.	n.a.	n.a.
35	1145	Responder	PR	neg	7	9	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
41	1158	Responder	PR	pos	18	17	T1	0	M0	n.a.	I	Bisulfitseq.	n.a.	n.a.
42	1165	Responder	PR	pos	23	237	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
43	1167	Responder	PR	neg	1	1	T3	1-3	M0	n.a.	IIIa	Bisulfitseq.	n.a.	n.a.
44	1171	Responder	PR	neg	0	0	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
46	1175	Responder	PR	pos	15	81	T1	>3	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
48	1182	Responder	PR	pos	11	0	T1	0	M0	DCIS+IDC	I	Bisulfitseq.	n.a.	n.a.
52	873	Responder	PR	pos	45	2	T2	0	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
54	967	Responder	PR	pos	30	89	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
55	975	Responder	PR	pos	86	0	T3	0	M0	IDC	IIb	Bisulfitseq.	n.a.	n.a.
56	339	Responder	PR	pos	610	72	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
57	997	Responder	PR	neg	10	0	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
61	1164	Responder	PR	neg	0	2	T2	0	M0	DCIS+IDC	IIa	Bisulfitseq.	n.a.	n.a.
62	1168	Responder	PR	pos	239	692	T2	>3	M0	n.a.	IIIa	Bisulfitseq.	n.a.	n.a.
33	1143	Non-Responder	PD	n.a.	n.a.	n.a.	T1	0	M0	DCIS+IDC	I	Bisulfitseq.	n.a.	n.a.
34	1144	Non-Responder	PD	pos	82	401	T2	0	M0	DCIS+IDC	IIa	Bisulfitseq.	n.a.	n.a.
36	1146	Non-Responder	PD	pos	116	75	T1	0	M0	selten	I	Bisulfitseq.	n.a.	n.a.
58	343	Non-Responder	PD	neg	0	0	T2	0	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
59	1159	Non-Responder	PD	neg	3	0	T2	0	M0	DCIS+IDC	IIa	Bisulfitseq.	n.a.	n.a.
60	1161	Non-Responder	PD	neg	3	69	T2	0	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
63	1186	Non-Responder	PD	neg	4	0	T1	1-3	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
31	353	fraglich	SD	neg	0	0	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
37	1147	fraglich	SD	pos	56	0	T2	0	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
38	1148	fraglich	SD	neg	0	0	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
39	1152	fraglich	SD	pos	48	83	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
40	1154	fraglich	SD	neg	0	8	T1	0	M0	TUB	I	Bisulfitseq.	n.a.	n.a.
45	1174	fraglich	SD	neg	4	7	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
47	1179	fraglich	SD	pos	84	504	T2	0	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
49	1185	fraglich	SD	pos	20	47	T1	0	M0	DCIS+IDC	I	Bisulfitseq.	n.a.	n.a.
50	1187	fraglich	SD	pos	12	12	T1	0	M0	IDC	I	Bisulfitseq.	n.a.	n.a.
51	1193	fraglich	SD	neg	0	0	T2	>3	M0	ILC	IIb	Bisulfitseq.	n.a.	n.a.
13	1184	fraglich	SD	neg	5	0	T2	0	M0	IDC	IIa	Bisulfitseq.	n.a.	n.a.
3	1194	fraglich	SD	neg	0	8	T2	0	M0	n.a.	IIa	Bisulfitseq.	n.a.	n.a.

Tabelle 14: Proben die zusätzlich bisulfitsequenziert wurden.

5.9. Listen für Kandidaten der Tumormarker

Tabelle 15: Kandidatenliste für Hypermethylierung in den Tumoren von in gesundem Gewebe hypomethylierten Fragmenten

Fragment	Mittlere Methylierung gesunden		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel		TSS	Gen	Gen Name
	Gewebe	Blut		methyliert	semi- methyliert	un- methyliert		-	+			
10953	11%	34%	7,0	7	13	8	12q24.13	-	-		ENSG00000135144	DTX1
3944	2%	91%	6,9	12	10	6	16q12.2	-	-		-	-
4563	29%	202%	6,5	25	3	0	6p22.1	-	-		ENSG00000096654	ZNF184
2376	28%	186%	6,1	26	1	1	3p14.1	-	ENSG00000163378	ENSG00000163378	C3orf64	
2097	-10%	26%	5,9	5	10	13	2p13.3	-	-		-	-
4233	24%	77%	5,6	8	18	2	6q13	+	-		-	-
4493	20%	71%	5,5	15	9	4	17p13.1	-	ENSG00000189353	ENSG00000189353	novel trans	
11604	22%	78%	5,4	10	15	3	1q25.3	+	ENSG00000143324	ENSG00000143324	XPR1	
11671	22%	101%	5,3	24	4	0	1p32.3	-	-		ENSG00000169174	PCSK9
9713	27%	51%	5,3	4	24	0	14q24.3	-	ENSG00000140043	ENSG00000140043	ZADH1	
7524	14%	28%	5,2	8	16	4	12p12.1	+	-		ENSG00000060982	BCAT1
5969	16%	29%	4,8	0	25	3	12q23.3	+	ENSG00000139420	-	-	Q9BVA6
4235	12%	20%	4,7	4	19	5	2q32.1	+	-		-	-
656	15%	29%	4,5	11	12	5	5q32	-	-		-	-
11509	26%	45%	4,5	12	14	2	12q23.1	-	-		ENSG00000185046	ANKS1B
12787	26%	17%	4,4	14	12	2	15q25.2	+	-		-	-
2180	23%	34%	4,4	6	16	6	10q21.3	-	-		ENSG00000183230,E NSG00000181858	CTNNA3
10136	-8%	14%	4,2	5	13	10	12q12	+	-		-	-
10422	22%	50%	4,0	10	13	5	14q13.3	+	ENSG00000165414	-	-	Q96HD6_HUMAN
6204	24%	32%	3,9	0	26	2	10p11.22	+	ENSG00000120616	ENSG00000120616	EPC1	
1674	19%	66%	3,9	8	20	0	10p13	-	ENSG00000148481	ENSG00000148481	C10orf197	
1678	23%	54%	3,9	9	19	0	8p11.23	+	ENSG00000168619	ENSG00000168619	ADAM18	
4148	28%	64%	3,9	8	20	0	7q32.1	-	-		-	-
4416	11%	39%	3,8	8	19	1	3q26.2	-	-		-	-
4664	27%	19%	3,6	8	7	13	7p15.2	-	ENSG00000105996	ENSG00000105996	HOXA2	
2043	22%	35%	3,6	2	13	13	9q34.12	+	ENSG00000130711	ENSG00000130711	PRDM12	
9857	22%	17%	3,6	6	10	12	5q22.3	+	ENSG00000129596	ENSG00000129596	CDO1	
1665	22%	40%	3,6	5	18	5	Xq11.2	-	-		-	-
12406	22%	27%	3,4	1	24	3	1q22	+	ENSG00000132677	ENSG00000132677	RHBG	
2272	29%	39%	3,3	11	12	5	18q23	+	ENSG00000151514	ENSG00000151514	SALL3	

Fragment	Mittlere Methylierung gesunden		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel	TSS	Gen	Gen Name
	Gewebe	Blut		methyliert	semi- methyliert	un- methyliert					
12682	28%	46%	3,3	5	22	1	10q23.32	+	ENSG00000180628	ENSG00000180628	PCGF5
11770	29%	39%	3,3	6	18	4	7q33	-	ENSG00000105894	ENSG00000105894	PTN
1523	-15%	22%	3,2	1	9	18	14q23.1	-	-	-	-
598	-26%	3%	3,2	1	7	20	12q12	-	-	-	-
4392	13%	-10%	3,2	12	14	2	Xq26.1	-	ENSG00000056277	ENSG00000056277	SUHW3
11545	26%	84%	3,1	14	14	0	8q24.3	-	ENSG00000180838	ENSG00000167653	PSCA und Q8NAM3
5422	27%	47%	3,0	0	21	7	7p22.1	-	-	ENSG00000164539	novel trans
4569	24%	33%	2,9	0	22	6	17p13.1	+	ENSG00000183011, ENSG00000182224	ENSG00000183011, ENSG00000182224	CYB5D1; LSMD1
11084	22%	37%	2,9	0	16	12	22q12.2	+	-	-	-
74	26%	30%	2,9	15	12	1	13q31.1	-	-	ENSG00000178235	SLITRK1
3712	18%	39%	2,8	3	16	9	6q21	-	-	-	-
8960	4%	22%	2,8	0	14	14	13q12.3	+	ENSG00000132952	-	USPL1
178	18%	54%	2,7	5	18	5	18q11.2	+	ENSG00000101773	ENSG00000101773	RBBP8
11968	14%	29%	2,7	0	10	18	3p21.31	-	-	-	-
7508	27%	39%	2,7	1	16	11	17q11.2	+	ENSG00000108671	-	PSMD11
3350	15%	27%	2,7	0	8	20	19p13.2	+	ENSG00000105518	ENSG00000105518	novel gene
3704	21%	32%	2,7	5	17	6	22q13.1	-	-	-	-
8998	20%	48%	2,6	1	19	8	1p36.12	+	ENSG00000117318	ENSG00000117318	ID3
13656	27%	49%	2,6	4	20	4	11q24.2	+	ENSG00000187686	-	novel gene
12401	27%	24%	2,6	8	11	9	4q21.21	-	-	-	-
1322	16%	23%	2,6	3	10	15	19q13.33	-	-	-	-
3869	17%	37%	2,6	4	16	8	3q26.31	-	-	ENSG00000169760	NLGN1
10666	26%	36%	2,6	1	24	3	1p36.22	-	-	ENSG00000175262	C1orf127
2125	29%	51%	2,5	0	25	3	3q23	+	ENSG00000114125	ENSG00000114125	RNF7
2764	24%	30%	2,5	1	14	13	1p36.33	+	ENSG00000107404	ENSG00000107404	DVL1
10977	18%	56%	2,4	2	16	10	10q25.2	-	ENSG00000119927	-	GPAM
8242	13%	24%	2,4	0	8	20	1q23.2	-	ENSG00000143315	ENSG00000143315	PIGM
3027	20%	33%	2,4	2	19	7	9p21.1	-	-	-	-
10064	18%	37%	2,4	0	21	7	2p13.3	+	ENSG00000115977	-	AAK1
11681	16%	29%	2,3	0	22	6	1q21.1	+	ENSG00000131788	-	PIAS3

Fortsetzung Tabelle 15

Anhang

Fragment	Mittlere Methylierung gesunden		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel	TSS	Gen	Gen Name
	Gewebe	Blut		methyliert	semi- metyhliert	un- methyliert					
6660	24%	37%	2,3	0	23	5	1_NT_079 +	ENSG00000160679	ENSG00000160679	C1orf77	
13219	21%	26%	2,2	1	9	18	1_NT_077 +	ENSG00000143502	-	SUSD4	
9252	13%	26%	2,2	0	21	7	1p36.23 +	-	-	-	
4827	22%	32%	2,2	0	13	15	11p11.2 +	ENSG00000180423,E NSG00000175224	ENSG00000175224	KIAA0652 ; C11orf77	
8670	19%	20%	2,2	3	11	14	13q21.1 +	-	-	-	
7951	27%	59%	2,2	16	12	0	13q33.1 +	ENSG00000134899	ENSG00000134899	ERCC5	
5525	25%	41%	2,2	1	20	7	19q13.31 -	ENSG00000131115	ENSG00000131115	ZNF227	
12680	26%	55%	2,2	2	23	3	3q23 +	ENSG00000193482,E NSG00000184432	ENSG00000184432	COPB2	
690	30%	20%	2,1	1	19	8	6p24.3 -	-	-	-	
9861	27%	51%	2,1	4	21	3	14q23.3 -	ENSG00000126822	-	PLEKHG3	
13072	15%	37%	2,1	0	18	10	6p12.2 -	ENSG00000096092	ENSG00000096092	TMEM14A	
1866	11%	31%	2,1	0	11	17	1p34.2 +	ENSG00000117399	-	CDC20	
4078	17%	32%	2,0	0	12	16	10p13 +	ENSG00000151465,E NSG00000165609	ENSG00000165609,E NSG00000151465	NUDT5; C10orf7	
869	12%	47%	2,0	15	5	8	Xq21.31 +	ENSG00000174740	ENSG00000174740	PABPC5	
9318	23%	30%	2,0	1	16	11	3q13.31 -	ENSG00000185565	ENSG00000185565	LSAMP	
3015	19%	23%	2,0	3	10	15	5q32 -	ENSG00000156463	ENSG00000156463	SH3RF2	
9466	23%	21%	2,0	3	5	20	7q11.21 +	ENSG00000183501	ENSG00000181342,E NSG00000183501	ZNF25/ZNF680	
3144	22%	43%	2,0	0	23	5	3p21.31 -	ENSG00000076201	ENSG00000076201	PTPN23	
10996	26%	48%	2,0	1	22	5	14q21.1 -	ENSG00000150527	ENSG00000150527	CTAGE5	
8295	18%	50%	2,0	2	20	6	5q22.2 +	ENSG00000047188	ENSG00000047188	YTHDC2	
8729	20%	52%	1,9	1	25	2	12q13.11 -	-	-	-	
2758	23%	25%	1,9	2	6	20	11q14.3 -	ENSG00000077616	ENSG00000077616	NAALAD2	
13165	27%	37%	1,9	4	17	7	2p15 -	ENSG00000173215	-	Q96HQ8	
12186	17%	27%	1,9	5	12	11	8q21.13 +	-	-	-	
1589	26%	38%	1,9	7	15	6	11q23.2 +	ENSG00000109906	-	ZBTB16	
6151	24%	33%	1,9	0	20	8	4q21.21 +	ENSG00000138759	ENSG00000138759	FRAS1	
10056	29%	21%	1,9	6	6	16	Xq22.1 +	ENSG00000102362	ENSG00000102362	SYTL4	
7729	11%	29%	1,9	6	8	14	1q25.3 -	ENSG00000058085	ENSG00000058085	LAMC2	

Fortsetzung Tabelle 15

Fragment	Mittlere Methylierung gesunden		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel	TSS	Gen	Gen Name
	Gewebe	Blut		methyliert	semi- methyliert	un- methyliert					
12240	27%	47%	1,9	3	25	0	15q21.1	+	ENSG00000171763	-	SPATA5L1
4280	13%	19%	1,8	0	9	19	3q21.1	-	ENSG00000180501	ENSG00000180501	WDR5
506	23%	31%	1,8	0	16	12	19p13.2	-	ENSG00000105576	ENSG00000105576	TNPO2
6132	24%	93%	1,8	5	17	6	19q13.42	-	ENSG00000189430	ENSG00000189430	NCR1
9195	23%	42%	1,8	1	23	4	11q21	+	ENSG00000077458,E NSG00000166037	ENSG00000166037	FAM76B; CEP57
11959	27%	67%	1,8	10	13	5	6p22.2	-	ENSG00000111802,E NSG00000112304	ENSG00000112304	TTRAP and THEM2
7472	21%	32%	1,8	3	19	6	4p15.2	+	-	-	-
12750	4%	28%	1,7	0	7	21	7p22.1	-	ENSG0000008256	ENSG0000008256	PSCD3
10883	29%	52%	1,7	3	22	3	1q23.3	+	ENSG00000152332	ENSG00000152332	UHMK1
13365	22%	52%	1,7	7	19	2	12q24.21	-	-	ENSG00000089225	TBX5
1458	26%	36%	1,6	3	20	5	10q24.2	-	ENSG00000119919	-	NKX2-3
9863	23%	23%	1,6	0	14	14	20q12	+	ENSG00000124183	ENSG00000124183	TOP1
12104	23%	23%	1,6	4	8	16	5q22.3	-	-	-	-
3610	25%	28%	1,6	0	16	12	10q24.1	+	-	-	-
10594	25%	43%	1,6	1	20	7	5q35.2	-	ENSG00000122203	ENSG00000122203	KIAA1191
7696	28%	43%	1,6	6	15	7	6q22.33	-	-	-	-
11129	29%	56%	1,5	15	9	4	17q21.32	+	-	-	-
13196	8%	23%	1,5	0	10	18	Xp11.3	-	ENSG00000147050	ENSG00000147050	UTX
6947	24%	28%	1,5	3	16	9	2p13.2	-	-	ENSG00000135638	EMX1
10512	28%	31%	1,5	4	10	14	8q12.3	+	ENSG00000180828	ENSG00000180828	BHLHB5
12213	28%	45%	1,5	1	22	5	19q13.2	+	-	ENSG00000079435	LIPE
8752	18%	38%	1,4	0	15	13	7q22.3	-	ENSG00000005483	ENSG00000005483	MLL5
8486	29%	51%	1,4	3	17	8	2q24.2	+	-	-	-
10183	2%	25%	1,4	0	10	18	2p23.3	-	ENSG00000115204	-	MPV17
7882	20%	40%	1,4	0	15	13	22q12.2	+	-	-	-
11619	20%	33%	1,4	0	16	12	7p15.1	-	ENSG00000106080,E NSG00000106086	-	FKBP14; PLEKHA8
11701	19%	30%	1,4	0	21	7	7q11.21	+	ENSG00000067601,E NSG00000106610	ENSG00000067601,E NSG00000106610	PMS2L4; NM_022906
373	19%	34%	1,4	0	17	11	11q22.3	-	ENSG00000152578	ENSG00000152578	GRIA4

Fortsetzung Tabelle 15

Fragment	Mittlere Methylierung gesunden		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel	TSS	Gen	Gen Name
	Gewebe	Blut		methyliert	semi- metyhliert	un- methyliert					
7780	23%	35%	1,4	8	13	7	12q13.2	-	-	-	-
10357	10%	21%	1,4	0	11	17	20p12.1	+	ENSG00000125864	ENSG00000125864	BFSP1
315	27%	32%	1,4	1	16	11	4q25	-	ENSG00000155011	ENSG00000155011	DKK2
4791	24%	36%	1,4	0	19	9	1q32.1	+	ENSG00000163444	ENSG00000163444	TMEM183B
13370	23%	32%	1,4	0	15	13	17_NT_07	+	ENSG00000182481	ENSG00000182481	KPNA2
										ENSG00000138303,E	
7152	22%	31%	1,3	0	19	9	10q22.1	+	ENSG00000166295	NSG00000166295	ASCC1; C10orf104
11869	28%	60%	1,3	3	19	6	1q24.3	+	ENSG00000117523	ENSG00000117523	BAT2D1
12496	13%	33%	1,3	0	11	17	6q21	+	ENSG00000057663	ENSG00000057663	ATG5
10278	25%	35%	1,3	0	16	12	10q25.2	-	ENSG00000108055	-	SMC3
4385	5%	20%	1,3	0	10	18	6p25.2	-	ENSG00000137267	ENSG00000137267	TUBB2A
356	24%	34%	1,3	0	21	7	17q12	+	ENSG00000108292	-	MLLT6
12353	25%	26%	1,3	0	18	10	8q11.23	+	-	-	-
5770	18%	57%	1,3	3	16	9	1p36.11	-	ENSG00000133226	ENSG00000133226	SRRM1
8481	26%	36%	1,3	0	12	16	2p23.3	-	ENSG00000138018	ENSG00000138018	SELI_HUMAN
5319	12%	30%	1,3	0	9	19	15q15.2	+	ENSG00000128881	ENSG00000128881	TTBK2
5202	28%	29%	1,3	2	15	11	10p11.23	-	ENSG00000107968	ENSG00000107968	MAP3K8
10354	23%	31%	1,3	0	13	15	11q22.1	+	ENSG00000187151	ENSG00000187151	ANGPTL5
4067	27%	29%	1,2	2	10	16	11q14.2	+	ENSG00000123892	ENSG00000123892	RAB38
11604	9%	20%	1,2	0	13	15	1q25.3	+	ENSG00000143324	ENSG00000143324	XPR1
4884	19%	24%	1,2	0	8	20	2q33.3	+	ENSG00000118257	ENSG00000118257	NRP2
3615	28%	30%	1,2	3	9	16	14q24.3	+	ENSG00000165553	ENSG00000165553	NGB
4777	25%	28%	1,1	0	14	14	18q21.33	+	ENSG00000081913	-	PHLPP
										ENSG00000168415,E	
6668	26%	65%	1,1	2	20	6	6p21.32	-	ENSG00000168468	NSG00000168468	ATF6B_HUMAN
6196	16%	30%	1,1	0	10	18	3p22.1	+	ENSG00000168036	-	CTNNB1
7295	24%	25%	1,1	0	9	19	12q13.12	+	ENSG00000169884	ENSG00000169884	Wnt10B
2357	26%	28%	1,1	0	17	11	17q25.1	-	ENSG00000177728	ENSG00000177728	KIAA0195
2577	21%	22%	1,1	1	8	19	6p12.1	-	ENSG00000137252	ENSG00000137252	HCRTR2
										ENSG00000153188,E	
2580	21%	18%	1,1	2	20	6	1q44	-	NSG00000179576	-	novel gene
12334	24%	29%	1,1	1	14	13	8q24.22	-	ENSG00000155897	ENSG00000155897	ADCY8

Fortsetzung Tabelle 15

Anhang

Fragment	Mittlere Methylierung		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel	TSS	Gen	Gen Name
	gesunden Gewebe	Blut		methyliert	semi- methyliert	un- methyliert					
1262	1,210	0,0	-44,4	0	0	28	2p21	-	-	-	-
9395	0,744	0,2	-18,2	0	0	28	22q13.31	+	-	ENSG00000186654	ARHGAP8
3533	0,753	-0,1	-17,0	0	1	27	6p21.31	+	-	-	-
10173	0,887	0,6	-16,3	0	12	16	16p13.3	-	ENSG00000127554	ENSG00000127554	GFER
1691	1,144	0,2	-12,3	0	0	28	2p25.1	+	ENSG00000151694	-	ADAM17
10088	0,925	0,0	-10,9	0	0	28	10q25.3	-	-	-	-
9761	0,891	0,1	-10,7	0	0	28	16p13.2	-	-	-	-
13178	1,031	0,9	-10,1	15	13	0	6q15	-	-	-	-
9997	0,970	0,8	-10,1	13	15	0	12q13.11	-	-	ENSG00000174233	ADCY6
5180	1,332	0,5	-9,8	1	3	24	3p25.3	-	-	ENSG00000156983	ARPC4
8493	0,783	0,5	-9,8	0	16	12	Xp11.23	+	ENSG00000147155	ENSG00000147155	EBP
6925	1,107	0,6	-9,4	0	8	20	3p21.1	-	-	ENSG00000157445	CACNA2D3
3206	0,822	0,1	-8,7	0	4	24	12p13.32	-	ENSG00000139180	ENSG00000139180	NDUFA9
8937	0,882	0,7	-8,7	2	26	0	10q26.3	+	ENSG00000151650	ENSG00000151650	VENTX2
689	0,739	0,3	-8,5	0	12	16	21q21.3	-	-	ENSG00000142192	APP
4349	0,744	0,2	-8,4	0	3	25	11q12.1	-	ENSG00000110042	ENSG00000172324,E NSG00000110042	DTX4; OR5A2
4659	1,140	0,5	-8,3	0	14	14	11q13.1	+	ENSG00000146670,E NSG00000162300	ENSG00000168065,E NSG00000162300	CDCA5; SLC22A11; ZFPL1
1994	0,936	0,7	-8,2	8	20	0	21q22.3	-	-	ENSG00000160194	NDUFV3
7041	1,062	0,7	-7,6	18	10	0	1q23.1	-	-	ENSG00000117029	NTRK1
8600	0,726	0,7	-7,1	3	25	0	1p36.12	-	-	ENSG00000158803	novel gene
13772	0,776	0,5	-6,9	1	16	11	11p13	+	ENSG00000149100	ENSG00000149100	PCID1
11299	1,272	0,7	-6,7	13	13	2	12q24.33	-	-	-	-
13118	0,946	0,9	-6,5	19	9	0	1p36.21	-	ENSG00000181952	ENSG00000189337	Q8N4O9; Q674X7
10276	0,873	0,5	-6,4	1	20	7	6p21.1	+	ENSG00000164663	ENSG00000164663	USP49
8038	0,830	0,8	-6,4	8	19	1	19q13.33	-	-	ENSG00000142530	Q9BSM9
8156	1,438	-0,1	-6,2	9	6	13	2p13.3	-	-	-	-
798	1,109	0,4	-5,9	8	11	9	1q42.3	-	-	ENSG00000168243	GNG4
7251	0,986	0,7	-5,0	15	13	0	11p11.2	-	ENSG00000187181	ENSG00000166199,E NSG00000187181	ALKBH3
2382	0,845	-0,1	-4,9	5	12	11	19q13.43	-	ENSG00000176593	ENSG00000176293,E NSG00000166704,EN SG00000176593	Q8N9G5; ZNF135; ZNF606
13491	0,907	0,4	-4,7	0	4	24	7p22.3	-	-	ENSG00000106266	SNX8

Tabelle 16: Kandidatenliste für Hypomethylierung in den Tumoren von in
gesundem Gewebe hypermethylierten Fragmenten

Fragment	Mittlere Methylierung		T- Statistik	Tumorproben pro Methylierungsgruppe			Chr	CpG Insel	TSS	Gen	Gen Name
	gesunden Gewebe	Blut		methyliert	semi- metyhliert	un- methyliert					
								ENSG00000083520,E			
9239	0,714	0,6	-4,7	1	20	7	13q22.1	+	NSG00000083535	ENSG00000083520	DIS3; C13orf24
9687	0,881	0,6	-4,7	0	17	11	2q37.1	-	-	ENSG00000144535	DIS3L2
1869	0,799	0,7	-4,5	2	24	2	16q23.2	-	-	-	-
4651	0,930	0,9	-4,4	18	10	0	8p21.1	-	-	ENSG00000104299	Q9NV88
8844	0,962	0,6	-4,3	4	20	4	1p34.2	+	ENSG00000117399	ENSG00000117399	CDC20
9427	0,756	0,5	-4,3	2	22	4	19q13.42	+	ENSG00000167615	ENSG00000167615	LENG8
2321	1,066	0,6	-4,2	5	14	9	11p11.2	-	ENSG00000165923	ENSG00000165923	AGBL2
7782	0,748	0,5	-4,0	0	17	11	4q22.1	+	ENSG00000180346	-	TIGD2
9916	0,735	0,9	-4,0	2	26	0	1q32.2	-	-	-	-
4575	1,059	0,1	-3,9	4	7	17	19p13.3	-	-	-	-
2305	0,851	0,5	-3,9	1	10	17	11p15.4	-	-	ENSG00000110319	SH2BP1
676	0,764	0,7	-3,8	6	19	3	21q22.11	-	-	-	-
4196	0,818	0,5	-3,8	1	11	16	3q21.2	-	-	ENSG00000121531	novel gene
9870	0,750	0,6	-3,7	1	21	6	4q12	-	ENSG00000134852	-	CLOCK
3194	0,898	0,8	-3,7	12	13	3	Xp22.11	-	ENSG00000184735	ENSG00000184735	DDX53
2105	0,831	0,7	-3,6	14	14	0	16q24.2	-	-	ENSG00000103264	FBXO31
2141	0,733	0,5	-3,5	0	5	23	17p13.3	-	ENSG00000127804	ENSG00000127804	METT10D
8485	0,782	0,7	-3,5	6	18	4	17q12	-	-	ENSG00000178655,E	
640	0,830	-0,2	-3,5	7	12	9	3q29	+	-	-	MRM1
361	1,266	0,6	-3,4	9	15	4	2p11.2	+	ENSG00000115525	ENSG00000115525	ST3GAL5
4792	1,190	0,7	-3,3	17	10	1	5q31.2	-	ENSG00000112981	ENSG00000112981	NME5
3712	0,835	0,6	-3,3	0	15	13	6q21	-	-	-	-
12337	0,786	0,7	-3,2	6	22	0	3p14.3	-	-	ENSG00000163947	ARHGEF3
11909	0,759	0,5	-3,2	0	4	24	7q22.1	-	-	ENSG00000106327	TRF2
6593	0,814	0,7	-3,2	5	23	0	20q11.22	-	-	-	-
12231	0,913	0,7	-3,2	3	21	4	2q33.2	+	ENSG00000163596	ENSG00000163596	ICA1L
612	0,748	0,5	-3,1	2	25	1	17q21.31	+	ENSG00000186834	ENSG00000186834	HEXIM1
1667	0,847	0,7	-3,1	2	14	12	3q22.1	-	-	-	-
8019	0,985	0,8	-3,1	18	10	0	2q14.3	-	-	-	-
3833	0,812	0,5	-3,0	0	18	10	21q22.2	-	ENSG00000183412	-	C21orf104
12176	0,824	0,3	-3,0	9	15	4	16p13.3	-	-	ENSG00000103253,E NSG00000103245	HAGHL; NARFL

Fortsetzung Tabelle 16

Fragment	Mittlere Methylierung		T- Statistik	Tumorproben pro			Chr	CpG		Gen	Gen Name
	gesunden Gewebe	Blut		methyliert	semi- methyliert	un- methyliert		Insel	TSS		
8532	0,801	0,7	-3,0	2	20	6	4p16.3	-	-	-	-
5153	0,791	0,1	-3,0	2	11	15	14q32.11	-	-	-	-
4196	0,826	0,8	-3,0	8	19	1	3q21.2	-	-	ENSG00000121531	novel gene
11370	0,706	0,2	-2,9	0	17	11	Xq26.1	-	-	ENSG00000156697	UTP14A
1265	0,784	0,7	-2,9	10	18	0	15q13.1	-	-	ENSG00000104067	TJP1
11730	0,941	0,9	-2,9	15	13	0	17q21.2	-	-	ENSG00000131473	ACLY
2972	1,139	0,7	-2,8	15	13	0	6p21.1	-	-	ENSG00000137218	FRS3
5975	0,711	0,6	-2,8	2	21	5	21q22.11	-	-	-	-
11117	1,443	0,4	-2,8	13	8	7	5q23.2	-	-	ENSG00000145794	MEGF10
714	0,703	0,8	-2,8	0	6	22	5p15.1	-	ENSG00000185296	ENSG00000185296	novel gene
12343	1,5	1,2	-2,7	16	10	2	3p22.2	-	-	ENSG00000185313	SCN10A
5475	0,877	0,9	-2,7	14	14	0	12p13.31	-	-	ENSG00000166535	A2ML1
3019	0,992	0,7	-2,7	16	11	1	1q44	-	-	ENSG00000162711	NLRP3
2796	0,778	0,3	-2,7	4	10	14	Xq28	+	ENSG00000147403	ENSG00000147403	RPL10
3665	0,848	0,7	-2,6	5	21	2	21q22.11	-	-	ENSG00000142183	ITSN1
12156	0,805	0,5	-2,6	12	9	7	20p11.21	-	-	ENSG00000101463	C20orf39
11003	0,777	0,5	-2,6	1	12	15	1p31.2	+	-	-	-
5971	0,719	0,6	-2,6	6	17	5	16p11.2	+	-	ENSG00000180209	MYLPF
6734	0,795	1,0	-2,5	10	18	0	11q13.1	-	-	-	-
762	0,947	0,6	-2,5	16	12	0	10p15.3	-	-	ENSG00000185736	ADARB2
11721	0,712	0,5	-2,4	0	18	10	2q37.2	-	-	ENSG00000157985	CENTG2
1297	0,959	0,9	-2,4	7	19	2	11q24.1	-	-	ENSG00000109944	C11orf36
6233	0,703	0,5	-2,4	5	23	0	7q32.2	+	-	ENSG00000158623	TSGA13
12294	0,825	0,6	-2,4	3	18	7	6p21.31	-	ENSG00000112079	ENSG00000112079	STK38
9138	1,160	0,6	-2,3	18	10	0	Xp22.31	-	-	ENSG00000130021	HDHD1A
479	0,795	0,3	-2,3	6	22	0	7q22.1	-	ENSG00000160991	-	TMEM142B
13391	0,905	0,7	-2,3	12	14	2	14q32.12	+	ENSG00000089723	ENSG00000089723	OTUB2
12934	0,828	0,7	-2,3	12	15	1	2q37.1	-	-	-	-
9871	0,830	0,8	-2,3	13	14	1	1p22.1	-	ENSG00000137948	ENSG00000137948	BRDT
12885	0,701	0,5	-2,2	2	21	5	17q25.1	+	ENSG00000132471	ENSG00000132471	WBP2
10535	0,990	0,9	-2,2	17	11	0	4q31.21	-	-	ENSG00000170153	RNF150
2090	0,720	0,5	-2,2	2	20	6	4q28.2	+	ENSG00000151466,E	ENSG00000151466	SCLT1; C4orf33
6716	0,713	0,6	-2,2	5	23	0	1q25.3	-	ENSG00000181014	-	SMG7
6687	1,024	0,8	-2,2	11	17	0	17p13.3	-	-	-	-
7126	1,1	0,6	-2,1	14	11	3	1q24.2	-	-	ENSG00000143178	TBX19

Fortsetzung Tabelle 16

Anhang

Fragment	Mittlere Methylierung gesunden		T- Statistik	Tumorproben pro			Chr	CpG Insel	TSS	Gen	Gen Name
	Gewebe	Blut		methyliert	semi- methyliert	un- methyliert					
10876	0,728	0,8	-2,1	5	23	0	2q32.2	+	ENSG00000115404	ENSG00000115404	novel gene
5417	0,718	0,5	-2,1	11	17	0	Xq12	+	ENSG00000181704	-	YIPF6
11332	0,900	0,8	-2,1	14	14	0	16q21	-	-	-	-
12844	0,754	0,6	-2,1	4	19	5	11q23.2	-	ENSG00000086827	ENSG00000086827	ZW10
11290	0,797	1,0	-2,1	10	16	2	18q12.3	-	-	ENSG00000132874	SLC14A2
761	0,838	1,1	-2,1	5	10	13	16q24.2	-	-	-	-
6036	0,915	0,8	-2,0	14	14	0	5q35.3	-	ENSG00000146063	-	TRIM41
8708	0,931	1,0	-2,0	18	10	0	2q14.3	-	-	ENSG00000183300	novel gene
5968	1,3	0,4	-2,0	15	7	6	22q13.2	-	ENSG00000100380,E NSG00000100382	ENSG00000100380	ST13
1805	0,739	0,4	-2,0	1	19	8	4q25	-	ENSG00000138804,E NSG00000192257	-	Q9H5S2
4645	0,898	0,9	-2,0	12	16	0	15q15.1	-	ENSG00000168907	ENSG00000168907	PLA2G4F
12024	0,953	0,9	-2,0	11	17	0	16q24.2	-	-	ENSG00000179588	ZFPM1
858	0,785	0,6	-2,0	7	21	0	20q13.12	-	-	ENSG00000185513	L3MBTL
11653	0,883	1,1	-2,0	19	9	0	4q31.22	-	-	ENSG00000151617	EDNRA
7428	0,792	0,7	-2,0	3	23	2	20q13.13	+	ENSG00000000419,E NSG00000124217	ENSG00000000419	DPM1; MOCS3
3757	1,0	0,6	-1,9	10	12	6	Xq13.2	-	ENSG00000186462	ENSG00000186462	NAP1L2
9032	0,711	0,4	-1,9	7	10	11	Xp11.4	+	ENSG00000156313	ENSG00000156313	RPGR
1996	0,785	0,7	-1,9	5	21	2	7p22.1	+	-	ENSG00000164916	FO XK1
4075	0,865	0,5	-1,9	4	16	8	17p13.1	-	-	ENSG00000133026	MYH10
5581	0,706	0,7	-1,9	6	22	0	15q21.1	+	ENSG00000104177	ENSG00000104177	MYEF2
12695	0,844	0,6	-1,9	12	16	0	11p14.1	-	-	-	-
5693	0,725	0,6	-1,9	1	17	10	6q24.1	+	ENSG00000164442	-	CITED2
2808	0,857	0,9	-1,8	14	14	0	5q33.3	-	-	ENSG00000135083	CCNJL
7048	1,075	0,8	-1,8	20	8	0	Xq24	-	-	-	-
9748	0,842	0,5	-1,8	7	18	3	13q13.1	-	ENSG00000139597	ENSG00000139597,E NSG00000183037	NM_052818
8602	0,726	0,7	-1,8	5	23	0	7q21.3	-	-	ENSG00000158560	DYNC111
5430	0,822	0,9	-1,8	12	16	0	19p13.2	-	-	ENSG00000171105	INSR
1261	0,906	0,8	-1,8	19	9	0	7p22.3	-	-	ENSG00000146540	NM_032350
5608	0,852	0,7	-1,7	16	12	0	16q24.2	-	-	ENSG00000103257	SLC7A5
11411	0,887	0,7	-1,7	11	17	0	12q13.13	-	-	ENSG00000123395	C12orf44

Fortsetzung Tabelle 16

Fragment	Mittlere Methylierung		T- Statistik	Tumorproben pro			Chr	CpG		Gen	Gen Name
	gesunden Gewebe	Blut		methyliert	semi- methyliert	un- methyliert		Insel	TSS		
10491	0,745	0,5	-1,7	0	18	10	17p12	+	-	-	-
4641	0,822	0,7	-1,7	4	15	9	5p15.33	-	-	-	-
8683	1,1	0,9	-1,7	12	11	5	17q21.2	-	-	ENSG00000131759	RARA
1931	0,706	0,6	-1,7	6	18	4	1p13.1	+	ENSG00000134263	ENSG00000134263	SYCP1
8236	0,801	0,6	-1,7	8	20	0	4q24	-	-	-	-
12973	0,797	0,7	-1,7	5	21	2	6p21.1	-	ENSG00000137225	ENSG00000137225	CAPN11
2584	0,802	0,9	-1,7	14	14	0	10q26.11	-	-	ENSG00000165646	SLC18A2
755	0,796	0,7	-1,6	15	13	0	6p21.31	+	ENSG00000065060	ENSG00000065060	C6orf107
10852	0,721	0,5	-1,6	3	22	3	3q25.33	-	-	-	-
7383	0,827	0,7	-1,6	10	18	0	18q11.2	+	ENSG00000167088	ENSG00000167088	SNRPD1
68	1,085	0,9	-1,5	19	9	0	3q25.2	+	ENSG00000174948	ENSG00000174948	GPR149
1359	0,904	1,0	-1,5	17	11	0	14q24.3	-	-	-	-
8551	0,911	0,8	-1,5	17	11	0	2q35	+	-	ENSG00000127838	PNKD
3371	0,988	1,1	-1,5	20	8	0	2q35	-	ENSG00000072182	ENSG00000072182	ACCN4
6861	0,830	0,7	-1,5	16	12	0	11p15.4	-	-	-	-
4208	0,759	0,6	-1,5	16	12	0	11q24.2	+	ENSG00000110063	ENSG00000110063	DCPS
6342	0,805	0,8	-1,5	12	16	0	8p22	-	-	ENSG00000104763	ASAH1
5758	0,925	0,9	-1,5	19	9	0	2q37.1	-	-	ENSG00000144481	TRPM8
2158	0,848	0,9	-1,5	5	18	5	11q23.3	-	-	ENSG00000137731,E	
2987	0,764	0,8	-1,5	12	16	0	3p21.2	+	ENSG00000164086	NSG00000137746	FXYP2; C11X
6670	0,871	0,8	-1,4	13	9	6	20q13.32	-	-	-	DUSP7
6887	0,802	0,6	-1,4	3	21	4	11q13.4	-	ENSG00000137522	ENSG00000124224	PPP4R1L
9756	0,763	0,9	-1,4	14	14	0	1q25.3	-	-	-	RNF121
3810	1,024	0,8	-1,4	9	17	2	Xp22.13	-	-	ENSG00000044446	-
8925	0,804	0,2	-1,4	11	11	6	20q13.33	-	-	ENSG00000101150	PHKA2
3003	0,862	0,7	-1,4	17	11	0	16q22.1	-	-	ENSG00000183723	TPD52L2
9294	0,821	0,8	-1,4	10	17	1	1p36.22	-	-	ENSG00000171621	CMTM4
8278	0,915	0,8	-1,4	16	12	0	1p36.13	-	-	-	SFSB1
11013	0,739	0,7	-1,4	6	13	9	2p24.1	-	-	ENSG00000115884	-
3606	0,832	0,9	-1,4	9	19	0	12q24.33	-	-	ENSG00000177192	SDC1
9533	0,876	0,6	-1,4	8	14	6	6q25.3	+	ENSG00000122335	ENSG00000177192	PUS1
6686	1,096	0,1	-1,4	17	3	8	12p13.32	+	ENSG00000111224	ENSG00000122335	SERAC1
7383	0,796	0,8	-1,4	11	17	0	18q11.2	+	ENSG00000167088	ENSG00000111224	PARP11
										ENSG00000167088	SNRPD1

Fortsetzung Tabelle 16

Anhang

Fragment	Mittlere Methylierung		T- Statistik	Tumorproben pro			Chr	CpG Insel	TSS	Gen	Gen Name
	gesunden Gewebe	Blut		methyliert	semi- methyliert	un- methyliert					
6545	0,932	0,8	-1,3	15	13	0	1q32.1	-	-	ENSG00000170498	KISS1
3856	0,767	0,6	-1,3	8	17	3	8q21.3	-	-	ENSG00000147613	PSKH2
										ENSG00000184678,E NSG00000184260,EN	HIST2H2AB; HIST2H2BE;
9835	0,751	0,6	-1,3	8	17	3	1q21.2	+	SG00000184270	ENSG00000184270	HIST2H2AC
5140	0,765	0,7	-1,3	8	20	0	14q12	+	ENSG00000129480	ENSG00000129480	C14orf126
4709	0,711	0,7	-1,3	5	22	1	3q26.1	+	ENSG00000169251	ENSG00000169251	NMD3
10567	0,940	0,9	-1,3	14	14	0	1p36.33	-	-	-	-
5969	0,726	0,6	-1,3	4	24	0	12q23.3	+	ENSG00000139420	ENSG00000139420	Q9BVA6
1632	0,795	1,0	-1,3	11	13	4	19p13.12	-	ENSG00000105085	ENSG00000105085	CRSP7
6686	0,813	0,9	-1,3	16	12	0	12p13.32	+	ENSG00000111224	ENSG00000111224	PARP11
7739	0,887	0,3	-1,3	9	18	1	12q21.1	-	ENSG00000111615	ENSG00000111615	KRR1
6921	0,797	-0,3	-1,3	13	6	9	11q13.4	-	ENSG00000149357	ENSG00000149357	C11orf59
2443	0,846	1,0	-1,3	13	13	2	8p23.1	-	-	-	-
6905	0,976	0,8	-1,3	13	15	0	13q14.2	-	-	ENSG00000136161	RCBTB2
7625	1,041	0,7	-1,3	8	14	6	10q11.21	-	-	ENSG00000107551	RASSF4
1170	0,784	0,7	-1,2	12	16	0	1p34.1	-	ENSG00000117448	-	AKR1A1
5863	0,715	0,6	-1,2	1	27	0	7p14.2	+	ENSG00000122545	-	SEPT7
5619	0,853	0,7	-1,2	10	18	0	7p11.2	-	-	ENSG00000185177	ZNF479
3673	0,753	0,9	-1,2	11	16	1	17q21.31	-	-	ENSG00000182963	GJA7 alias CX45
10466	1,010	0,9	-1,2	20	8	0	17q21.31	-	-	ENSG00000186868	MAPT
2312	0,795	0,6	-1,2	9	16	3	1q25.1	-	-	-	-
12099	0,849	0,7	-1,2	9	18	1	20q11.22	+	-	-	-
7191	0,831	0,7	-1,2	12	16	0	4q22.2	-	ENSG00000179931	ENSG00000179931	-
12867	0,840	-0,4	-1,2	5	10	13	12q13.13	-	-	-	-
10878	0,753	0,8	-1,2	10	18	0	10q21.1	-	-	-	-
12392	0,797	0,9	-1,2	19	9	0	17q23.1	-	-	-	-
5392	0,936	0,4	-1,2	11	17	0	3q24	-	ENSG00000175040	-	CHST2
7624	1,061	0,8	-1,2	20	8	0	1p36.12	-	-	ENSG00000117242	DDOST
12939	0,864	0,7	-1,2	6	19	3	17q25.3	-	ENSG00000055468	ENSG00000055468	DNAH17
										ENSG00000121716,E NSG00000078319	ENSG00000188146,E NSG00000078319
7513	0,782	0,7	-1,2	8	20	0	7q22.1	+	NSG00000078319	NSG00000078319	Q9UKJ0-3; PMS2L1
9837	0,748	0,6	-1,2	8	20	0	13q34	-	-	ENSG00000139842	CUL4A
5900	0,850	1,1	-1,1	11	7	10	20q11.23	-	-	ENSG00000080845	DLGAP4
10285	0,752	0,9	-1,1	13	14	1	4q31.3	-	-	ENSG00000145423	SFRP2
4160	0,732	0,9	-1,1	8	18	2	7q11.23	-	-	ENSG00000077800	FKBP6

Fortsetzung Tabelle 16

Anhang

Fragment	Mittlere Methylierung		T- Statistik	Methylierungsgruppe			Chr	CpG		Gen	Gen Name
	gesunden Gewebe	Blut		methyliert	semi- methyliert	un- methyliert		Insel	TSS		
417	0,702	0,6	-1,1	5	20	3	20p12.3	-	ENSG00000171984	ENSG00000171984	C20orf196
13296	0,704	0,4	-1,1	5	12	11	Xq12	-	-	ENSG00000169083	AR
7069	0,778	0,6	-1,1	12	16	0	10p12.2	-	-	-	-
12617	0,857	0,9	-1,1	15	13	0	1_NT_079	-	-	-	-
4740	0,786	0,6	-1,1	10	18	0	16p13.3	-	-	ENSG00000167977	KCTD5
9154	0,917	1,2	-1,1	13	15	0	13q34	+	-	-	-
3084	0,791	0,9	-1,1	15	13	0	7q22.1	+	ENSG00000146833	ENSG00000146833	TRIM4
12671	0,786	0,7	-1,0	3	20	5	17q21.32	-	ENSG00000182742	ENSG00000182742	HOXB4
3664	0,719	0,9	-1,0	8	19	1	7p22.1	-	-	-	-
2307	0,760	0,7	-1,0	9	19	0	3_NT_079	-	ENSG00000183662	ENSG00000183662	FAM19A1

Fortsetzung Tabelle 16

5.10. Listen für Kandidaten der Tumormarker

Tabelle 17: Kandidatenliste für das Ansprechen auf eine Chemotherapie von in gesundem Gewebe hypomethylierten Fragmenten

Fragment ID	T- Statistik	Responder pro Methylierungsgruppe			Non-Responder pro Methylierungsgruppe			Differenz methyliert	Res - Non methyliert	Chr	CpG Insel	TSS	Gen	Gen Name
		semi-methyliert	un-methyliert	un-methyliert	semi-methyliert	un-methyliert	un-methyliert							
1322	4,43	3	9	2	0	1	13	-11	3	19q13.33	-	NA	NA	No gene
1639	3,12	1	6	7	0	1	13	-6	1	19p13.2	+	NA	ENSG00000174652	ZNF266
5044	2,95	0	7	7	0	3	11	-4	0	7p14.1	+	ENSG00000065883	NA	CDC2L5
13370	2,83	0	11	3	0	4	10	-7	0	17_NT_07	+	ENSG00000182481	ENSG00000182481	KPNA2
5319	2,66	0	7	7	0	2	12	-5	0	15q15.2	+	ENSG00000128881	ENSG00000128881	TTBK2
2659	2,46	0	9	5	0	4	10	-5	0	4p16.3	+	ENSG00000127418	NA	FGFRL1
10156	2,39	0	11	3	0	2	12	-9	0	19q13.43	+	NA	ENSG00000127903,E	novel gene; ZNF582
13159	2,39	0	7	7	0	3	11	-4	0	1p21.2	+	ENSG00000156876,E	ENSG00000156876	SASS6; CCDC76
480	2,36	0	6	8	0	2	12	-4	0	6q14.1	+	ENSG00000118418	ENSG00000118418	HMG3
13569	2,30	0	9	5	0	4	10	-5	0	5p15.2	+	ENSG00000145491	ENSG00000145491	ROPN1L
6401	2,25	0	6	8	0	2	12	-4	0	6q25.1	+	ENSG00000181472	ENSG00000181472	ZBTB2
3502	2,24	0	9	5	0	5	9	-4	0	7q21.11	-	ENSG00000170381	ENSG00000170381	SEMA3E
1047	2,23	0	8	6	0	2	12	-6	0	4q21.22	-	ENSG00000138663	ENSG00000138663	COPS4
2403	2,21	0	8	6	1	2	11	-5	-1	8q13.2	+	NA	NA	No gene
5036	2,15	0	5	9	0	0	14	-5	0	17q25.3	-	ENSG00000169689,E	ENSG00000169683	STRA13; LRRC45
11968	2,15	0	8	6	0	2	12	-6	0	3p21.31	-	NA	NA	No gene
8242	2,13	0	6	8	0	2	12	-4	0	1q23.2	-	ENSG00000143315	ENSG00000143315	PIGM
13595	2,09	0	11	3	0	6	8	-5	0	11q23.3	+	ENSG00000175913,E	ENSG00000175913	MLL; Q9BRP9
7029	2,07	0	11	3	0	5	9	-6	0	5q22.1	+	ENSG00000164211	ENSG00000164211	STARD4
7250	2,04	0	9	5	0	5	9	-4	0	3p22.2	+	ENSG00000157036	ENSG00000157036	ENDOGL1
12749	2,02	1	11	2	0	7	7	-5	1	8p21.2	+	NA	NA	No gene
4402	2,01	0	11	3	0	6	8	-5	0	18p11.32	-	NA	NA	No gene
12799	1,97	1	9	4	0	4	10	-6	1	19q13.32	+	ENSG00000142252	ENSG00000142252	GEMIN7
3252	1,96	0	13	1	0	8	6	-5	0	16p13.12	+	ENSG00000103381	ENSG00000103381	CSTP1
13196	1,89	0	8	6	0	2	12	-6	0	Xp11.3	-	ENSG00000147050	ENSG00000147050	UTX
3615	1,89	2	7	5	1	2	11	-6	1	14q24.3	+	ENSG00000165553	ENSG00000165553	NGB
4235	1,87	3	11	0	1	8	5	-5	2	2q32.1	+	NA	NA	No gene
5035	1,83	0	7	7	0	3	11	-4	0	11p13	-	ENSG00000166016	NA	ABTB2
5485	1,78	0	11	3	0	4	10	-7	0	15q22.31	+	ENSG00000140455	NA	USP3
4885	1,77	0	11	3	0	7	7	-4	0	3p14.3	-	ENSG00000174844	ENSG00000174844	NM_198564
8505	1,74	2	4	8	0	2	12	-4	2	2p22.2	+	ENSG00000138061	ENSG00000138061	CYP1B1
3015	1,73	2	7	5	1	3	10	-5	1	5q32	-	ENSG00000156463	ENSG00000156463	SH3RF2
753	1,73	0	8	6	0	4	10	-4	0	1p13.2	-	ENSG00000081019	ENSG00000081019	RSBN1

Fragment ID	T-Statistik	Responder pro Methylierungsgruppe			Non-Responder pro Methylierungsgruppe			Differenz	Res - Non	Chr	CpG Insel	TSS	Gen	Gen Name
		semi-methyliert	un-methyliert	un-methyliert	semi-methyliert	un-methyliert	un-methyliert							
10053	1,68	1	3	10	0	0	14	-4	1	3q26.2	-	NA	NA	No gene
8752	1,68	0	10	4	0	5	9	-5	0	7q22.3	-	ENSG00000005483	ENSG00000005483	MLL5
8481	1,61	0	8	6	0	4	10	-4	0	2p23.3	-	ENSG00000138018	ENSG00000138018	SELL_HUMAN
6009	1,60	0	11	3	0	6	8	-5	0	17q23.2	+	ENSG00000062725	ENSG00000062725	APPBP2
1523	1,58	1	6	7	0	3	11	-4	1	14q23.1	-	NA	NA	No gene
1589	1,56	4	10	0	3	5	6	-6	1	11q23.2	+	ENSG00000109906	NA	ZBTB16
13165	1,56	4	6	4	0	11	3	1	4	2p15	-	ENSG00000173215	NA	Q96HQ8
1866	1,55	0	8	6	0	3	11	-5	0	1p34.2	+	ENSG00000117399	NA	CDC20
10183	1,54	0	7	7	0	3	11	-4	0	2p23.3	-	ENSG00000115204	NA	MPV17
9827	1,47	0	9	5	0	4	10	-5	0	6q21	+	ENSG00000112297	NA	AIM1
4569	1,44	0	13	1	0	9	5	-4	0	17p13.1	+	ENSG00000183011,E	ENSG00000183011,E	
5272	1,44	0	12	2	0	7	7	-5	0	11q13.1	+	NSG00000182224	NSG00000182224	CYB5D1; LSMD1
708	1,39	0	13	1	0	7	7	-6	0	1p22.2	+	ENSG00000097046	ENSG00000097046	VEGFB
8295	1,354	1	12	1	1	8	5	-4	0	5q22.2	+	ENSG00000047188	ENSG00000047188	CDC7
9863	1,338	0	9	5	0	5	9	-4	0	20q12	+	ENSG00000124183	ENSG00000124183	YTHDC2
4784	-1,337	0	1	13	0	6	8	5	0	11q22.3	+	ENSG00000170962	ENSG00000170962	TOP1
10422	1,308	7	4	3	3	9	2	1	4	14q13.3	+	ENSG00000165414	NA	PDGFD
5666	1,255	0	10	4	0	6	8	-4	0	20p13	+	ENSG00000088826	ENSG00000088826	Q96HD6
12517	-1,179	0	4	10	0	8	6	4	0	17p13.1	+	ENSG00000174292	ENSG00000174292	SMOX
11084	1,089	0	10	4	0	6	8	-4	0	22q12.2	+	NA	NA	TNK1
13376	1,036	0	6	8	0	0	14	-6	0	6p21.1	-	ENSG00000048544	ENSG00000048544	No gene
5202	1,023	1	10	3	1	5	8	-5	0	10p11.23	-	ENSG00000107968	ENSG00000107968	MRPS10
														MAP3K8

Fortsetzung Tabelle 17

Tabelle 18: Kandidatenliste für das Ansprechen auf eine Chemotherapie von in gesundem Gewebe hypermethylierten Fragmenten

Fragment ID	T-Statistik	Responder pro Methylierungsgruppe			Non-Responder pro Methylierungsgruppe			Differenz methyliert	Res - Non un-methyliert	Chr	CpG Insel	TSS	Gen	Gen Name
		methyliert	semi-methyliert	un-methyliert	methyliert	semi-methyliert	un-methyliert							
3757	3,92	8	6	0	2	6	6	-6	6	Xq13.2	-	ENSG00000186462	ENSG00000186462	NAP1L2
8156	3,16	6	5	3	3	1	10	-7	3	2p13.3	-	NA	NA	-
													ENSG00000176293,E NSG00000166704,EN	
2382	3,07	5	5	4	0	7	7	-3	5	19q13.43	-	ENSG00000176593	SG00000176593	Q8N9G5; ZNF135; ZNF606
11069	3,00	10	4	0	5	3	6	-6	5	8p23.1	-	NA	NA	-
640	3,01	6	5	3	1	7	6	-3	5	3q29	+	NA	NA	-
7295	2,92	9	5	0	1	12	1	-1	8	12q13.12	+	ENSG00000169884	ENSG00000169884	Wnt10B
1126	2,83	10	2	2	5	5	4	-2	5	9q33.2	-	ENSG00000136861	ENSG00000136861	CDK5RAP2
4575	2,83	4	4	6	0	3	11	-5	4	19p13.3	-	NA	NA	-
3712	2,85	0	11	3	0	4	10	-7	0	6q21	-	NA	NA	-
9427	2,80	2	12	0	0	10	4	-4	2	19q13.42	+	ENSG00000167615	ENSG00000167615	LENG8
798	2,77	5	8	1	3	3	8	-7	2	1q42.3	-	NA	ENSG00000168243	GNG4
7126	2,77	10	4	0	4	7	3	-3	6	1q24.2	-	NA	ENSG00000143178	TBX19
													ENSG00000100380,E	
5968	2,74	11	2	1	4	5	5	-4	7	22q13.2	-	NSG00000100382	ENSG00000100380	ST13
10521	2,73	4	9	1	1	7	6	-5	3	3q26.33	-	ENSG00000058056	ENSG00000058056	USP13
9032	2,69	4	8	2	3	2	9	-7	1	Xp11.4	+	ENSG00000156313	ENSG00000156313	RPGR
													ENSG00000149357,E	
6921	2,66	8	4	2	5	2	7	-5	3	11q13.4	-	NSG00000193294	ENSG00000149357	C11orf95
5039	2,61	13	1	0	7	7	0	0	6	13q33.2	-	NA	NA	-
6686	2,60	10	2	2	7	1	6	-4	3	12p13.32	+	ENSG00000111224	ENSG00000111224	PARP11
10053	2,55	13	1	0	8	6	0	0	5	3q26.2	+	NA	NA	-
7656	2,53	14	0	0	9	3	2	-2	5	9q32	-	NA	ENSG00000165185	KIAA1958
10380	2,51	10	4	0	4	10	0	0	6	9q21.31	+	ENSG00000106829	NA	TLE4
11219	2,50	11	3	0	7	5	2	-2	4	1q23.3	-	NA	NA	-
11003	2,44	1	9	4	0	3	11	-7	1	1p31.2	+	NA	NA	-
361	2,35	7	6	1	2	9	3	-2	5	2p11.2	+	ENSG00000115525	ENSG00000115525	ST3GAL5
5153	2,34	1	8	5	1	3	10	-5	0	14q32.11	-	NA	NA	-
1666	2,29	9	5	0	4	10	0	0	5	3q26.33	+	ENSG00000114416	ENSG00000114416	FXR1
12156	2,29	7	7	0	5	2	7	-7	2	20p11.21	-	NA	ENSG00000101463	C20orf39
3833	2,26	0	12	2	0	6	8	-6	0	21q22.2	-	ENSG00000183412	NA	C21orf104
12343	2,23	12	1	1	4	9	1	0	8	3p22.2	-	NA	ENSG00000185313	SCN10A
8683	2,22	9	4	1	3	7	4	-3	6	17q21.2	-	NA	ENSG00000131759	RARA

Fragment ID	T- Statistik	Responder pro Methylierungsgruppe			Non-Responder pro Methylierungsgruppe			Differenz methyliert	Res - Non methyliert	Chr	CpG Insel	TSS	Gen	Gen Name	
		methyliert	semi-methyliert	un-methyliert	methyliert	semi-methyliert	un-methyliert								
8639	2,22	7	7	0	3	9	2	-2	4	18q12.2	+	ENSG00000141428	ENSG00000141428	C18orf21	
374	2,23	11	3	0	7	7	0	0	4	16q13	-	NA	ENSG00000140848	CPNE2	
10535	2,20	10	4	0	4	6	4	-4	6	4q31.21	-	NA	ENSG00000170153	RNF150	
7624	2,20	12	2	0	8	6	0	0	4	1p36.12	-	NA	ENSG00000117242	DDOST	
10780	2,19	13	1	0	6	8	0	0	7	8q22.3	+	NA	NA	-	
3716	2,13	14	0	0	10	4	0	0	4	21q22.3	-	NA	NA	-	
2972	2,11	10	4	0	5	9	0	0	5	6p21.1	-	NA	ENSG00000137218	FRS3	
7100	2,05	10	3	1	4	6	4	-3	6	21q22.2	-	NA	ENSG00000157578	C21orf13	
													ENSG00000179738,E		
4328	2,01	13	1	0	7	7	0	0	6	1p36.22	+	ENSG00000179840	NSG00000179840	Q8N853; novel gene	
4300	1,90	12	2	0	5	9	0	0	7	10p15.1	-	NA	ENSG00000065675	PRKCQ	
12028	1,86	12	2	0	7	7	0	0	5	7q36.1	-	NA	ENSG00000033050	ABCF2	
8309	1,83	11	3	0	7	7	0	0	4	5q14.1	+	ENSG00000145685	ENSG00000145685	LHFPL2	
11653	-1,86	6	8	0	13	1	0	0	-7	4q31.22	-	NA	ENSG00000151617	EDNRA	
9519	-1,78	9	5	0	14	0	0	0	-5	20p13	-	ENSG00000171878	ENSG00000171878	novel gene	
													ENSG00000188121,E		
11781	1,75	12	2	0	8	6	0	0	4	19p13.3	+	NA	NSG00000116014	KISS1R	
5475	1,69	9	5	0	5	9	0	0	4	12p13.31	-	NA	ENSG00000166535	A2ML1	
6925	1,68	0	6	8	0	2	12	-4	0	3p21.1	-	NA	ENSG00000157445	CACNA2D3	
													ENSG00000138804,E		
1805	1,64	1	11	2	0	8	6	-4	1	4q25	-	NSG00000192257	NA	Q9H5S2	
9294	1,63	7	7	0	3	10	1	-1	4	1p36.22	-	NA	ENSG00000171621	SPSB1	
4075	1,63	4	7	3	0	9	5	-2	4	17p13.1	-	NA	ENSG00000133026	MYH10	
4606	1,63	10	4	0	5	9	0	0	5	8q24.3	+	NA	NA	-	
10491	1,62	0	11	3	0	7	7	-4	0	17p12	+	NA	NA	-	
6969	1,62	12	2	0	8	6	0	0	4	7q34	+	ENSG00000146963	NA	LUC7L2	
3764	-1,60	5	9	0	9	5	0	0	-4	7q36.2	+	ENSG00000157219	ENSG00000157219	HTR5A	
11102	-1,58	8	6	0	13	1	0	0	-5	10q25.2	-	NA	ENSG00000148737	TCF7L2	
12844	1,57	2	12	0	2	7	5	-5	0	11q23.2	-	ENSG00000086827	ENSG00000086827	ZW10	
10377	-1,56	9	5	0	13	1	0	0	-4	10q26.3	-	NA	ENSG00000175470	PPP2R2D	
11290	-1,55	3	9	2	7	7	0	2	-4	18q12.3	-	NA	ENSG00000132874	SLC14A2	
6652	-1,55	10	4	0	14	0	0	0	-4	9q21.2	-	NA	ENSG00000156049	GNA14	
7666	1,54	9	4	1	5	7	2	-1	4	12q13.3	+	NA	ENSG00000135454	B4GALNT1	
4208	1,52	11	3	0	5	9	0	0	6	11q24.2	+	ENSG00000110063	ENSG00000110063	DCPS	
7048	-1,50	8	6	0	12	2	0	0	-4	Xq24	-	NA	NA	-	
12018	1,47	10	4	0	6	8	0	0	4	22q13.1	-	NA	ENSG00000100324	MAP3K7IP1	
7548	-1,47	10	4	0	14	0	0	0	-4	6p22.1	-	NA	NA	-	

Fragment ID	T-Statistik	Responder pro Methylierungsgruppe			Non-Responder pro Methylierungsgruppe			Differenz methyliert	Res - Non methyliert	Chr	CpG Insel	TSS	Gen	Gen Name
		methyliert	semi-methyliert	un-methyliert	methyliert	semi-methyliert	un-methyliert							
6861	-1,46	6	8	0	10	4	0	0	-4	11p15.4	-	NA	NA	-
5850	-1,43	4	10	0	9	5	0	0	-5	19q13.42	-	NA	NA	-
10598	1,41	14	0	0	10	4	0	0	4	2p13.3	-	NA	NA	-
4005	1,41	14	0	0	8	6	0	0	6	20q11.21	-	NA	ENSG00000101337	TM9SF4
5140	1,39	6	8	0	2	12	0	0	4	14q12	+	ENSG00000129480	ENSG00000129480	C14orf126
10276	1,39	0	14	0	1	6	7	-7	-1	6p21.1	+	ENSG00000164663	ENSG00000164663	USP49
3327	1,32	13	1	0	9	5	0	0	4	15q22.31	-	NA	ENSG00000035664	DAPK2
4026	1,30	12	2	0	8	6	0	0	4	6q14.1	-	ENSG00000146242	ENSG00000146242	TPBG
8135	-1,27	7	7	0	12	2	0	0	-5	12q13.11	-	ENSG00000134294	NA	SLC38A2
10202	1,26	14	0	0	9	5	0	0	5	19q13.32	-	NA	NA	-
10066	-1,26	3	11	0	8	6	0	0	-5	1p36.22	-	NA	ENSG00000171603	CLSTN1
12867	-1,18	3	2	9	2	8	4	5	1	12q13.13	-	NA	NA	-
6342	1,17	8	6	0	4	10	0	0	4	8p22	-	NA	ENSG00000104763	ASAHI
9016	1,17	11	3	0	6	8	0	0	5	17p13.1	-	NA	ENSG00000060010	O94853
8493	1,16	0	10	4	0	6	8	-4	0	Xp11.23	+	ENSG00000147155	ENSG00000147155	EBP
3974	1,16	8	5	1	4	7	3	-2	4	1q31.3	-	ENSG00000162687	NA	KCNT2
1667	-1,14	1	5	8	1	9	4	4	0	3q22.1	-	NA	NA	-
4792	1,13	11	3	0	6	7	1	-1	5	5q31.2	-	ENSG00000112981	ENSG00000112981	NME5
3427	-1,05	8	5	1	12	1	1	0	-4	17q21.2	-	NA	ENSG00000131737	KRT34
762	-1,05	3	10	1	8	5	1	0	-5	10p15.3	-	NA	ENSG00000185736	ADARB2
4943	-1,04	9	5	0	13	1	0	0	-4	2p21	-	NA	ENSG00000143921	ABCG8

Fortsetzung Tabelle 15

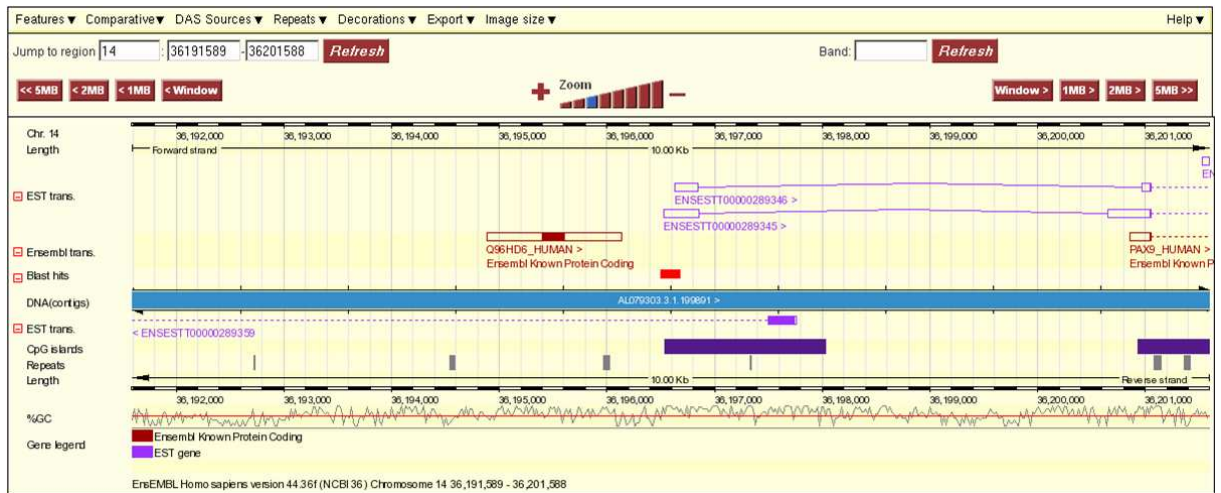


Abbildung 27: Fragment 10422

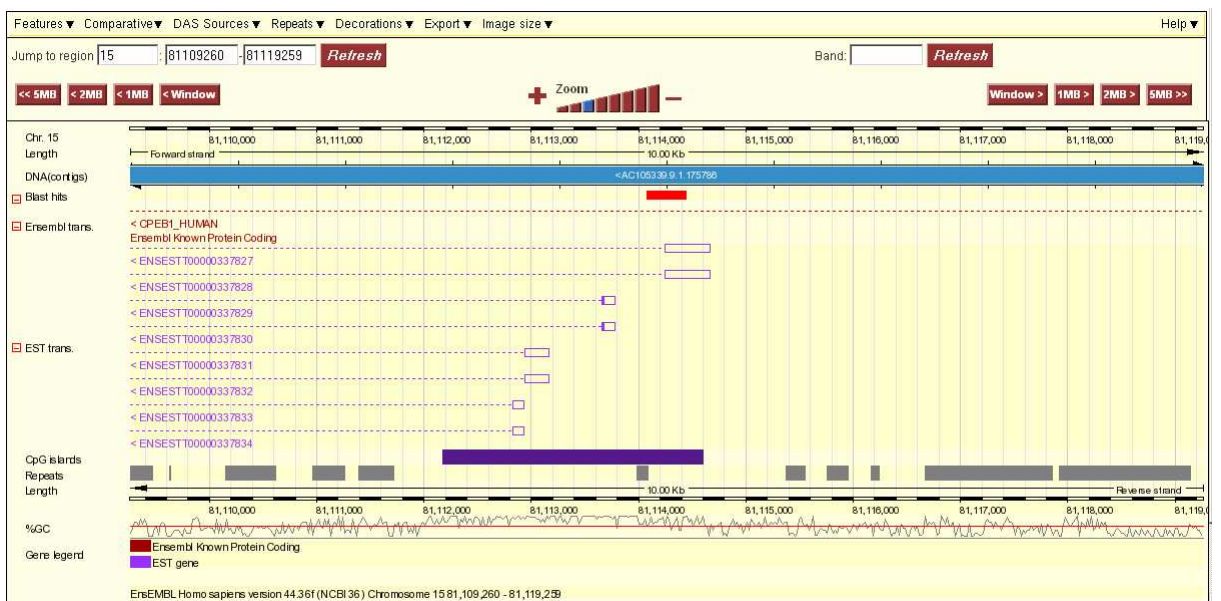


Abbildung 28: Fragment 12787

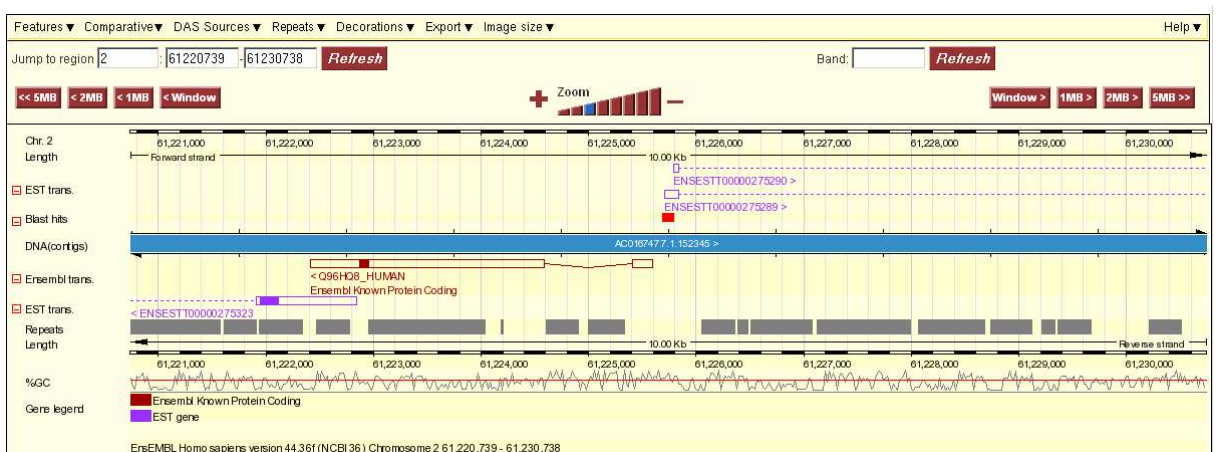


Abbildung 29: Fragment 13165

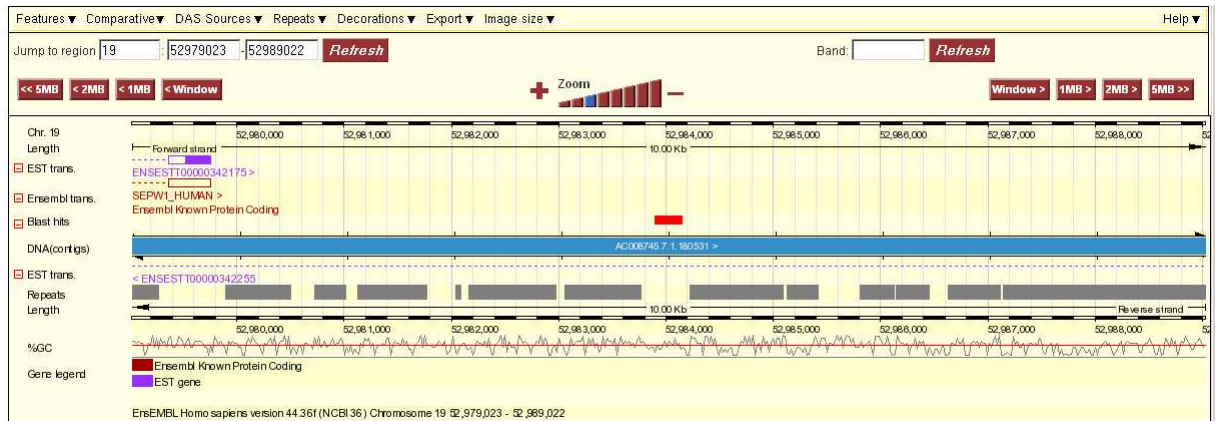


Abbildung 30: Fragment 1322

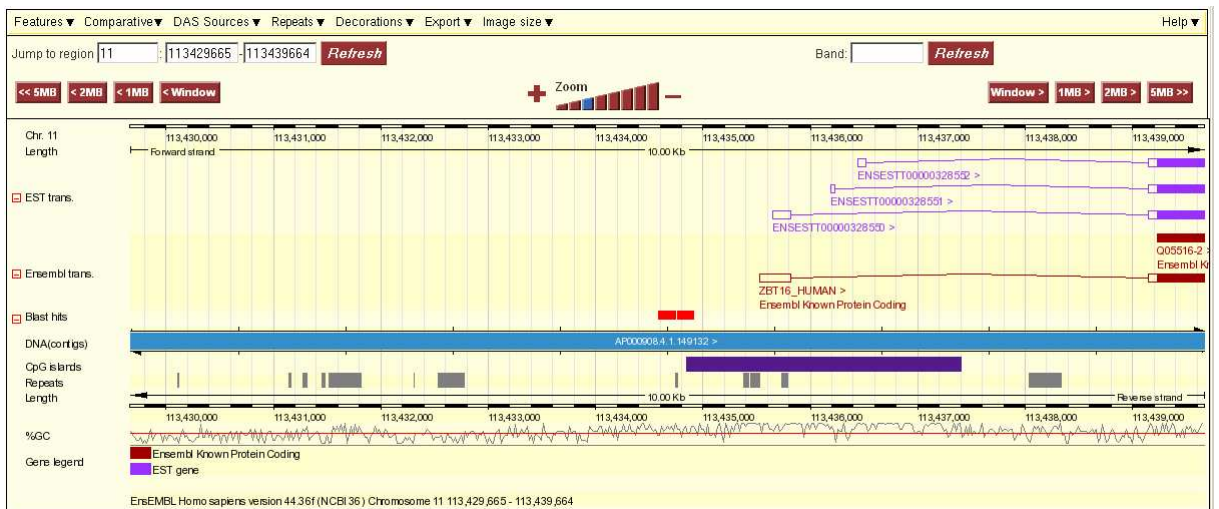


Abbildung 31: Fragment 1589

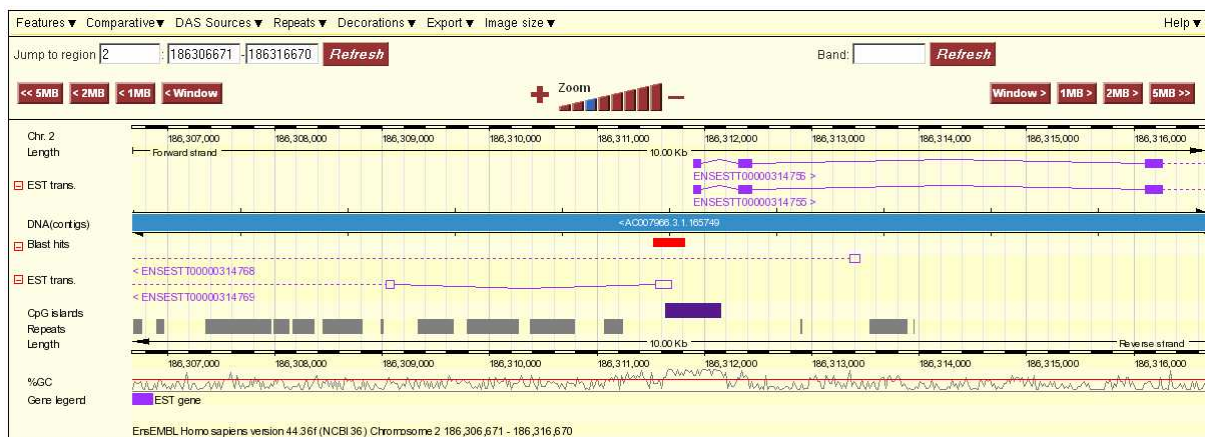


Abbildung 32: Fragment 4235

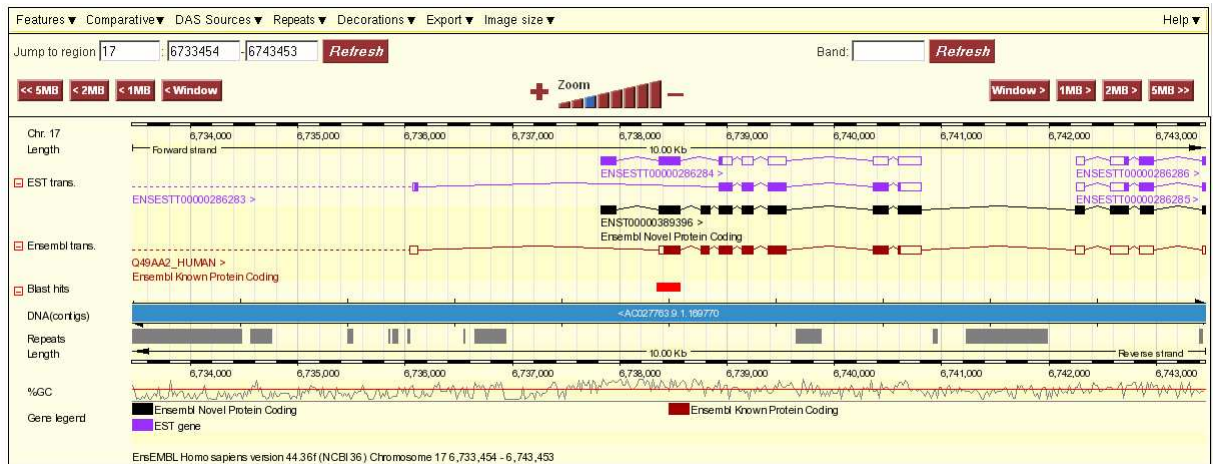


Abbildung 33: Fragment 4493

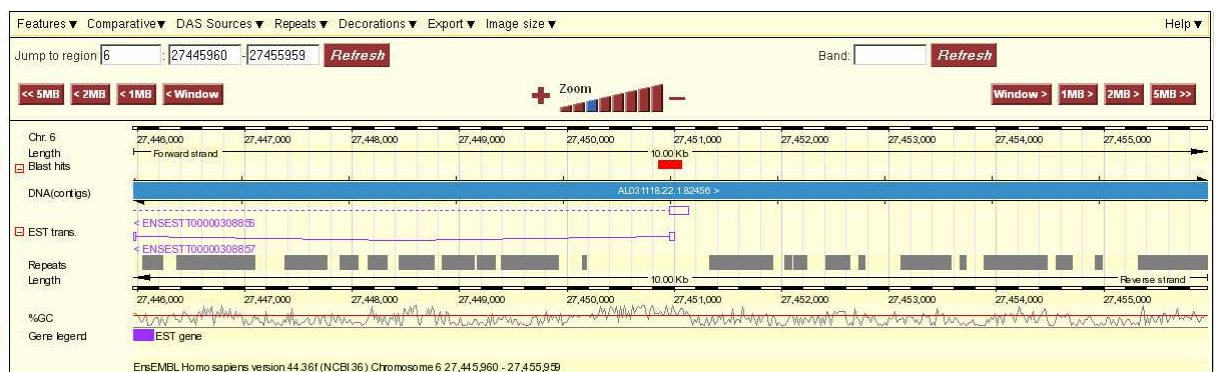


Abbildung 34: Fragment 4563

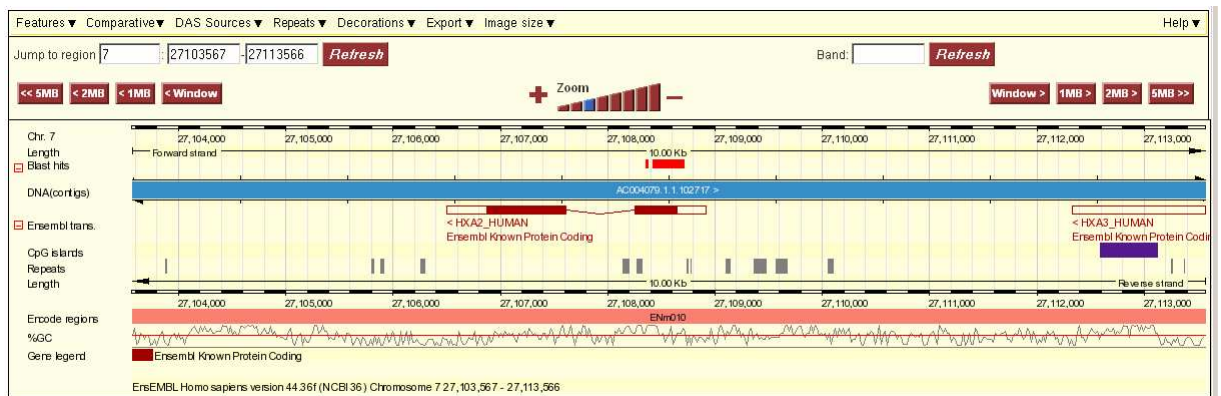


Abbildung 35: Fragment 4664

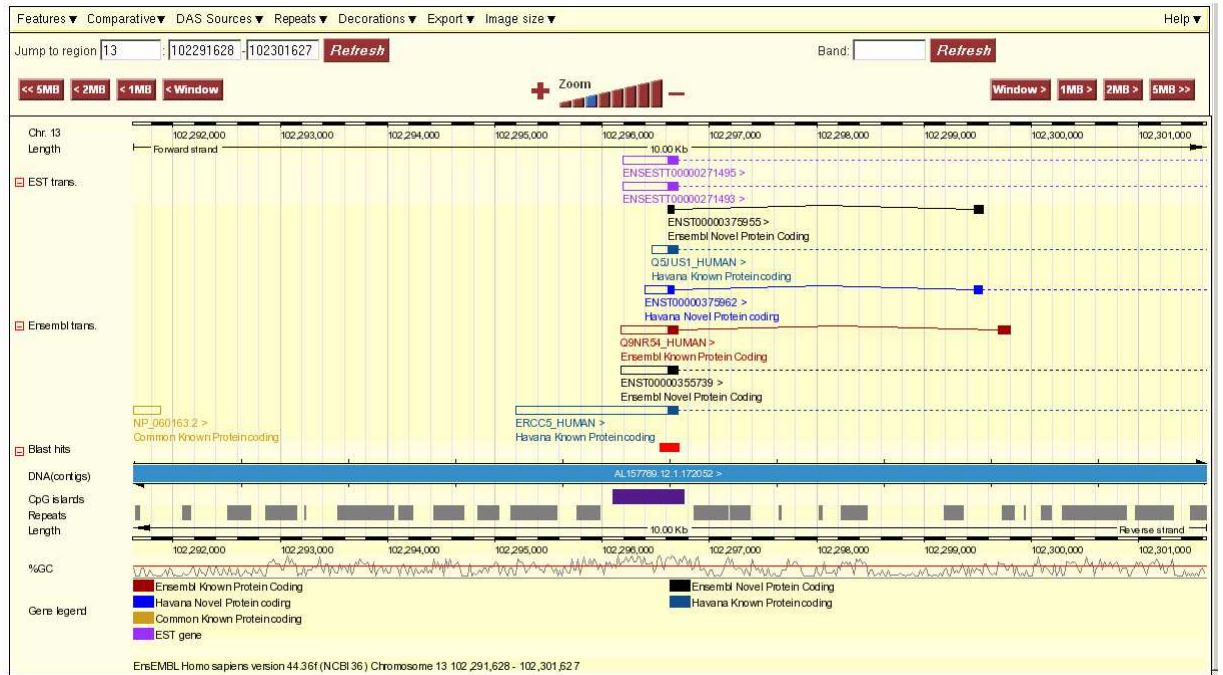


Abbildung 36: Fragment 7951

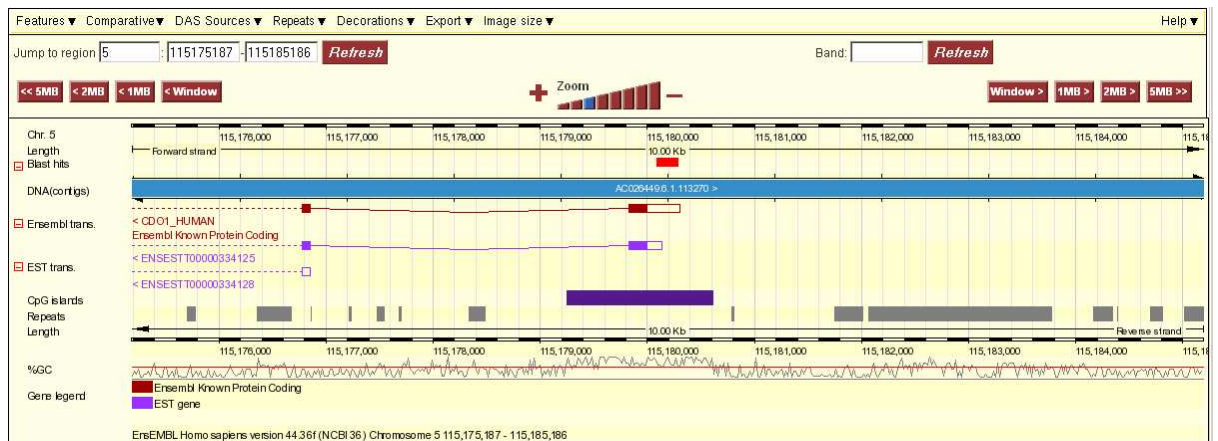


Abbildung 37: Fragment 9857

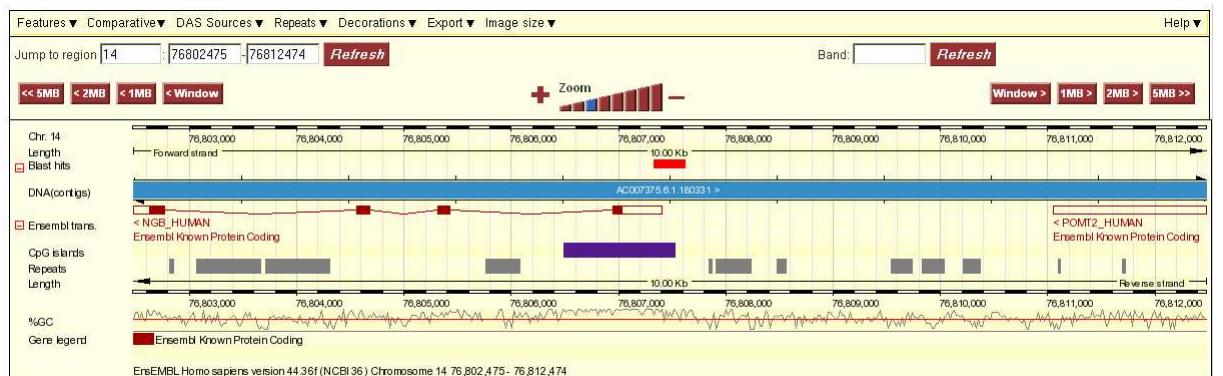


Abbildung 38: Fragment 3615

5.11. Abkürzungsverzeichnis und Gennamen

Abkürzungen

°C	Grad Celsius
°K	Grad Kelvin
μ	Mikro
5' UTR	5' untranslatierte Region
5-meC	5-Methylcytosin
A	2'-Desoxyadenin
AG	Aktiengesellschaft
ATP	Adenosintriphosphat
B	Belgien
BAC	engl. Bacterial Artificial Chromosome
Bisulfit-DNA	mit Natriumbisulfit behandelte und desulfonierte DNA
bp	Basenpaar
BSA	Rinderserumalbumin (engl. Bovine Serum Albumin)
C	2'-Desoxycytosin
CDN	Kanada
CGH	Vergleichende Genomhybridisierung (engl. Comparative Genome Hybridization)
cm	Zentimeter
CMF	Chemotherapieschema aus Cyclophosphamid, Methotrexat und 5-Fluorouracil
COBRA	Kombinationsassay aus Bisulfit und Restriktion (engl. Combined Bisulfite Restriction Assay)
CpG	5'-CG-Dinukleotid
CR	engl. Complete Remission (vollständige Remission einer Metastase)
D	Deutschland
Da	Dalton
dATP	2'-Desoxyadenosin-5'-triphosphat
DCIS	Duktales Karzinom in situ
dCTP	2'-Desoxycytidin-5'-triphosphat
ddNTPs	Didesoxyribonukleosid-Triphosphate
dGTP	2'-Desoxyguanosin-5'-triphosphat
DMH	Differentielle Methylierungs Hybridisierung (engl. Differential Methylation Hybridization)
DNA	Desoxyribonukleinsäure (engl. Desoxyribonucleic Acid)
dNTP	2'-Desoxyribonukleosid-5'-triphosphat
dsDNA	doppelsträngige DNA
dTTP	2'-Desoxythymidin-5'-triphosphat
dUTP	2'-Desoxyuridin-5'-triphosphat
EC/AC	Epirubicin oder Adriamycin und Cyclophosphamid
EDTA	Ethylendiamintetraessigsäure
ER	Östrogenrezeptor (engl. Estrogen Receptor)
ESME	Epigenetische Sequenzierungs-Methylierungsanalyse
EU	Europäische Union
FDA	US-amerikanische Behörde für Lebensmittel- und Arzneimittelsicherheit (engl. Food and Drug Administration)
FEC/FAC	5-Fluorouracil, Epirubicin oder Adriamycin und Cyclophosphamid
G	2'-Desoxyguanosin
g	Gramm
h	Stunde
HMEC	Humane Brustepithelzellen (engl. Human Mammary Epithelial Cells)
HPLC	Hochleistungs-Flüssigchromatographie (engl. High Performance Liquid Chromatography)
ICEAMP	Englisch: Identification of CpG Islands Exhibiting Altered Methylation Patterns
IDC	Invasives duktales Karzinom

IHC	Immunohistochemie
ILC	Invasives lobuläres Karzinom
k	Kilo
Laser	Englisch: Light Amplification by Stimulated Emission of Radiation
LCIS	Lobuläres Karzinom in situ
log	Logarithmus
LOH	Verlust der Heterozygotie (engl. Loss of Heterozygosity)
LOI	Verlust des Imprinting (engl. Loss of Imprinting)
M	Methylierungsindex
M	Molar
m	Meter
m	Mili
MALDI	Englisch: Matrix-Assisted Laser Desorption/Ionization
MCA	Englisch: methylated CGI amplification
MIAMI	Englisch: Integrated Analysis of Methylation by Isoschizomers Methylierte DNA Immunopräzipitation
min	Minuten
mol	Mol
MS	Methylierungs-sensitiv
MS-AP-PCR	Methylierungs-sensitiv Arbiträre Primer PCR (engl. Methylation-Sensitive Arbitrarily Primed PCR)
MSDK	Methylierungs-sensitiv digitale Karyotypisierung (engl. Methylation-Sensitive Digital Karyotyping)
MSP	Methylierungsspezifische PCR (engl. Methylation-Sensitive PCR)
MS-RLGS	Landmark Genomisches Scannen (engl. Methylation-Sensitive Restriction Landmark Genomic Scanning)
MS-SNuPE	Methylierungsspezifische Einzelnukleotid-Primerverlängerung (engl. Methylation Specific Single Nucleotide Primer Extension)
NCBI	Englisch: National Center for Biotechnology Information
NL	Niederlande
nm	Nanometer
nmol	Nanomol
OD	Optische Dichte
p	Pico-
PBL	Periphere Blut-Lymphozyten
PCR	Polymerasekettenreaktion (engl. Polymerase Chain Reaction)
PD	Progressive Erkrankung (engl. Progressive Disease)
PgR	Progesteronrezeptor
Phi100%-DNA	DNA, hergestellt durch genomweite Amplifikation mit Hilfe der phi29-DNA-Polymerase
PR	Teilweise Remission einer Metastase (engl. Partial Remission)
RNA	Ribonukleinsäure
rpm	Umdrehungen pro Minute
RT PCR	Reverse Transkriptase PCR
s	Sekunde
SAM	S-Adenosyl-L-Methionin
SD	Stabile Erkrankung (engl. Stable Disease)
SDS	Sodiumdodecylsulfat
SNP	Single nucleotide polymorphism
ssDNA	Einzelsträngige DNA (Single Stranded DNA)
T	2'-Desoxythymidin
TBE-Puffer	Tris/Borsäure/EDTA-Puffer
TdT	Terminale Desoxyribonukleotidyltransferase
TE-Puffer	Tris/EDTA-Puffer
TNM	Zeitpunkt T; nodaler Status N; Metastasen M
Tris	Trishydroxymethylaminomethan
TSS	Transkriptionsstart
U	Uracil

UICC	Union International Contre le Cancer
UK	Großbritannien
USA	Vereinigten Staaten von Amerika (engl. United States of America)
UV	ultraviolett
V	Volt
WHO	Weltgesundheitsorganisation (engl. World Health Organization)
Xist	X inactivation center
ABC-Transporter	ATP bindenden Kasette (engl. ATP binding cassette)

Gennamen

<i>ABTB2</i>	ankyrin repeat and BTB (POZ) domain containing 2
<i>ADAM17</i>	ADAM metalloproteinase domain 17
<i>ADAM18</i>	ADAM metalloproteinase domain 18
<i>ADCY6</i>	adenylate cyclase 6
<i>ALKBH3</i>	alkB, alkylation repair homolog 3 (E. coli)
<i>ANKS1B</i>	ankyrin repeat and sterile alpha motif domain containing 1B
<i>APC</i>	adenomatosis polyposis coli
<i>APP</i>	amyloid beta (A4) precursor protein
<i>ARHGAP8</i>	Rho GTPase activating protein 8
<i>ARPC4</i>	actin related protein 2/3 complex, subunit 4, 20kDa
<i>ATM</i>	ataxia telangiectasia mutated
<i>BCAT1</i>	branched chain aminotransferase 1, cytosolic
<i>BRCA1</i>	breast cancer 1, early onset
<i>BRCA2</i>	breast cancer 2, early onset
<i>CACNA2D3</i>	calcium channel, voltage-dependent, alpha 2/delta subunit 3
<i>CCDC76</i>	coiled-coil domain containing 76
<i>CDC2L5</i>	cell division cycle 2-like protein kinase 5
<i>CDCA5</i>	cell division cycle associated
<i>CDK5RAP2</i>	CDK5 regulatory subunit associated protein 2
<i>CDO1</i>	cystein dioxygenase, type I
<i>cMyc</i>	v-myc myelocytomatosis viral oncogene homolog (avian)
<i>COPS4</i>	COP9 constitutive photomorphogenic homolog subunit 4
<i>COX-2</i>	cyclooxygenase 2b
<i>CSTP1</i>	cystatin pseudogene 1
<i>CTNNA3</i>	catenin (cadherin-associated protein), alpha 3
<i>CCND2</i>	<i>cyclin D2</i>
<i>DTX1</i>	deltex homolog 1 (Drosophila)
<i>DTX4</i>	deltex 4 homolog (Drosophila)
<i>DVL1</i>	dishevelled, dsh homolog 1 (Drosophila)
<i>EBP</i>	emopamil binding protein (sterol isomerase)
<i>CDH1</i>	<i>E-Cadherin</i>
<i>ENDOGL1</i>	endonuclease G like 1
<i>EPC1</i>	enhancer of polycomb homolog 1 (Drosophila)
<i>ERCC5</i>	excision repair cross-complementing rodent repair deficiency, complementation group 5
<i>ESR1</i>	estrogen receptor 1
<i>ERBB2</i>	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
<i>FGFRL1</i>	fibroblast growth factor receptor-like 1 precursor
<i>FXR1</i>	fragile X mental retardation, autosomal homolog 1
<i>GEMIN7</i>	Gem-associated protein 7
<i>GFER</i>	growth factor, augmenter of liver regeneration
<i>GNG4</i>	guanine nucleotide binding protein (G protein), gamma 4
<i>GnRH</i>	gonadotropin releasing hormon
<i>GSTp1</i>	glutathione S-transferase p1
<i>CDH13</i>	<i>H-Cadherin</i>
<i>HER-2/neu</i>	siehe ERBB2

<i>HMGN3</i>	high-mobility group nucleosome binding domain 3
<i>MLH1</i>	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)
<i>HOXA2</i>	homeobox A2
<i>IGFBP-3</i>	insulin-like growth factor binding protein 3
<i>int-2</i>	INT-2 proto-oncogene protein
<i>KPNA2</i>	Importin alpha-2 subunit (Karyopherin alpha-2 subunit)
<i>LAMC2</i>	laminin, gamma 2
<i>LAMA3</i>	laminin-5
<i>LENG8</i>	leukocyte receptor cluster (LRC) member 8
<i>LRRC45</i>	leucine rich repeat containing 45
<i>MDR1</i>	multidrug resistance 1
<i>MGMT</i>	O-6-methylguanin-DNA-methyltransferase
<i>MLL</i>	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)
<i>NAP1L2</i>	nucleosome assembly protein 1-like 2
<i>NDUFA9</i>	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex
<i>NDUFV3</i>	NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa
<i>NGB</i>	neuroglobin
<i>NTRK1</i>	neurotrophic tyrosine kinase, receptor, type 1
<i>OR5A2</i>	olfactory receptor, family 5, subfamily A, member 2
<i>CDKN2B</i>	cyclin-dependent kinase inhibitor 2B
<i>CDKN2A</i>	cyclin-dependent kinase inhibitor 2A
<i>p15</i>	tumor suppressor gene p15 (CDKN2B)
<i>p16</i>	tumor suppressor gene p16 (CDKN2A)
<i>CDKN1A</i>	cyclin-dependent kinase inhibitor 1A
<i>p21</i>	tumor suppressor gene p16 (CDKN1A)
<i>p53</i>	tumor suppressor gene p53 (Tp53)
<i>PARP11</i>	poly (ADP-ribose) polymerase family, member 11
<i>PCID1</i>	PCI domain containing 1
<i>PCSK9</i>	proprotein convertase subtilisin/kexin type 9
<i>PIGM</i>	PIG-M mannosyltransferase
<i>PITX2</i>	paired-like homeodomain 2
<i>PRDM12</i>	PR domain containing 12
<i>PSAT1</i>	phosphoserine aminotransferase 1
<i>RARA</i>	retinoic acid receptor, alpha
<i>RARB</i>	retinoic acid receptor, beta
<i>RASSF1A</i>	ras association domain family 1
<i>RHBG</i>	Rh family, B glycoprotein
<i>ROPN1L</i>	ropporin 1-like
<i>RPGR</i>	retinitis pigmentosa GTPase regulator
<i>SALL3</i>	sal-like 3 (Drosophila)
<i>SASS6</i>	spindle assembly 6 homolog (C. elegans)
<i>SCN10A</i>	sodium channel, voltage-gated, type X, alpha subunit
<i>SEMA3E</i>	semaphorin 3E precursor
<i>SFRP1</i>	secreted frizzled-related protein 1
<i>SLC22A11</i>	solute carrier family 22 (organic anion/urate transporter), member 11
<i>SNX8</i>	sorting nexin 8
<i>ST13</i>	suppression of tumorigenicity 13
<i>ST3GAL5</i>	ST3 beta-galactoside alpha-2,3-sialyltransferase 5
<i>STARD4</i>	StAR-related lipid transfer protein 4
<i>STRA13</i>	stimulated by retinoic acid 13
<i>TBX19</i>	T-box 19
<i>TLE4</i>	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)
<i>TTBK2</i>	tau tubulin kinase 2
<i>TWIST</i>	twist homolog 1 (Drosophila)
<i>PLAU</i>	plasminogen activator, urokinase
<i>uPA</i>	U-plasminogen activator (PLAU)
<i>USP13</i>	ubiquitin specific peptidase 13 (isopeptidase T-3)

<i>USP3</i>	ubiquitin carboxyl-terminal hydrolase 3
<i>USP49</i>	ubiquitin specific peptidase 49
<i>UTX</i>	ubiquitously transcribed X chromosome tetratricopeptide repeat protein
<i>VEGF</i>	vascular Endothelial Growth Factor A
<i>VENTX2</i>	VENT homeobox homolog (<i>Xenopus laevis</i>)
<i>Wnt10B</i>	wingless-type MMTV integration site family, member 10B
<i>XPR1</i>	xenotropic and polytropic retrovirus receptor
<i>ZADH1</i>	zinc binding alcohol dehydrogenase, domain containing 1
<i>ZBTB2</i>	zinc finger and BTB domain containing protein 2
<i>ZFPL1</i>	zinc finger protein-like 1
<i>ZNF 606</i>	zinc finger protein 606
<i>ZNF135</i>	zinc finger protein 135
<i>ZNF145</i>	zinc finger protein 145 (Kruppel-like, expressed in promyelocytic leukemia)
<i>ZNF184</i>	zinc finger protein 184
<i>ZNF266</i>	zinc finger protein 266 (Zinc finger protein HZF1)
<i>ZNF582</i>	zinc finger protein 582
<i>ZNFN1A1</i>	zinc finger protein, subfamily 1A, 1 (Ikaros)

5.12. Literaturverzeichnis

1. Klaus Giersiepen, C.H., Katrin Janhsen, and C. Lange, *Gesundheitsberichterstattung des Bundes; Heft 25; Brustkrebs*. 2005, Robert Koch-Institut. p. 33.
2. Sant, M., et al., *Breast carcinoma survival in Europe and the United States*. *Cancer*, 2004. **100**(4): p. 715-22.
3. Mallon, E., et al., *The basic pathology of human breast cancer*. *J Mammary Gland Biol Neoplasia*, 2000. **5**(2): p. 139-63.
4. Sobin, L.H., *TNM: principles, history, and relation to other prognostic factors*. *Cancer*, 2001. **91**(8 Suppl): p. 1589-92.
5. Elston, C.W. and I.O. Ellis, *Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up*. *Histopathology*, 1991. **19**(5): p. 403-10.
6. Osborne, C.K., *Steroid hormone receptors in breast cancer management*. *Breast Cancer Res Treat*, 1998. **51**(3): p. 227-38.
7. Ross, J.S., et al., *Breast cancer biomarkers*. *Adv Clin Chem*, 2005. **40**: p. 99-125.
8. Harvey, J.M., et al., *Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer*. *J Clin Oncol*, 1999. **17**(5): p. 1474-81.
9. Bertucci, F., et al., *Gene expression profiling of primary breast carcinomas using arrays of candidate genes*. *Hum Mol Genet*, 2000. **9**(20): p. 2981-91.
10. Locker, G.Y., *Hormonal therapy of breast cancer*. *Cancer Treat Rev*, 1998. **24**(3): p. 221-40.
11. McPherson, K., C.M. Steel, and J.M. Dixon, *ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics*. *Bmj*, 2000. **321**(7261): p. 624-8.
12. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1*. *Science*, 1994. **266**(5182): p. 66-71.
13. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2*. *Nature*, 1995. **378**(6559): p. 789-92.
14. Gospodarowicz, M., et al., *Prognostic factors in clinical decision making: the future*. *Cancer*, 2001. **91**(8 Suppl): p. 1688-95.
15. Partridge, A.H., H.J. Burstein, and E.P. Winer, *Side effects of chemotherapy and combined chemohormonal therapy in women with early-stage breast cancer*. *J Natl Cancer Inst Monogr*, 2001(30): p. 135-42.
16. Andre, F. and L. Pusztai, *Molecular classification of breast cancer: implications for selection of adjuvant chemotherapy*. *Nat Clin Pract Oncol*, 2006. **3**(11): p. 621-32.
17. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. *Cell*, 2000. **100**(1): p. 57-70.
18. Murphy, N., E. Millar, and C.S. Lee, *Gene expression profiling in breast cancer: towards individualising patient management*. *Pathology*, 2005. **37**(4): p. 271-7.

19. Ross, J.S., et al., *The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy*. *Oncologist*, 2003. **8**(4): p. 307-25.
20. Gilbey, A.M., et al., *The detection of circulating breast cancer cells in blood*. *J Clin Pathol*, 2004. **57**(9): p. 903-11.
21. Cristofanilli, M., et al., *Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer*. *J Clin Oncol*, 2005. **23**(7): p. 1420-30.
22. Weigelt, B., et al., *Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer*. *Cancer Res*, 2005. **65**(20): p. 9155-8.
23. Weigelt, B., et al., *Gene expression profiles of primary breast tumors maintained in distant metastases*. *Proc Natl Acad Sci U S A*, 2003. **100**(26): p. 15901-5.
24. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer*. *Nature*, 2002. **415**(6871): p. 530-6.
25. Widschwendter, M. and P.A. Jones, *DNA methylation and breast carcinogenesis*. *Oncogene*, 2002. **21**(35): p. 5462-82.
26. Bird, A., *DNA methylation patterns and epigenetic memory*. *Genes Dev*, 2002. **16**(1): p. 6-21.
27. Cooper, D.N. and M. Krawczak, *Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes*. *Hum Genet*, 1989. **83**(2): p. 181-8.
28. Bird, A.P., *CpG-rich islands and the function of DNA methylation*. *Nature*, 1986. **321**(6067): p. 209-13.
29. Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes*. *J Mol Biol*, 1987. **196**(2): p. 261-82.
30. Reik, W., W. Dean, and J. Walter, *Epigenetic reprogramming in mammalian development*. *Science*, 2001. **293**(5532): p. 1089-93.
31. Robertson, K.D. and A.P. Wolffe, *DNA methylation in health and disease*. *Nat Rev Genet*, 2000. **1**(1): p. 11-9.
32. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals*. *Nat Genet*, 2003. **33 Suppl**: p. 245-54.
33. Laird, P.W., *Cancer epigenetics*. *Hum Mol Genet*, 2005. **14 Spec No 1**: p. R65-76.
34. Reik, W. and J. Walter, *Genomic imprinting: parental influence on the genome*. *Nat Rev Genet*, 2001. **2**(1): p. 21-32.
35. Laird, P.W., *The power and the promise of DNA methylation markers*. *Nat Rev Cancer*, 2003. **3**(4): p. 253-66.
36. Issa, J.P., et al., *Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon*. *Nat Genet*, 1994. **7**(4): p. 536-40.
37. Ahuja, N., et al., *Aging and DNA methylation in colorectal mucosa and cancer*. *Cancer Res*, 1998. **58**(23): p. 5489-94.
38. Hoal-van Helden, E.G. and P.D. van Helden, *Age-related methylation changes in DNA may reflect the proliferative potential of organs*. *Mutat Res*, 1989. **219**(5-6): p. 263-6.
39. Yang, X., L. Yan, and N.E. Davidson, *DNA methylation in breast cancer*. *Endocr Relat Cancer*, 2001. **8**(2): p. 115-27.
40. Soares, J., et al., *Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression*. *Cancer*, 1999. **85**(1): p. 112-8.

41. Pike, M.C., et al., *Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk*. *Epidemiol Rev*, 1993. **15**(1): p. 17-35.
42. Ehrlich, M., *DNA methylation in cancer: too much, but also too little*. *Oncogene*, 2002. **21**(35): p. 5400-13.
43. Jones, P.A. and P.W. Laird, *Cancer epigenetics comes of age*. *Nat Genet*, 1999. **21**(2): p. 163-7.
44. Laird, P.W., *Oncogenic mechanisms mediated by DNA methylation*. *Mol Med Today*, 1997. **3**(5): p. 223-9.
45. Alaminos, M., et al., *Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma*. *J Natl Cancer Inst*, 2004. **96**(16): p. 1208-19.
46. Gonzalgo, M.L., et al., *Molecular profiling and classification of sporadic renal cell carcinoma by quantitative methylation analysis*. *Clin Cancer Res*, 2004. **10**(21): p. 7276-83.
47. Marchevsky, A.M., J.A. Tsou, and I.A. Laird-Offringa, *Classification of individual lung cancer cell lines based on DNA methylation markers: use of linear discriminant analysis and artificial neural networks*. *J Mol Diagn*, 2004. **6**(1): p. 28-36.
48. Tsou, J.A., et al., *Distinct DNA methylation profiles in malignant mesothelioma, lung adenocarcinoma, and non-tumor lung*. *Lung Cancer*, 2005. **47**(2): p. 193-204.
49. Field, J.K., et al., *Methylation discriminators in NSCLC identified by a microarray based approach*. *Int J Oncol*, 2005. **27**(1): p. 105-11.
50. Wei, S.H., et al., *Methylation microarray analysis of late-stage ovarian carcinomas distinguishes progression-free survival in patients and identifies candidate epigenetic markers*. *Clin Cancer Res*, 2002. **8**(7): p. 2246-52.
51. Wei, S.H., et al., *Prognostic DNA methylation biomarkers in ovarian cancer*. *Clin Cancer Res*, 2006. **12**(9): p. 2788-94.
52. Graziano, F., et al., *Prognostic analysis of E-cadherin gene promoter hypermethylation in patients with surgically resected, node-positive, diffuse gastric cancer*. *Clin Cancer Res*, 2004. **10**(8): p. 2784-9.
53. House, M.G., et al., *Tumor suppressor gene hypermethylation as a predictor of gastric stromal tumor behavior*. *J Gastrointest Surg*, 2003. **7**(8): p. 1004-14; discussion 1014.
54. Chang, H.W., et al., *Loss of E-cadherin expression resulting from promoter hypermethylation in oral tongue carcinoma and its prognostic significance*. *Cancer*, 2002. **94**(2): p. 386-92.
55. Leung, W.K., et al., *Potential diagnostic and prognostic values of detecting promoter hypermethylation in the serum of patients with gastric cancer*. *Br J Cancer*, 2005. **92**(12): p. 2190-4.
56. Sathyanarayana, U.G., et al., *Aberrant promoter methylation of laminin-5-encoding genes in prostate cancers and its relationship to clinicopathological features*. *Clin Cancer Res*, 2003. **9**(17): p. 6395-400.
57. Sathyanarayana, U.G., et al., *Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes*. *Cancer Res*, 2004. **64**(4): p. 1425-30.
58. Gonzalez-Quevedo, R., et al., *Differential impact of p16 inactivation by promoter methylation in non-small cell lung and colorectal cancer: clinical implications*. *Int J Oncol*, 2004. **24**(2): p. 349-55.

59. Wang, J., et al., *Value of p16INK4a and RASSF1A promoter hypermethylation in prognosis of patients with resectable non-small cell lung cancer*. Clin Cancer Res, 2004. **10**(18 Pt 1): p. 6119-25.
60. Chang, Y.S., et al., *Correlation between insulin-like growth factor-binding protein-3 promoter methylation and prognosis of patients with stage I non-small cell lung cancer*. Clin Cancer Res, 2002. **8**(12): p. 3669-75.
61. Brabender, J., et al., *Adenomatous polyposis coli gene promoter hypermethylation in non-small cell lung cancer is associated with survival*. Oncogene, 2001. **20**(27): p. 3528-32.
62. Chim, C.S., et al., *Methylation of p15 and p16 genes in acute promyelocytic leukemia: potential diagnostic and prognostic significance*. J Clin Oncol, 2001. **19**(7): p. 2033-40.
63. Roman, J., et al., *Hypermethylation of the calcitonin gene in acute lymphoblastic leukaemia is associated with unfavourable clinical outcome*. Br J Haematol, 2001. **113**(2): p. 329-38.
64. Roman-Gomez, J., et al., *5' CpG island hypermethylation is associated with transcriptional silencing of the p21(CIP1/WAF1/SDI1) gene and confers poor prognosis in acute lymphoblastic leukemia*. Blood, 2002. **99**(7): p. 2291-6.
65. Rosenbaum, E., et al., *Promoter hypermethylation as an independent prognostic factor for relapse in patients with prostate cancer following radical prostatectomy*. Clin Cancer Res, 2005. **11**(23): p. 8321-5.
66. Janicke, F., et al., *Randomized adjuvant chemotherapy trial in high-risk, lymph node-negative breast cancer patients identified by urokinase-type plasminogen activator and plasminogen activator inhibitor type 1*. J Natl Cancer Inst, 2001. **93**(12): p. 913-20.
67. Pakneshan, P., B. Tetu, and S.A. Rabbani, *Demethylation of urokinase promoter as a prognostic marker in patients with breast carcinoma*. Clin Cancer Res, 2004. **10**(9): p. 3035-41.
68. Pakneshan, P., M. Szyf, and S.A. Rabbani, *Hypomethylation of urokinase (uPA) promoter in breast and prostate cancer: prognostic and therapeutic implications*. Curr Cancer Drug Targets, 2005. **5**(7): p. 471-88.
69. Chen, C.M., et al., *Methylation target array for rapid analysis of CpG island hypermethylation in multiple tissue genomes*. Am J Pathol, 2003. **163**(1): p. 37-45.
70. Shinozaki, M., et al., *Distinct hypermethylation profile of primary breast cancer is associated with sentinel lymph node metastasis*. Clin Cancer Res, 2005. **11**(6): p. 2156-62.
71. Chow, L.W., et al., *Aberrant methylation of cyclooxygenase-2 in breast cancer patients*. Biomed Pharmacother, 2005. **59 Suppl 2**: p. S264-7.
72. Veeck, J., et al., *Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis*. Oncogene, 2006. **25**(24): p. 3479-88.
73. Nimmrich, I., et al., *DNA hypermethylation of PITX2 is a marker of poor prognosis in untreated lymph node-negative hormone receptor-positive breast cancer patients*. Breast Cancer Res Treat, 2007.
74. Maier, S., et al., *Identifying DNA methylation biomarkers of cancer drug response*. Am J Pharmacogenomics, 2005. **5**(4): p. 223-32.
75. Kim, W.J., et al., *Aberrant methylation of the ATM promoter correlates with increased radiosensitivity in a human colorectal tumor cell line*. Oncogene, 2002. **21**(24): p. 3864-71.

76. Satoh, A., et al., *Epigenetic inactivation of CHFR and sensitivity to microtubule inhibitors in gastric cancer*. *Cancer Res*, 2003. **63**(24): p. 8606-13.
77. Fulda, S., et al., *Sensitization for death receptor- or drug-induced apoptosis by re-expression of caspase-8 through demethylation or gene transfer*. *Oncogene*, 2001. **20**(41): p. 5865-77.
78. Martens, J.W., et al., *Association of DNA methylation of phosphoserine aminotransferase with response to endocrine therapy in patients with recurrent breast cancer*. *Cancer Res*, 2005. **65**(10): p. 4101-17.
79. Widschwendter, M., et al., *Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen*. *Cancer Res*, 2004. **64**(11): p. 3807-13.
80. Fiegl, H., et al., *Circulating tumor-specific DNA: a marker for monitoring efficacy of adjuvant therapy in cancer patients*. *Cancer Res*, 2005. **65**(4): p. 1141-5.
81. Siegmund, K.D. and P.W. Laird, *Analysis of complex methylation data*. *Methods*, 2002. **27**(2): p. 170-8.
82. Schmitt, F., E.J. Oakeley, and J.P. Jost, *Antibiotics induce genome-wide hypermethylation in cultured *Nicotiana tabacum* plants*. *J Biol Chem*, 1997. **272**(3): p. 1534-40.
83. Kuo, K.C., et al., *Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA*. *Nucleic Acids Res*, 1980. **8**(20): p. 4763-76.
84. Fraga, M.F., et al., *High-performance capillary electrophoretic method for the quantification of 5-methyl 2'-deoxycytidine in genomic DNA: application to plant, animal and human cancer tissues*. *Electrophoresis*, 2002. **23**(11): p. 1677-81.
85. Oakeley, E.J., A. Podesta, and J.P. Jost, *Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation*. *Proc Natl Acad Sci U S A*, 1997. **94**(21): p. 11721-5.
86. Stach, D., et al., *Capillary electrophoretic analysis of genomic DNA methylation levels*. *Nucleic Acids Res*, 2003. **31**(2): p. E2.
87. Frommer, M., et al., *A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands*. *Proc Natl Acad Sci U S A*, 1992. **89**(5): p. 1827-31.
88. Herman, J.G., et al., *Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands*. *Proc Natl Acad Sci U S A*, 1996. **93**(18): p. 9821-6.
89. Xiong, Z. and P.W. Laird, *COBRA: a sensitive and quantitative DNA methylation assay*. *Nucleic Acids Res*, 1997. **25**(12): p. 2532-4.
90. Eads, C.A., et al., *MethyLight: a high-throughput assay to measure DNA methylation*. *Nucleic Acids Res*, 2000. **28**(8): p. E32.
91. Eads, C.A., et al., *CpG island hypermethylation in human colorectal tumors is not associated with DNA methyltransferase overexpression*. *Cancer Res*, 1999. **59**(10): p. 2302-6.
92. Lo, Y.M., et al., *Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction*. *Cancer Res*, 1999. **59**(16): p. 3899-903.
93. Colella, S., et al., *Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites*. *Biotechniques*, 2003. **35**(1): p. 146-50.

94. Lewin, J., et al., *Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplicates*. *Bioinformatics*, 2004. **20**(17): p. 3005-12.
95. Balog, R.P., et al., *Parallel assessment of CpG methylation by two-color hybridization with oligonucleotide arrays*. *Anal Biochem*, 2002. **309**(2): p. 301-10.
96. Gitan, R.S., et al., *Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis*. *Genome Res*, 2002. **12**(1): p. 158-64.
97. Adorjan, P., et al., *Tumour class prediction and discovery by microarray-based DNA methylation analysis*. *Nucleic Acids Res*, 2002. **30**(5): p. e21.
98. Bibikova, M., et al., *High-throughput DNA methylation profiling using universal bead arrays*. *Genome Res*, 2006. **16**(3): p. 383-93.
99. Gonzalgo, M.L., et al., *Identification and characterization of differentially methylated regions of genomic DNA by methylation-sensitive arbitrarily primed PCR*. *Cancer Res*, 1997. **57**(4): p. 594-9.
100. Toyota, M., et al., *Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification*. *Cancer Res*, 1999. **59**(10): p. 2307-12.
101. Costello, J.F., et al., *Aberrant CpG-island methylation has non-random and tumour-type-specific patterns*. *Nat Genet*, 2000. **24**(2): p. 132-8.
102. Kawai, J., et al., *Methylation profiles of genomic DNA of mouse developmental brain detected by restriction landmark genomic scanning (RLGS) method*. *Nucleic Acids Res*, 1993. **21**(24): p. 5604-8.
103. Brock, G.J., et al., *A novel technique for the identification of CpG islands exhibiting altered methylation patterns (ICEAMP)*. *Nucleic Acids Res*, 2001. **29**(24): p. E123.
104. Strichman-Almashanu, L.Z., et al., *A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes*. *Genome Res*, 2002. **12**(4): p. 543-54.
105. Hu, M., et al., *Distinct epigenetic changes in the stromal cells of breast cancers*. *Nat Genet*, 2005. **37**(8): p. 899-905.
106. Bock, C., et al., *CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure*. *PLoS Genet*, 2006. **2**(3): p. e26.
107. Yan, P.S., et al., *Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays*. *Cancer Res*, 2001. **61**(23): p. 8375-80.
108. Huang, T.H., M.R. Perry, and D.E. Laux, *Methylation profiling of CpG islands in human breast cancer cells*. *Hum Mol Genet*, 1999. **8**(3): p. 459-70.
109. Tompa, R., et al., *Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3*. *Curr Biol*, 2002. **12**(1): p. 65-8.
110. Ching, T.T., et al., *Epigenome analyses using BAC microarrays identify evolutionary conservation of tissue-specific methylation of SHANK3*. *Nat Genet*, 2005. **37**(6): p. 645-51.
111. Schumacher, A., et al., *Microarray-based DNA methylation profiling: technology and applications*. *Nucleic Acids Res*, 2006. **34**(2): p. 528-42.

112. Weber, M., et al., *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells*. Nat Genet, 2005. **37**(8): p. 853-62.
113. Yuan, E., et al., *A single nucleotide polymorphism chip-based method for combined genetic and epigenetic profiling: validation in decitabine therapy and tumor/normal comparisons*. Cancer Res, 2006. **66**(7): p. 3443-51.
114. Hatada, I., et al., *Genome-wide profiling of promoter methylation in human*. Oncogene, 2006. **25**(21): p. 3059-64.
115. Lippman, Z., et al., *Profiling DNA methylation patterns using genomic tiling microarrays*. Nat Methods, 2005. **2**(3): p. 219-24.
116. Cross, S.H., et al., *Purification of CpG islands using a methylated DNA binding column*. Nat Genet, 1994. **6**(3): p. 236-44.
117. Dean, F.B., et al., *Comprehensive human genome amplification using multiple displacement amplification*. Proc Natl Acad Sci U S A, 2002. **99**(8): p. 5261-6.
118. Team, R.D.C., *R: A language and environment for statistical computing*. 2006, R Foundation for Statistical Computing, Vienna, Austria. .
119. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics*. Genome Biol, 2004. **5**(10): p. R80.
120. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
121. Kennedy, G.C., et al., *Large-scale genotyping of complex DNA*. Nat Biotechnol, 2003. **21**(10): p. 1233-7.
122. Hayward, J.L., et al., *Assessment of response to therapy in advanced breast cancer. A project of the programme on clinical oncology of the International Union against Cancer, Geneva, Switzerland*. Eur J Cancer, 1978. **14**(11): p. 1291-2.
123. Yan, P.S., et al., *CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer*. Clin Cancer Res, 2000. **6**(4): p. 1432-8.
124. Yan, P.S., et al., *Use of CpG island microarrays to identify colorectal tumors with a high degree of concurrent methylation*. Methods, 2002. **27**(2): p. 162-9.
125. Lewin, J., et al., *Comparative DNA methylation analysis in normal and tumour tissues and in cancer cell lines using differential methylation hybridisation*. Int J Biochem Cell Biol, 2007. **39**(7-8): p. 1539-50.
126. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer*. Nat Rev Genet, 2002. **3**(6): p. 415-28.
127. Eckhardt, F., et al., *DNA methylation profiling of human chromosomes 6, 20 and 22*. Nat Genet, 2006. **38**(12): p. 1378-85.
128. Rakyan, V.K., et al., *The marks, mechanisms and memory of epigenetic states in mammals*. Biochem J, 2001. **356**(Pt 1): p. 1-10.
129. Yamada, Y., et al., *A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q*. Genome Res, 2004. **14**(2): p. 247-66.
130. Baylin, S.B. and J.E. Ohm, *Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction?* Nat Rev Cancer, 2006. **6**(2): p. 107-16.

131. Feinberg, A.P., R. Ohlsson, and S. Henikoff, *The epigenetic progenitor origin of human cancer*. Nat Rev Genet, 2006. **7**(1): p. 21-33.
132. Sathyanarayana, U.G., et al., *Aberrant promoter methylation and silencing of laminin-5-encoding genes in breast carcinoma*. Clin Cancer Res, 2003. **9**(17): p. 6389-94.
133. Yuan, B.Z., et al., *Aberrant gene expression in human non small cell lung carcinoma cells exposed to demethylating agent 5-aza-2'-deoxycytidine*. Neoplasia, 2004. **6**(4): p. 412-9.
134. Chen, J., et al., *The t(1;3) breakpoint-spanning genes LSAMP and NORE1 are involved in clear cell renal cell carcinomas*. Cancer Cell, 2003. **4**(5): p. 405-13.
135. Fujimori, A., et al., *Silencing and selective methylation of the normal topoisomerase I gene in camptothecin-resistant CEM/C2 human leukemia cells*. Oncol Res, 1996. **8**(7-8): p. 295-301.
136. Sato, H., et al., *Frequent epigenetic inactivation of DICKKOPF family genes in human gastrointestinal tumors*. Carcinogenesis, 2007. **28**(12): p. 2459-66.
137. Rahmatpanah, F.B., et al., *Differential DNA methylation patterns of small B-cell lymphoma subclasses with different clinical behavior*. Leukemia, 2006. **20**(10): p. 1855-62.
138. Yoshikawa, H., et al., *WNT10B functional dualism: beta-Catenin/Tcf-dependent growth promotion or independent suppression with deregulated expression in cancer*. Mol Biol Cell, 2007. **18**(11): p. 4292-303.
139. Grunwald, C., et al., *Expression of multiple epigenetically regulated cancer/germline genes in nonsmall cell lung cancer*. Int J Cancer, 2006. **118**(10): p. 2522-8.
140. Cho, B., et al., *Promoter hypomethylation of a novel cancer/testis antigen gene CAGE is correlated with its aberrant expression and is seen in premalignant stage of gastric carcinoma*. Biochem Biophys Res Commun, 2003. **307**(1): p. 52-63.
141. Lee, T.S., et al., *DNA hypomethylation of CAGE promoters in squamous cell carcinoma of uterine cervix*. Ann N Y Acad Sci, 2006. **1091**: p. 218-24.
142. Cho, B., et al., *Identification and characterization of a novel cancer/testis antigen gene CAGE*. Biochem Biophys Res Commun, 2002. **292**(3): p. 715-26.
143. McGowan, P.M., et al., *ADAM-17 expression in breast cancer correlates with variables of tumor progression*. Clin Cancer Res, 2007. **13**(8): p. 2335-43.
144. Longley, D.B. and P.G. Johnston, *Molecular mechanisms of drug resistance*. J Pathol, 2005. **205**(2): p. 275-92.
145. Gottesman, M.M., T. Fojo, and S.E. Bates, *Multidrug resistance in cancer: role of ATP-dependent transporters*. Nat Rev Cancer, 2002. **2**(1): p. 48-58.
146. Wilson, T.R., D.B. Longley, and P.G. Johnston, *Chemoresistance in solid tumours*. Ann Oncol, 2006. **17 Suppl 10**: p. x315-24.
147. Dahl, E., et al., *Molecular profiling of laser-microdissected matched tumor and normal breast tissue identifies karyopherin alpha2 as a potential novel prognostic marker in breast cancer*. Clin Cancer Res, 2006. **12**(13): p. 3950-60.

148. Christensen, C., et al., *Proteolytic processing converts the repelling signal Sema3E into an inducer of invasive growth and lung metastasis*. *Cancer Res*, 2005. **65**(14): p. 6167-77.
149. Chakrabarti, J., et al., *The transcription factor DEC1 (stra13, SHARP2) is associated with the hypoxic response and high tumour grade in human breast cancers*. *Br J Cancer*, 2004. **91**(5): p. 954-8.
150. Sourvinos, G., C. Tsatsanis, and D.A. Spandidos, *Overexpression of the Tpl-2/Cot oncogene in human breast cancer*. *Oncogene*, 1999. **18**(35): p. 4968-73.
151. Sissung, T.M., et al., *Pharmacogenetics and regulation of human cytochrome P450 1B1: implications in hormone-mediated tumor metabolism and a novel target for therapeutic intervention*. *Mol Cancer Res*, 2006. **4**(3): p. 135-50.
152. Ogawa, Y., et al., *Clinical role of ABCF2 expression in breast cancer*. *Anticancer Res*, 2006. **26**(3A): p. 1809-14.
153. Gill, P.K., A. Gescher, and T.W. Gant, *Regulation of MDR1 promoter activity in human breast carcinoma cells by protein kinase C isozymes alpha and theta*. *Eur J Biochem*, 2001. **268**(15): p. 4151-7.

5.13. Lebenslauf

Zur Person:	Anne Faßbender, geboren am 12.11.1975 in Erfurt
	ledig, einen Sohn: Julius Faßbender
Adresse:	Kadiner Str. 15, 10243 Berlin
email:	anne.fassbender@epigenomics.com
Schulausbildung:	1982 – 1991: Polytechnische Oberschule 1991 – 1994: Abitur am Albert Einstein Gymnasium Erfurt
Studium:	1994 – 2001: Studium der Biotechnologie an der Technischen Universität Berlin, Abschluss als Diplom-Ingenieurin für Biotechnologie mit dem Gesamturteil: Mit Auszeichnung
Wissenschaftliche Arbeit:	1997 – 1999: Tätigkeit am Institut für Biotechnologie der TU-Berlin im Fachgebiet Mikrobiologie und Genetik als studentische Hilfskraft. Arbeit an verschiedenen Projekten im Bereich Genetik und Biochemie von Hefen. (Veröffentlichung: Nevoigt, E.;Faßbender, A.; Stahl, U. Cells of the yeast <i>Saccharomyces cerevisiae</i> are transformable by DNA under non-artificial conditions. <i>Yeast</i> . 2000 Sep 15;16(12):1107-10.) Januar 1999 – Mai 1999: Studienarbeit, Thema Umlenkung des Glycerolstoffwechsels von <i>S. cerevisiae</i> . 2000 – 2001: Diplomarbeit bei der Epigenomics AG, Berlin in Kooperation mit dem Max Planck Institut für evolutionäre Anthropologie, Leipzig. Thema: Auf Mikroarrays basierende Analyse von DNA-Methylierungsmustern bei Menschen und Schimpansen (Veröffentlichung: Enard W, Fassbender A, Model F, Adorján P, Pääbo S, Olek A. Differences in DNA methylation patterns between humans and chimpanzees. <i>Curr Biol</i> . 2004 Feb 17;14(4):R148-9).
Berufstätigkeit:	Seit 01/2002 Angestellte bei der Epigenomics AG als Wissenschaftliche Mitarbeiterin im Bereich der Molekularbiologie.

5.14. Veröffentlichungen

Wissenschaftliche Publikationen:

Anne Fassbender, Jörn Lewin, Thomas König, Tamas Rujan, Cecile Pelet, Ralf Lesche, Jürgen Distler and Matthias Schuster.

Quantitative Genome-wide DNA Methylation Profiling on a High Density Oligonucleotide Microarray. *Methods in Molecular Medicine: Cancer Gene Profiling* (in press).

Oliver Hartmann, Frédérique Spyrtos, Nadia Harbeck, Dimo Dietrich, **Anne Fassbender** Manfred Schmitt, Serenella Eppenberger-Castori, Vincent Vuaroqueaux, Florence Lerebours, Katrin Welzel, Sabine Maier, Achim Plum, Stephan Niemann, John A. Foekens, Ralf Lesche, John W. M. Martens.

DNA methylation markers predict outcome in node-positive, estrogen receptor-positive breast cancer with adjuvant anthracycline-based chemotherapy. *Clinical Cancer Research* (in press).

Dimo Dietrich, Manuel Krispin, Jörn Dietrich, Anne Fassbender, Jörn Lewin, Nadia Harbeck, Manfred Schmitt, Serenella Eppenberger-Castori, Vincent Vuaroqueaux, Frédérique Spyrtos, John A. Foekens, Ralf Lesche, John W. M. Martens.

CDO1 Promotor Methylation is a Prognostic Biomarker for Anthracycline Treated, Estrogen Receptor-Positive, Lymph Node-Positive Breast Cancer Patients. *BMC Cancer* (in submission).

Veröffentlichte Patentanmeldungen:

Maier S., Nimmrich I., König Th., Jünemann K., Lesche R., Fassbender A., Wang-Gohrke, Foekens J., Martens J., Epigenetic markers for the treatment of breast cancer. 29.12.2005, WO2005123954

Lesche R., Fassbender A., Jünemann K., Foekens J., Martens J., Epigenetic methods and nucleic acids for the detection of breast cell proliferative disorders. 26.01.2006, WO 2006008128

Rujan T., Fassbender A., Verfahren zur Herstellung einer Mischung von Fragmenten einer Polynukleinsäure. 17.08.2006, DE 102005007185

Fassbender A., Lesche R., Piepenbrock Ch., Rujan T., Berlin K., Dislter J., König Th., Method for determining the methylation pattern of polynucleic acid. 24.08.2006, WO2006088978

Fassbender A., Rujan T., Lesche R., Verfahren zur Herstellung einer Mischung von Fragmenten einer Polynukleinsäure. 16.11.2006, DE 102005023055

Lesche R., Maier S., König Th., Fassbender A., Foekens J., Martens J., Method and nucleic acids for the improved treatment of breast cancers. 22.02.2007, WO2007019670

Maier S., Weiss G., Lesche R., Hildmann Th., Plum A., Foekens J., Martens J., Eppenberger-Castori, Vuaroquaux, Spyrtos, Schmitt, Harbeck, Höfler, Hartmann O., Fassbender A., Dietrich D., Method and nucleic acids for the improved detection and treatment of cancers. 02.08.2007, WO2007085497

Poster:

Anne Faßbender, Jana Papassotiriou, Ralf Lesche, Jürgen Distler, Kurt Berlin, Tamas Rujan, An Optimised Method for Genome Wide Discovery of Differential Methylated DNA Marker on a Newly Developed CpG Island Array. 2nd BMBF Colloquium 'Proteomics' Mai 2005.

Matthias Schuster, Anne Fassbender, Jörn Lewin and Thomas König, Quantitative Genome-Wide Methylation Profiling. NGFN Meeting 2006, November 2006.

Ralf Lesche, Oliver Hartmann, John W. M. Martens, Dimo Dietrich, Anne Fassbender, Sabine Maier, Gunter Weiss, Achim Plum, Nadia Harbeck, Manfred Schmitt, Serenella Eppenberger-Castori, Frédérique Spyrtos, John A. Foekens, Identification of Novel DNA Methylation Markers Predicting Outcome in Node-Positive, Anthracycline-Treated Breast Cancer Patients. San Antonio Breast Cancer Meeting 2006, December 2006.

Xiaoling Song, Ralf Lesche, Anne Fassbender, Fabian Model, Andrew Sledziewski, Development of DNA Methylation Biomarkers for Early Detection of Breast Cancer. Gordon conference - New Frontiers in cancer detection & diagnosis January 2007.