

AUS DEM LEHRSTUHL  
FÜR INNERE MEDIZIN III  
PROF. DR. MED. WOLFGANG HERR  
DER FAKULTÄT FÜR MEDIZIN  
DER UNIVERSITÄT REGENSBURG

EVOLUTION EPIGENETISCHER ABERRATIONEN BEI DER MALIGNEN  
PROGRESSION VON BRUSTKREBS

Inaugural – Dissertation  
zur Erlangung des Doktorgrades  
der Medizin

der  
Fakultät für Medizin  
der Universität Regensburg

vorgelegt von  
Sophie Gaertner

2021



AUS DEM LEHRSTUHL  
FÜR INNERE MEDIZIN III  
PROF. DR. MED. WOLFGANG HERR  
DER FAKULTÄT FÜR MEDIZIN  
DER UNIVERSITÄT REGENSBURG

EVOLUTION EPIGENETISCHER ABERRATIONEN BEI DER MALIGNEN  
PROGRESSION VON BRUSTKREBS

Inaugural – Dissertation  
zur Erlangung des Doktorgrades  
der Medizin

der  
Fakultät für Medizin  
der Universität Regensburg

vorgelegt von  
Sophie Gaertner

2021

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung:

Prof. Dr. Dirk Hellwig

Prof. Dr. Michael Rehli

Prof. Dr. Christoph Klein

18. November 2021

## Inhaltsverzeichnis

|       |  |    |
|-------|--|----|
| 1     | Einleitung.....  | 5  |
| 1.1   | Brustkrebs .....   | 5  |
| 1.1.1 | Definition.....  | 5  |
| 1.1.2 | Inzidenz und Prävalenz .....                                 | 5  |
| 1.1.3 | Risikofaktoren.....  | 6  |
| 1.1.4 | Der Biomarker HER2/neu .....                                 | 7  |
| 1.1.5 | Das BalbNeuT-Mausmodell.....                                 | 7  |
| 1.2   | Epigenetische Grundlagen .....                               | 7  |
| 1.2.1 | DNA-Methylierung .....                                       | 8  |
| 1.2.2 | DNA-Demethylierung.....                                      | 9  |
| 1.2.3 | Histonmodifikationen .....                                   | 10 |
| 1.3   | Genetische Mechanismen der Entstehung von Brustkrebs .....   | 14 |
| 1.4   | Epigenetische Mechanismen der Entstehung von Brustkrebs..... | 14 |
| 1.4.1 | DNA-Methylierung .....                                       | 15 |
| 1.4.2 | DNA-Demethylierung.....                                      | 15 |
| 1.4.3 | Histonmodifizierungen .....                                  | 16 |
| 1.5   | Mechanismen der malignen Progression von Brustkrebs.....     | 16 |
| 1.5.1 | Genetische Mechanismen der malignen Progression.....         | 17 |
| 1.5.2 | Epigenetische Mechanismen der malignen Progression .....     | 18 |
| 1.6   | Theorie der <i>early dissemination</i> .....                 | 19 |
| 1.7   | Zielsetzung der Arbeit und Fragestellung .....               | 20 |
| 2     | Material.....  | 22 |
| 2.1   | Versuchstiere.....   | 22 |
| 2.2   | Geräte.....  | 22 |
| 2.3   | Verbrauchsmaterialien.....                                   | 23 |
| 2.4   | Chemikalien.....   | 24 |
| 2.5   | Enzyme, Nukleotide und molekularbiologische Kits.....        | 24 |
| 2.6   | Größenstandards.....   | 25 |
| 2.7   | Datenbanken und Software .....                               | 25 |
| 3     | Methoden.....  | 26 |
| 3.1   | Arbeiten mit Mäusen.....                                     | 26 |
| 3.1.1 | Züchtung, Transplantation und Operation der Mäuse.....       | 26 |
| 3.1.2 | Tötung und Dissektion der Mäuse .....                        | 26 |
| 3.2   | Verarbeitung der Gewebeproben.....                           | 27 |
| 3.2.1 | Verarbeitung von Tumoren und Metastasen.....                 | 27 |
| 3.2.2 | Gewebeverdaу des Kontrollgewebes .....                       | 27 |
| 3.3   | Allgemeine molekularbiologische Methoden.....                | 28 |
| 3.3.1 | Isolation von genomischer DNA aus Brustdrüsengewebe .....    | 28 |
| 3.3.2 | Isolation von genomischer DNA aus Zellen .....               | 28 |
| 3.3.3 | Quantifizierung der DNA.....                                 | 28 |
| 3.3.4 | <i>In vitro</i> Methylierung der DNA.....                    | 29 |
| 3.3.5 | Methyl-CpG-Immunopräzipitation (MCIp) .....                  | 29 |

|       |  |     |
|-------|--|-----|
| 3.3.6 | Chromatin-Immunopräzipitation (ChIP) .....                                     | 31  |
| 3.3.7 | Library Präparation für Next Generation Sequencing .....                       | 39  |
| 3.3.8 | Assay for Transposase-Accessible Chromatin (ATAC) .....                        | 44  |
| 3.4   | Next Generation Sequencing.....  | 48  |
| 3.5   | Bioinformatische Auswertung .....  | 51  |
| 3.5.1 | DNA-Methylierung durch MChIP.....  | 51  |
| 4     | Ergebnisse.....  | 54  |
| 4.1   | Übersicht über Proben und durchgeführte Experimente.....                       | 54  |
| 4.2   | Übersicht über die MChIP-Daten .....   | 57  |
| 4.3   | Clustering der BalbNeuT-Proben.....  | 59  |
| 4.3.1 | Multidimensionale Skalierung (MDS).....  | 59  |
| 4.3.2 | Principal Component Analyse (PCA).....   | 60  |
| 4.3.3 | T-distributed Stochastic Neighbor Embedding (t-SNE).....                       | 64  |
| 4.4   | Auswertung der Regionen .....  | 65  |
| 4.4.1 | Mean-versus-Average-Plot (MVA-Plot).....                                       | 65  |
| 4.4.2 | Annotation mittels Genomic Regions Enrichment of Annotations Tool (GREAT)..... | 67  |
| 4.4.3 | Darstellung der assoziierten Gene mittels Metascape.....                       | 77  |
| 5     | Diskussion .....   | 81  |
| 5.1   | Bewertung der Arbeit.....  | 81  |
| 5.2   | Zusammenfassung .....  | 86  |
| 6     | Literaturverzeichnis.....  | 88  |
| 7     | Danksagung .....   | 100 |

# 1 Einleitung

## 1.1 Brustkrebs

### 1.1.1 Definition

Die menschliche Brust besteht funktionell aus Milchdrüsenläppchen (Lobuli) und Milchdrüsengängen (Ducti). Dazwischen liegen Fett- und Bindegewebe, sowie lymphatisches Gewebe. Brustkrebs entsteht, wie andere Krebsformen auch, durch Mutationen in einzelnen Zellen, die den programmierten Zelltod (Apoptose) außer Kraft setzen, wodurch sich die Zellen anschließend unbegrenzt teilen. Dies führt zu einem unphysiologischen Zellwachstum, das verdrängend oder invasiv sein kann. Im Falle von Brustkrebs gehen diese mutierten Zellen aus dem Drüsengewebe, also den Lobuli oder Ducti hervor (1).

Zu den verdrängend wachsenden, nicht invasiven (*in situ*) Läsionen zählen das duktale Carcinoma in situ (DCIS), das 83 % der *in situ* Fälle ausmacht, und das lobuläre Carcinoma in situ. Diese *in situ* Läsionen sind oftmals Vorgängerläsionen des invasiven Brustkrebs und ihr Auftreten erhöht das Risiko, dass im weiteren Verlauf invasiver Brustkrebs entsteht (2). Invasiver Brustkrebs durchbricht die Gänge oder Drüsen und infiltrierte das umliegende Gewebe. Die Eindringtiefe des Tumors und das Vorliegen von Metastasen bestimmen das Stadium der Erkrankung im Sinne der TNM-Klassifikation (1). Das Stadium wiederum korreliert stark mit der Prognose. Dabei ist der Lymphknotenmetastasen-Status der wichtigste prognostische Faktor für die Überlebenszeit. Patientinnen mit positivem Lymphknotenstatus haben außerdem eine höhere Rezidivwahrscheinlichkeit als Patientinnen ohne betroffene Lymphknoten (3).

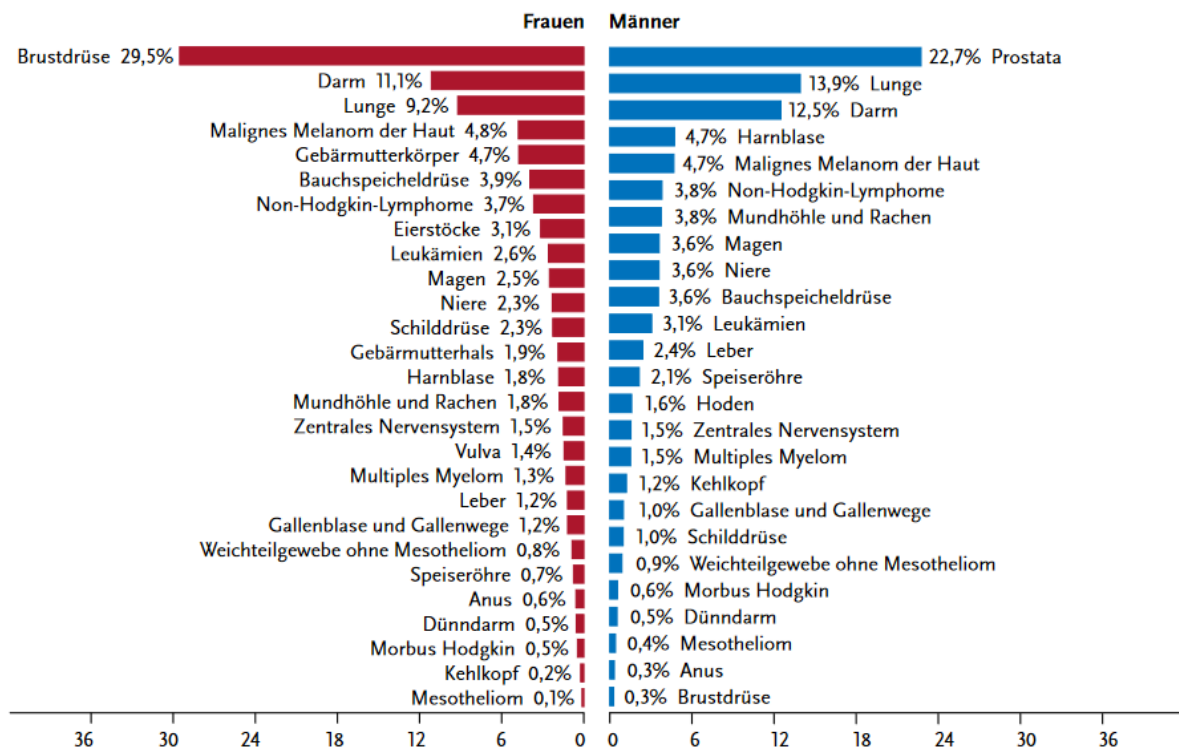
### 1.1.2 Inzidenz und Prävalenz

In Deutschland treten jährlich rund 71.000 Neuerkrankungen bei Frauen und rund 700 Neuerkrankungen bei Männern auf (4). Somit ist Brustkrebs mit 29,5 % die häufigste Krebserkrankung bei Frauen in Deutschland (s. Abb. 1). Außerdem werden bei weiteren 6.500 Frauen *in situ* Tumore festgestellt (5). Das Erkrankungsalter bei Brustkrebs ist deutlich niedriger als bei anderen Krebsarten. 18 % der Frauen sind bei Diagnosestellung jünger als 50 Jahre und 45 % sind zwischen 50 und 69 Jahre alt, was auch dem Alter entspricht, in dem das Mammographie-Screening empfohlen wird (4). Seit Einführung des

Mammographie-Screenings im Jahr 2005 ist die Inzidenz von Brustkrebs gestiegen. Dies liegt vermutlich teilweise an einer erhöhten Rate an kleinen Läsionen, die durch das Screening detektiert werden, diese Beobachtung könnte aber auch mit der zunehmend älter werdenden Gesellschaft zusammenhängen (3).

### 1.1.3 Risikofaktoren

Brustkrebs ist eine komplexe Erkrankung mit vielfältigen Einflüssen auf Entstehung und Progression. Zu den bekannten Risikofaktoren für Brustkrebs gehören die genetische Prädisposition, sowie eine längere Exposition gegenüber den weiblichen Geschlechtshormonen Östrogen und Progesteron (z.B. über Hormonersatztherapien, frühe Menarche, späte Menopause, wenige Schwangerschaften, oder eine kurze Stillzeit). Da die in Adipozyten vorhandene Aromatase Testosteron in Östrogen umwandelt, sind auch Adipositas und mangelnde Bewegung bekannte Risikofaktoren für Brustkrebs. Des Weiteren können ionisierende Strahlung, andere atypische Hyperplasien in der Vorgeschichte und Alkoholkonsum zum Auftreten von Brustkrebs beitragen (3,4).



**Abbildung 1:** Prozentualer Anteil der häufigsten Tumorlokalisationen an allen Krebsneuerkrankungen in Deutschland 2016. Männer n = 258.520 Frauen n = 233.570. Aus „Krebs in Deutschland 2015/2016“, Robert-Koch-Institut, 2019 (5)



#### **1.1.4 Der Biomarker HER2/neu**

Brustkrebs ist durch unterschiedliche Biomarker charakterisiert. Die wichtigsten Marker in Bezug auf therapeutisches Management, adjuvante Therapie und Metastasierungsstadium sind der Östrogen- und Progesteronrezeptorstatus sowie der HER2/neu-Status (6–8). Das *human epidermal growth factor receptor 2 (HER2)* Onkogen kodiert für einen Tyrosinkinase-Rezeptor (9), der zu der Gruppe der epidermalen Wachstumsfaktor-Rezeptoren (*epidermal growth factor receptor, EGFR*) gehört. Diese spielen eine wichtige Rolle bei der Aktivierung von zellulären Signaltransduktionswegen, indem sie das epitheliale Zellwachstum und die Differenzierung (10,11), sowie möglicherweise auch die Angiogenese kontrollieren (12,13). Die Vervielfältigung von *HER2* oder die Überexpression des Proteins, für das es kodiert, findet man in 18-20 % der menschlichen Brustkrebspatienten (14–16). Als zielgerichtete Therapie wurde der monoklonale anti-HER2/neu Antikörper Trastuzumab (Herceptin®) entwickelt (17).

#### **1.1.5 Das BalbNeuT-Mausmodell**

Für diese Arbeit wurde ein transgenes Mausmodell verwendet, um die Evolution von Metastasen möglichst naturgetreu nachzustellen. Das BalbNeuT-Mausmodell stellt eines der aggressivsten Modelle der Entstehung von Brustkrebs dar. BalbNeuT-Mäuse sind transgene BALB/c-Mäuse, die das mutierte Ratten-HER2/neu Onkogen tragen. HER2/neu positive, aktiv proliferierende Zellen treten bereits in den Brustdrüsen von drei Wochen alten, jungfräulichen transgenen Weibchen auf. Diese Zellen bilden Herde einer lobulär atypischen Hyperplasie. Die multiplen Krebsherde entwickeln sich fokal, verschmelzen und sind ab der 20. Woche tastbar. Dieser Prozess geschieht beinahe synchron in allen zehn Brustdrüsen und im Alter von 30 Wochen haben sich bei 100 % der transgenen Weibchen tastbare Tumore in allen Brustdrüsen gebildet. Ab der 34. Woche tauchen HER2/neu positive Lungenmetastasen auf (18).

## **1.2 Epigenetische Grundlagen**

Der Begriff „Epigenetik“ wurde im Jahr 1942 erstmals von Conrad Waddington eingeführt. Er bezeichnete damit die kausale Interaktion von Genen und ihren Produkten, die die phänotypischen Veränderungen im Laufe der Entwicklung

kontrolliert (19). Über die Jahrzehnte hat sich die Definition von Epigenetik gewandelt. Sie konzentriert sich nun auf Veränderungen, die in der Mitose und/oder Meiose vererbt werden, ohne dabei die zugrundeliegende DNA-Sequenz zu verändern. Es sind mehrere epigenetische Mechanismen beschrieben, die die Genexpression von Säugetieren regulieren. Die vier Hauptkomponenten sind DNA-Methylierung, Histonmodifizierung, Chromatinkonformation und nicht-kodierende RNA (ncRNAs) (20). Da DNA-Methylierung und Histonmodifikationen für die Transkription von Genen entscheidend sind (21), werden diese zwei Mechanismen im Weiteren näher beschrieben.

### **1.2.1 DNA-Methylierung**

Die DNA-Methylierung ist der am besten erforschte Aspekt der Epigenetik und spielt eine entscheidende Rolle in der Entwicklung von Säugetieren. Dabei wird eine Methylgruppe an das fünfte Kohlenstoffatom am Ring des Cytosins angehängt (Position C-5) und es entsteht 5-Methylcytosin (5mC), die 1948 erstmals entdeckte „fünfte Base“ des menschlichen Genoms. Diese Reaktion wird von Methyltransferasen katalysiert und findet für gewöhnlich an sogenannten CpG-Dinukleotiden (Cytosin-Guanin-Dinukleotiden) mit der Sequenz 5'-CG-3' statt (22). Jedoch wurden in embryonalen Stammzellen auch geringe Mengen an nicht-CG-Methylierung nachgewiesen, und zwar an CHG- und CHH-Sequenzen (H steht für A, T oder C, also Adenin, Thymin oder Cytosin) (20).

Im Genom von Säugetieren sind in etwa 75 % der CpG-Dinukleotide methyliert. Es gibt CpG-Regionen, die stabil methyliert sind, jedoch auch eine kleine Anzahl dynamisch methylierter Regionen (20). Ansammlungen von CpG-Dinukleotiden nennt man CpG-Inseln, sie finden sich an vielen 5'-Enden von menschlichen Genen. CpG-Inseln sind im Säugetiergenom häufig unmethyliert, selbst wenn die assoziierten Gene nicht exprimiert werden (23). Bei malignen Zellveränderungen werden CpG-Inseln allerdings häufig methyliert (24).

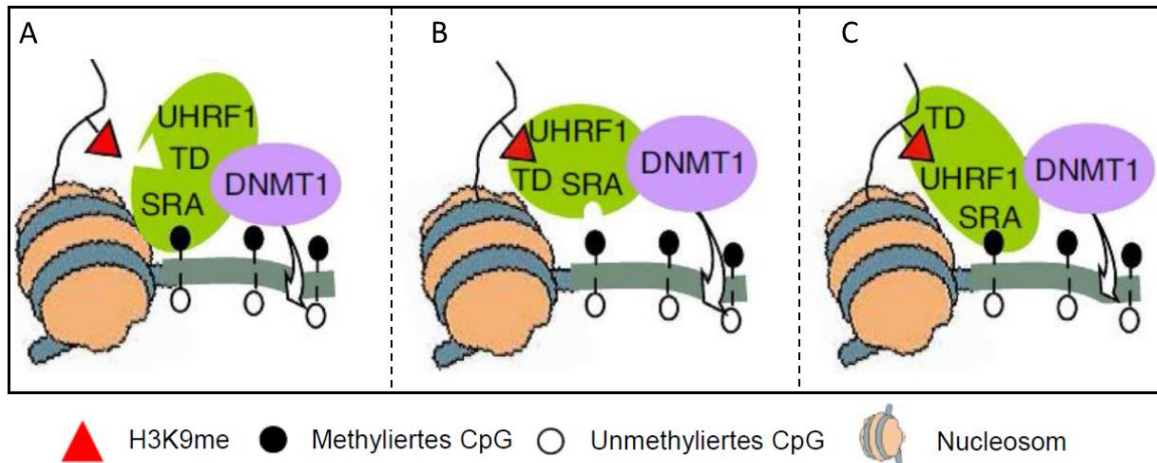
Die Hauptaufgabe der DNA-Methylierung ist es, die Bindung von Transkriptionsfaktoren an regulatorischen Sequenzen zu verhindern und auf diese Art und Weise die Genexpression zu kontrollieren. Unlängst fand man jedoch auch Repressorproteine, die dazu in der Lage sind, an methylierte DNA zu binden und sie abzulesen und damit zur Repression der Transkription beizutragen. Des

Weiteren spielt DNA-Methylierung auch bei der Prägung, der Inaktivierung des X-Chromosoms und dem Abschalten repetitiver DNA eine Rolle (20,25).

DNA-Methylierung wird durch eine Familie von DNA-Methyltransferasen (DNMTs) katalysiert und kann in zwei unterschiedliche Prozesse unterteilt werden. Zum einen tritt DNA-Methylierung als Erhaltungs-Methylierung während der Zellteilung auf. Diese sogenannte *Maintenance*-Methylierung wird durch die DNA-Methyltransferase DNMT1 ausgeführt und ist für die reziproke Methylierung des neu replizierten DNA-Strangs zuständig. Zum anderen gibt es *de novo*-Methyltransferasen, DNMT3a und DNMT3b, die unmethylierte doppelsträngige DNA methylieren können (21,26).

### **1.2.2 DNA-Demethylierung**

Es sind mehrere Mechanismen beschrieben, die zu einer DNA-Demethylierung führen könnten. Aktive DNA-Demethylierung involviert Methylcytosin-Dioxygenasen, auch TET-Proteine genannt, die 5-Methylcytosin (5mC) zu 5-Hydroxymethylcytosin (5hmC) oxidieren (27). Anschließend kommt es zu einer Desaminierung von 5hmC. Die daraus entstehende Basenfehlpaarung wird durch Basenexzisionsreparaturmechanismen korrigiert (28). Passive Demethylierung wird größtenteils durch relative Inaktivität der *Maintenance*-Methyltransferase DNMT1 verursacht, wodurch nach der Replikation der komplementäre Strang einer hemimethylierten DNA nicht methyliert wird. Zusätzlich kann passive Demethylierung durch aktive Demethylierung verstärkt werden, da das Vorhandensein von 5hmC die Aktivität von DNMT1 vermindert (29–32). Die zugrundeliegenden Mechanismen sind allerdings noch nicht eindeutig geklärt. Um an methylierte DNA zu binden, benötigt DNMT1 das nukleare Protein UHRF1, welches eine spezielle SET- und RING-assoziierte Domäne (SRA) besitzt. Damit bindet UHRF1 bevorzugt an hemimethylierte, soeben neu synthetisierte DNA und rekrutiert DNMT1 dorthin (33). *In vitro* erkennt UHRF1 sowohl methylierte als auch hydroxymethylierte Cytosine (34), die DNMT1-abhängige Methylierungsreaktion scheint aber durch 5hmC gestört zu werden (35) (Abb. 2).



**Abbildung 2:** Verschiedene Modelle für die UHRF1-vermittelte Rekrutierung von DNMT1 zur *Maintenance*-Methylierung. UHRF1 kann sowohl an H3K9me, als auch an hemimethylierte DNA binden (Abb. 2A und B). Am wahrscheinlichsten ist die gleichzeitige Bindung an H3K9me und hemimethylierte DNA (Abb. 2C). Modifiziert nach Liu, 2013 (33).

### 1.2.3 Histonmodifikationen

Die genetische Erbinformation liegt im menschlichen Zellkern als Chromatin vor. Chromatin besteht aus DNA, Histonproteinen und Nicht-Histonproteinen. Histone sind kleine Proteine, die aus einer kugelförmigen Domäne und einem beweglicheren, positiv geladenen NH<sub>2</sub>-Terminus, meist Arginin oder Lysin, aufgebaut sind. Diese positiv geladenen Histone lagern sich der DNA mit ihrem negativ geladenen Zucker-Phosphat-Rückgrat an, um räumlich besser in den Zellkern zu passen und vor Scherkräften geschützt zu sein. Die entstehende Struktur aus DNA-Doppelhelix, die zweimal um ein Oktamer aus Histon-Proteinen gewickelt ist, nennt man Nucleosom. Das Histon-Oktamer wird aus einem H3-H4-Tetramer und zwei H2A-H2B-Dimeren gebildet. Wenn DNA im Sinne der Replikation oder Transkription abgelesen werden soll, wird sie zur besseren Zugänglichkeit aufgelockert. Dies geschieht durch verschiedene posttranslationale Modifikationen (PTM) der positiv geladenen NH<sub>2</sub>-Termini der Histone (36). Zu diesen PTM gehören Methylierung, Acetylierung, Phosphorylierung, Citrullinierung, SUMOylierung, ADP-Ribosylierung, Deiminierung, und Crotonylierung (37). Durch einige dieser Histonmodifikationen kann sich die Ladung des Chromatins verändern, wodurch sich die Bindung zwischen DNA und Histonen lockert und die Zugänglichkeit des Chromatins für regulatorische Proteine, wie z.B. Transkriptionsregulatoren, erhöht wird (20).

### 1.2.3.1 Histonmethylierung

Histonmethylierung findet an den NH<sub>2</sub>-Termini der basischen Aminosäuren Arginin und Lysin statt und wird durch Histon-Methyltransferasen (HMTs) vermittelt (Abb. 3). Dabei können Lysinreste durch Lysin-Methyltransferasen (KMT) mono-, di- oder trimethyliert und Argininreste mono- oder dimethyliert werden (38). Durch die Methylierung ändert sich im Gegensatz zur Histonacetylierung die Ladung der NH<sub>2</sub>-Termini des Lysins nicht (38). Es gibt jedoch spezielle Enzyme, die diese methylierten Aminosäuren spezifisch erkennen (39). Diese ablesenden Enzyme, z.B. das ATP-abhängige Chromatin-Remodellierungs-Enzym (CHD1) und den Bromodomänen-PHD-Finger-Transkriptionsfaktor (BPTF) (40), können weitere enzymatische Veränderungen auslösen, welche schließlich das umgebende Chromatin verändern und die Transkription beeinflussen (38).

Im Gegensatz zu anderen Histonmodifikationen vermittelt Histonmethylierung aktive oder unterdrückte Transkription abhängig von ihrer Position und dem Methylierungsstatus (41). Generell wird Methylierung an H3K4, H3K36 und H3K79 aktiver Transkription zugeschrieben, während H3K9, H3K27 und H4K20 im methylierten Zustand als stillgelegtes Chromatin gelten (41). H3K4-Methylierung ist als Marker für aktive Transkription an Promotorregionen und Transkriptionsstartpunkten erhöht (42).

Seit der Entdeckung der ersten Histonmethyltransferasen (43) wurde verschiedenen Proteinfamilien die Fähigkeit zugeschrieben, die Übertragung von Methylgruppen von dem Methyl donor S-Adenosylmethionin (SAM) auf Histone zu katalysieren (41). Die Mitglieder der SET-Domänen-Familie und die Lysin-N-Methyltransferase DOT1L methylieren beide die Aminosäure Lysin, während Arginin durch die Familie der Arginin-Methyltransferasen methyliert wird (38).

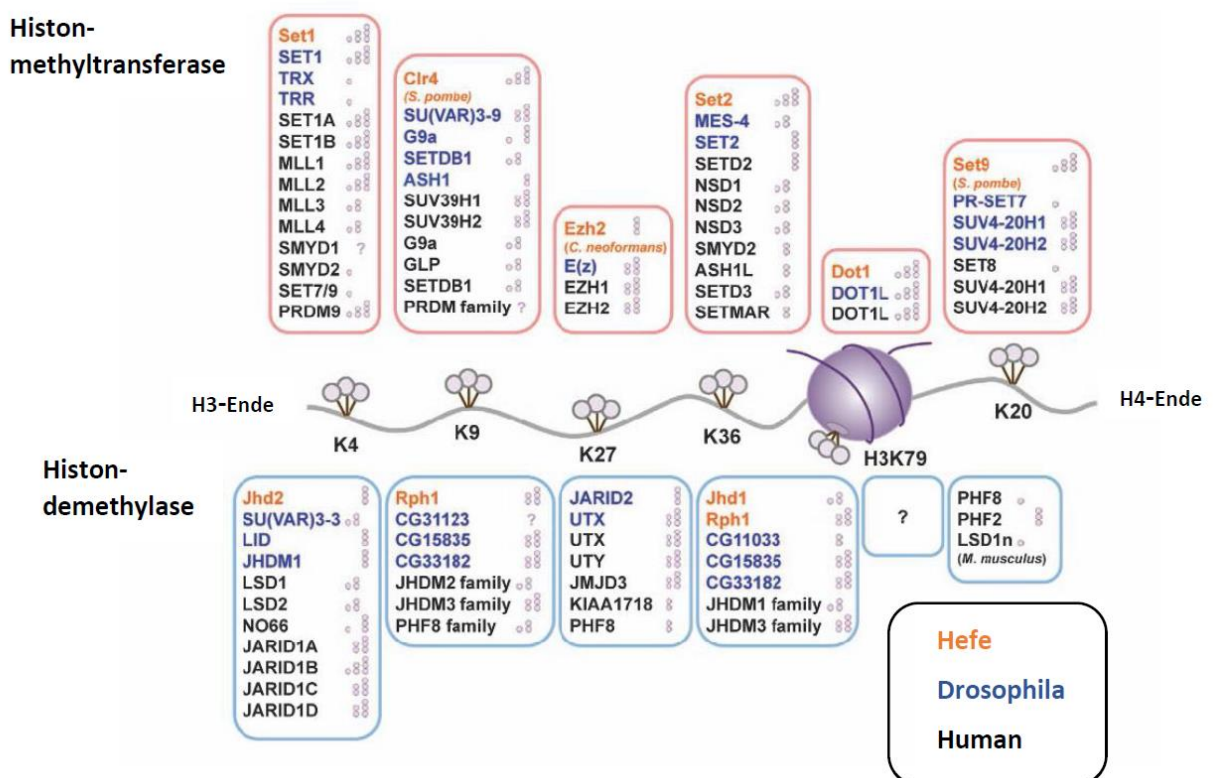
In Hefezellen werden alle H3K4-Methylierungen durch die Set1-Methyltransferase durchgeführt. Diese Methyltransferase blieb von der Hefe bis zum Menschen weitestgehend erhalten. In Säugetieren finden sich sechs Homologe von Set1 (SET1A/KMT2F, SET1B/KMT2G, MLL1 (mixed-lineage-leukemia1)/KMT2A, MLL2/KMT2B, MLL3/KMT2C, MLL4/KMT2D) (42). Diese Methyltransferasen bestehen jeweils aus einem Komplex von verschiedenen Untereinheiten. Wie genau diese Enzymkomplexe differenzielle Methylierung bewirken, konnte bisher noch nicht abschließend gezeigt werden (40).

Methyliertes H3K9 wird mit stillgelegtem Chromatin und Heterochromatin in Verbindung gebracht (40). In Säugetieren gibt es diverse Methyltransferasen, die die Methylierung von H3K9 vermitteln: SUV39H1/KMT1A, SUV39H2/KMT1B, SETDB1/KMT1E, dimeres G9a/KMT1C-GPL (G9a-like protein)/KMT1D und die Familie der PRDM-Proteine (PRDI-BF1- und RIZ-Homologie-Domäne) (44).

H3K27me<sub>3</sub> ist ebenfalls ein Marker für die Repression der Transkription (40). Die für alle drei Methylierungszustände (me<sub>1</sub>, me<sub>2</sub> und me<sub>3</sub>) von H3K27 zuständige Histonmethyltransferase ist die Untereinheit EZH2 (enhancer-of-zest homolog 2) des Polycomb Repressive Complex 2 (PRC2) (45,46).

Bei Säugetieren gibt es mindestens acht Histonmethyltransferasen, die H3K36 methylieren (47). Die wichtigsten hierfür sind NDS1-3 (nukleäre Rezeptor-bindende SET-Domäne) und SETD2, wobei SETD2 als einzige der acht die Trimethylierung von H3K36 katalysiert (48).

H3K79 befindet sich im Gegensatz zu den anderen Lysinmodifikationen am globulären Anteil des Histons H3. Die Methylierung wird auch nicht durch eine SET-Domäne vermittelt, sondern durch DOT1 (disruptor of telomeric silencing) (49).



**Abbildung 3:** Schematische Darstellung eines Nucleosoms mit den Methylierungsstellen an Lysin in H3 und H4 und den zugehörigen Histonmethyltransferasen und -demethylasen. Zusätzlich wird der spezifische Methylierungsstatus gezeigt: von einem Kreis (me<sub>1</sub>) über zwei Kreise (me<sub>2</sub>) bis drei Kreise (me<sub>3</sub>). Modifiziert nach Hyun, 2017 (40).

### 1.2.3.2 Histon-Demethylierung

Ursprünglich wurde angenommen, dass die Histonmethylierung stabil und vererbbar ist. Dies änderte sich, als die erste H3K4-Histon-Demethylase, die FAD-abhängige Aminooxidase LSD1 (= KDM1A), entdeckt wurde (50). Es gibt zwei verschiedene Klassen von Demethylasen, die verschiedene biochemische Reaktionen nutzen, um eine Demethylierung zu bewirken (21). Die Lysin-spezifische Demethylase 1 (LSD1) demethyliert H3K4me1 und H3K4me2, sowie H3K9me1 und H3K9me2 (50–52). Zur zweiten Gruppe der Histon-Demethylasen gehören mehrere Proteine, die eine JMJC-Domäne (=Jumonji C) besitzen und damit trimethylierte Histonreste demethylieren können (53). Hierzu gehört die JARID1-Proteinfamilie (Jumonji AT-reiche interaktive Domäne-1) und das Protein NO66 (Nucleolus-Protein 66), welches ebenfalls eine JMJC-Domäne enthält (38).

### 1.2.3.3 Histonacetylierung

Eine weitere wichtige Histonmodifikation ist die Histonacetylierung. Sie findet am Aminosäurerest Lysin an den Histonen H2A, H2B, H3 und H4 statt. Durch Acetylierung wechselwirkt die negative geladene Acetylgruppe mit dem ebenfalls negativ geladenen Zucker-Phosphat-Rückgrat der DNA und reduziert dadurch die Bindungsaffinität der Histone an die DNA (54). Man nimmt an, dass sich kondensiertes Heterochromatin dadurch lockert und die Transkription erleichtert wird. Niedrige Mengen an Acetylierung sind mit stillem Heterochromatin assoziiert. Aktiv transkribiertes Euchromatin ist typischerweise stark acetyliert (54).

Die Acetylgruppe wird durch Histonacetyltransferasen (HATs), die speziell bei Acetylierung von Lysin auch Lysin (K)-Acetyltransferasen (KATs) genannt werden, vermittelt. Die Histondeacetylasen (HDACs) sind für die Entfernung der Acetylgruppen zuständig (37). Acetyltransferasen werden in drei große Familien unterteilt: GNAT, MYST und CBP/p300 (55). Deacetylierung ist mit einer Unterdrückung der Transkription assoziiert (54). Es gibt vier Familien von Histon-Deacetylasen (Klasse I-IV), wobei Klasse III aus den 7 Homologen der Sirtuine, auch Sir2-like Proteine, besteht (56,57).

Histonacetyltransferasen können auch Nichthiston-Proteine wie p53, MYC und Rb acetylieren (37) und sind zusätzlich in verschiedene Transkriptionsaktivator-komplexen involviert. In Transkriptionsrepressorkomplexen hingegen findet man Histon-Deacetylierungsaktivität, was wiederum die transkriptionsaktivierende

Wirkung der Histonacetylierung unterstreicht (20). Obwohl eher die HATs mit transkriptioneller Aktivität assoziiert sind, benötigt es sowohl HATs als auch HDACs um eine regelrecht regulierte Gentranskription zu gewährleisten (58). Neben den strukturellen Auswirkungen der Acetylierung, kann diese auch durch „Reader“ wie Bromodomänen erkannt werden und kann somit als Zielstruktur für epigenetische Therapien dienen (37,54).

### **1.3 Genetische Mechanismen der Entstehung von Brustkrebs**

20-25 % aller Frauen mit Brustkrebs haben ein nahes Familienmitglied, das an Brustkrebs erkrankt war, was dafür spricht, dass Brustkrebs zum Teil genetische Ursachen haben könnte (59).

Allerdings findet man nur bei rund 5-10 % der Frauen mit Brustkrebs eine genetische Veranlagung, zum Beispiel in den bekannten prädisponierenden Hochrisikogenen *BRCA1* oder *BRCA2* (*Breast Cancer Gene 1 & 2*). Frauen mit einer Mutation in einem *BRCA*-Gen haben ein Risiko von 50-80 % im Laufe ihres Lebens an einem Mammakarzinom zu erkranken (60,61). Ein ähnlich hohes Risiko scheinen die Gene *RAD51C* und *RAD51D* (*RAD51 Rekombinase Homolog C und D*) zu vermitteln, wobei das Ovarialkarzinomrisiko bei diesen Genen höher als das Mammakarzinomrisiko sein könnte (62,63).

### **1.4 Epigenetische Mechanismen der Entstehung von Brustkrebs**

Neben genetischen sind auch epigenetische Mechanismen bekannt, die einen Einfluss auf die Entstehung von Brustkrebs haben. Wie bereits oben erläutert, beeinflusst der epigenetische Status die Genexpression, was nahelegt, dass epigenetische Veränderungen einen Beitrag zur Tumorentstehung leisten könnten. So konnte gezeigt werden, dass sich DNA-Methylierungsmuster und die Verteilung von Histonmodifikationen zwischen normalem Gewebe und Tumoren, die aus diesem Gewebe entstanden sind, unterscheiden. Da dies allerdings nur als Zusammenhang beschrieben ist, ist bisher nicht klar, warum diese epigenetischen Veränderungen auftreten und ob sie den Grund für das Tumorwachstum darstellen, oder lediglich als Antwort auf die veränderte Umgebung entstehen (21).



### 1.4.1 DNA-Methylierung

In Krebszellen tritt eine globale Hypomethylierung auf, sowie eine DNA-Hypermethylierung in vielen CpG-Inseln, also oftmals im Promotorbereich bestimmter Gene (64). Wie unter 1.2.1 bereits erläutert, führt DNA-Methylierung in Promotorregionen meist zu verminderter Transkription und könnte daher die Stilllegung von Genen zur Folge haben (25).

Diese Methylierungen sind häufig in Promotoren von Tumorsuppressorgenen zu finden. Dies könnte zu einer verminderten Kontrolle des Zellzyklus führen und damit zu einer erhöhten Proliferationsrate, was eine mögliche Erklärung dafür ist, wie Methylierung zu dem Veränderungsprozess im Rahmen von Krebserkrankungen beitragen könnte (21). So konnte beispielsweise vor kurzem das Gen *FLRT2* (*Fibronectin leucine rich transmembrane protein 2*) als neues Tumorsuppressorgen bei Brustkrebs identifiziert werden, das durch Hypermethylierung inaktiviert wird (65). Gleichmaßen konnte gezeigt werden, dass Promotorregionen von Proto-Onkogenen in Krebszellen oftmals hypomethyliert sind, was eine erhöhte Transkriptionsrate und eine vermehrte Aberrationswahrscheinlichkeit nach sich zieht (66). Deshalb wird angenommen, dass sich eine aberrante Methylierung auf die Entstehung von Krebs im Allgemeinen und Brustkrebs im Speziellen auswirkt (67,68).

### 1.4.2 DNA-Demethylierung

Es wird angenommen, dass auch fehlerhafte DNA-Demethylierung zur Entstehung von Krebs beiträgt. Trotz der regionalen Hypermethylierung, die bei Krebs nachgewiesen werden konnte, ist sind die globalen Spiegel an 5-Methylcytosin in Tumoren 5-10 % niedriger als in normalen Zellen (69,70). Die funktionelle Relevanz dieser globalen Hypomethylierung konnte noch nicht nachgewiesen werden, doch es wird diskutiert, ob sie zu genomischer Instabilität führen könnte (71).

Da die DNA von benignen Kolonpolyphen und von maligne entartetem Gewebe im Vergleich zu normalem Gewebe deutlich hypomethyliert ist, lässt sich vermuten, dass die Demethylierung zwischen den Stadien Hyperplasie und benigne Neoplasie auftreten könnte (72). Das legt nahe, dass Demethylierung eine entscheidende Rolle bei der Tumorentstehung spielt (21).

### 1.4.3 Histonmodifizierungen

Die Dysregulation der Histon-Methyltransferasen (HMTs) und der Histon-Demethylasen (HDMs) ist mit verschiedenen Krebsarten assoziiert, unter anderem auch mit Brustkrebs (73,74). Ähnlich wie die DNA-Methylierung könnten auch Histonmodifikationen dabei helfen, die Prognose verschiedener Krebsarten, unter anderem auch bei Brustkrebs, zu präzisieren (75–77). Auch Histonacetylierungen wie H3K4ac werden mit deregulierten Signalwegen bei Krebs und mit fortschreitender Progression in Verbindung gebracht (78). Als Beispiel bei Brustkrebs lässt sich H4K16ac und dessen zugehörige Acetyltransferase hMOF (*human males absent on the first*) nennen, welche beim primären Mammakarzinom reduziert sind (79,80). Die Histonlysinttransferase EZH2 als Untergruppe des Polycomb-Repressive-Complex 2 ist bei Brustkrebs vervielfältigt und überexprimiert (81). Auch die Histon-Lysin-Demethylase LSD1 ist bei Brustkrebs herunterreguliert und korreliert negativ mit der Expression von TGF $\beta$ 1 (*Transforming Growth Factor  $\beta$* ), was wiederum die Invasion von Mammakarzinomzellen *in vitro* verhindert und *in vivo* das metastatische Potenzial von Brustkrebs unterdrückt (82).

### 1.5 Mechanismen der malignen Progression von Brustkrebs

Die maligne Progression beschreibt die Ausbreitung von Tumorzellen im Organismus (83). Durch die Dissemination und das anschließende Wachstum der Zellen eines soliden Tumors an anderer Stelle im Körper, wird der Tumor zu einer systemischen Erkrankung. Dies geschieht in zwei Phasen, der klinisch unauffälligen Phase der Krebsausbreitung und der Phase der manifesten Metastasierung (84). Metastasierung ist für bis zu 90 % der krebsbezogenen Mortalität verantwortlich. Allerdings ist Metastasierung wohl auch die am wenigsten verstandene Komponente der Tumorphathogenese (85). Bis eine Tumorzelle des Primärtumors eine Fernmetastase formen kann, muss sie eine Reihe von Schritten durchlaufen. Zunächst dringt sie in das umliegende Gewebe und die kleinen Gefäße der Blut- und Lymphbahnen ein (*Intravasation*), wo sie größtenteils über den Blutstrom in die Kapillaren entfernter Gewebe transportiert wird. Dort verlässt die Tumorzelle den Blutstrom (*Extravasation*), überlebt im Mikromilieu dieses Gewebes und passt sich schließlich so an diese fremde Umgebung an, dass eine erneute Zellproliferation möglich ist (*Kolonisierung*) (85).

Circa 10-15 % aller Patientinnen mit der Diagnose Brustkrebs entwickeln innerhalb von 3 Jahren Fernmetastasen (86), am häufigsten im Knochen und in der Lunge (87). Metastasierter Brustkrebs ist schwierig zu behandeln, da die disseminierten Tumorzellen (DCCs, *disseminated cancer cells*), wenn sie erst einmal gestreut haben, schwer zu lokalisieren sind und auch nach Entfernung des Primärtumors für Jahre „schlafend“ überleben können (87). Um Mikrometastasen frühzeitig zu entdecken und neue Therapiestrategien zu entwickeln, ist es wichtig, die der Metastasierung zugrundeliegenden Mechanismen zu verstehen (87).

### **1.5.1 Genetische Mechanismen der malignen Progression**

Metastasierung lässt sich teilweise durch genetische Mechanismen erklären. Metastasierte Zellen weisen eine höhere Rate an Spontanmutationen auf (88), das Metastasierungspotenzial könnte also durch den Erwerb von genetischen Variationen getrieben sein (87). Außerdem wurden Gene identifiziert, die speziell in Mammatumoren exprimiert sind, die Lungenmetastasen ausbildeten und nicht in Tumoren ohne Metastasenbildung (89). Diese Gene kodieren für Proteine wie z.B. Epiregulin (ein Ligand, der spezifisch für Mitglieder der HER-Rezeptor-Familie ist), Matrix-Metalloproteinasen 1 und 2 (MMP-1 und MMP-2) und das Zelladhäsionsmolekül VCAM1 (*Vascular cell adhesion protein 1*), von denen bekannt ist, dass sie eine Rolle beim Prozess der Metastasierung spielen (89).

In der Zusammenschau verschiedener Studien lässt sich vermuten, dass Gene, die in der Metastasierung involviert sind, nicht über einen gemeinsamen molekularen Signalweg wirken. Stattdessen bilden sie eine heterogene Gruppe. So erhöhen die genetischen Veränderungen lediglich die Wahrscheinlichkeit, dass die verschiedenen Schritte der Metastasierung absolviert werden, indem sie die Aktivität krebserregender Signalwege modulieren (90). Es ist möglich, dass über eine genauere genetische Sequenzierung weitere Mutationen gefunden werden, die bei der Metastasierung eine Rolle spielen. In einem aktuellen Review vermuten Patel und Vanharanta jedoch, dass es angesichts der großen noch unerklärten Varianz wahrscheinlich ist, dass genetische Mutationen nicht der hauptsächliche Weg sind, durch den die Aktivierung von Metastasierungseigenschaften vermittelt wird (90).

### 1.5.2 Epigenetische Mechanismen der malignen Progression

Da genetische Mechanismen den Vorgang der Metastasierung nicht ausreichend erklären, wird auch epigenetischen Mechanismen ein Anteil an der malignen Progression von Krebs zugeschrieben. Experimentell konnte bei der Isolation von metastasenbildenden Zellpopulationen gezeigt werden, dass sie sich genetisch nur wenig vom Primärtumor unterschieden. Sie exprimierten jedoch Gene, die auch bei Patienten mit Metastasen besonders aktiv sind. Daraus lässt sich schließen, dass auch in menschlichen Tumoren metastasenfördernde, epigenetische Regulationsprogramme vorhanden sein müssen (90,91). Viele dieser Gene sind auch funktionell mit der Metastasierung in Verbindung gebracht worden (92–94). Bei Brustkrebs gibt es zum Beispiel eine Assoziation zwischen einem speziellen Hypermethylierungsphänotyp (B-CIMP) und einem niedrigeren Metastasierungsrisiko (95).

#### 1.5.2.1 DNA-Methylierung

Genau wie bei der Entstehung von Krebs, wurde der Methylierungsstatus auch bei metastasierten Krebszellen untersucht. So zeigte sich beim Prostatakarzinom kein Unterschied im Methylierungsmuster zwischen Primärtumor und Metastasen (96). Bei Brustkrebs konnten jedoch Metastasen-spezifische Veränderungen der Methylierung zwischen Primärtumoren und regionalen Metastasen, vor allem außerhalb CpG-reicher Regionen, festgestellt werden (97). Allerdings gilt es bei der Interpretation der Daten zu beachten, dass die aberrante Methylierung in den meisten Fällen in den Regionen auftritt, die bereits vor der malignen Entartung reprimiert waren (98–100).

Es gibt einige Beispiele dafür, dass aberrante DNA-Methylierung zur Erlangung von funktionell relevanten, metastasenfördernden Eigenschaften führt. Zum Beispiel verursacht der Verlust von DNA-Methylierung beim Nierenzellkarzinom einen erhöhten Output des tumorauslösenden Signalwegs HIF2A (Hypoxie-induzierbarer Faktor 2 Alpha), was die Metastasierung fördert (90).

DNA-Methylierung kann auch zur Stilllegung von Metastasensuppressorgenen führen (68). So wird z.B. das Tumorsuppressorgen *PAQR3* (*Progesterin And AdipoQ Receptor Family Member 3*) durch Hypermethylierung herunterreguliert, was wiederum mit Lymphknotenmetastasen assoziiert ist (101). Die Dysregulation von DNA-Methylierung ist auch mit hämatologischen Krebserkrankungen assoziiert. Bei

der akuten myeloischen Leukämie (AML) lassen sich anhand der DNA-Methylierungsprofile unterschiedliche Subgruppen unterscheiden (102). Mutationen in DNMT3A (tritt bei 25% der AML-Patienten auf) (103), in TET2 (ca. 17% der AML-Patienten) (104) und im mutierten *IDH1/2*-Gen (Isocitratdehydrogenase 1/2) beeinflussen die DNA-Methylierung (105). Die Mutation von DNMT3A führt bei den AML-Subtypen M4/M5 *in vitro* zu vermehrter Aggressivität mit verstärkter extramedullärer Infiltration, was teilweise durch die Hochregulation von TWIST1 (*Twist family BHLH transcription factor 1*) geschieht (106).

### 1.5.2.2 Histonmodifikationen

Neben dem Einfluss von DNA-Methylierung auf die maligne Progression von Krebs ist es wahrscheinlich, dass auch Histonmodifikationen zur Metastasierung beitragen. Dabei spielt zum Beispiel PRC2 eine wichtige Rolle, da dieser Komplex die Repression von Genen reguliert. Über seine katalytische Untereinheit EZH1/EZH2 methyliert er H3K27 dreifach, was zur Unterdrückung der Transkription führt. Dies geschieht u.a. an Promotoren, die sich bereits in einem inaktiven Zustand befinden, der dadurch stabilisiert wird (90).

PRC2 kann die Tumorprogression, je nach Kontext, sowohl fördern als auch inhibieren. Die Aktivität von EZH2 ist einerseits entscheidend für die EMT (epithelial-mesenchymale Transition) bei Brustkrebs (107,108) und für die maligne Progression bei malignem Melanom (109). Jedoch kann auch eine verminderte Aktivität von PRC2 die Metastasierung unterstützen, indem sie zu epigenetischer Instabilität und damit zur Aktivierung metastasenfördernder Gene, wie z.B. *CXCR4* (CXC-Motiv-Chemokinrezeptor 4) führt (90).

## 1.6 Theorie der *early dissemination*

Obwohl es viele verschiedene Krebsarten gibt, wird angenommen, dass es bestimmte Eigenschaften gibt, die alle miteinander gemein haben. Hanahan und Weinberg nannten diese Eigenschaften die „*Hallmarks of Cancer*“ (110,111). Dabei beschreiben sie *Invasion und Metastasierung* als eines der entscheidenden Kennzeichen von malignen Tumoren. Diese Eigenschaft wird während der Tumorentstehung durch die Anhäufung von genetischen und epigenetischen Aberrationen erworben (112). Manifeste Metastasen sind meist unheilbar. Um diese

manifeste Metastasierung zu verhindern, müssten disseminierte Tumorzellen (*disseminated cancer cells*; DCCs) so früh wie möglich therapeutisch behandelt werden (84). Dabei stellt sich die Frage, zu welchem Zeitpunkt der Tumor streut und ob sich die DCCs von den Tumorzellen des Primärtumors unterscheiden. Das Auftreten von Metastasen Jahre nach Entfernung des Primärtumors lässt vermuten, dass die Aussaat von DCCs relativ früh stattfindet.

Erste Belege dafür, dass diese früh disseminierten Tumorzellen eine charakteristische Biologie aufweisen, zeigten Analysen von DCCs aus dem Knochenmark von Patienten mit Brustkrebs vor und nach der Manifestation von Metastasen (M0 und M1 Stadium, entsprechend der Leitlinien der *Union for International Cancer Control*) (113,114). Aus der Entdeckung, dass Primärtumor und DCCs sich biologisch gesehen unterscheiden, entstand die Theorie der *early dissemination*. Diese nimmt an, dass DCCs bereits im M0-Stadium disseminieren und sich parallel zum Primärtumor entwickeln (115). Studien mit transgenen Mausmodellen (116–118) und mit Patienten mit Präkanzerosen oder *in situ* Karzinomen (116,119,120) haben weitere Hinweise für die Richtigkeit dieser Hypothese geliefert, doch die Relevanz von DCCs wird immer noch angezweifelt (84,94).

Man weiß also bereits, dass bei Brustkrebs, ähnlich wie bei anderen Krebsarten, DCCs vorkommen (121). Jedoch haben nicht alle disseminierten Zellen die Fähigkeit, sich in fremdem Gewebe festzusetzen. Es gilt deshalb zu bedenken, dass es neben der Fähigkeit zur Invasion weiterer Eigenschaften bedarf, um zu metastasieren. Neu erworbene genetische und epigenetische Veränderungen scheinen eine Rolle bei der Einnistung in und Anpassung an eine neue Nische zu spielen (112).

### **1.7 Zielsetzung der Arbeit und Fragestellung**

Die Bedeutung veränderter DNA-Methylierung für die Entstehung und Progression von Brustkrebs ist bis heute nicht vollständig geklärt. Große Studien konzentrieren sich auf Chromosomenaberrationen, Mutationen und Transkriptionsprofile, doch wurde das Epigenom noch nicht umfassend auf einem genomweiten Level untersucht. Vor kurzem legten Studien nahe, dass DNA-Methylierungsmuster mit Krebsentstehung und Progression korrelieren (122).

Ziel dieser Arbeit war es daher, den Beitrag epigenetischer Mechanismen zur malignen Progression von Brustkrebs in einem Mausmodell systematisch und umfassend zu untersuchen. Dabei lag ein besonderer Fokus auf der Frage, ob die Metastasierung mit spezifischen epigenetischen Veränderungen im Vergleich zu Primärtumoren korreliert und ob auf epigenetischer Ebene weitere Hinweise für die Theorie der *early dissemination* detektierbar sind.

## 2 Material

### 2.1 Versuchstiere

Alle BALB/c-Versuchstiere wurden, soweit nicht anders angegeben, von Charles River Laboratories erworben. Die BalbNeuT-Mäuse stammten von Prof. Guido Forni, Università di Torino. Alle Experimente wurden nach den Standards der „*good laboratory practise*“ sowie den Richtlinien der Europäischen Union und der Universität Regensburg durchgeführt und die tierexperimentellen Untersuchungen wurden von der Regierung der Oberpfalz unter dem Aktenzeichen 55.2-2532.1-27/14 genehmigt. Alle Experimente wurden unter Aufsicht eines zertifizierten Tierarztes durchgeführt.

### 2.2 Geräte

|                                    |  |
|------------------------------------|--|
| Heizblock Thermomixer compact      | Eppendorf, Hamburg, Deutschland              |
| Heizblock                          | BIOER, Hangzhou, China                       |
| Magnetic Particle Concentrator     | Thermo Fisher Scientific, Waltham, USA       |
| Magnetrührer                       | IKA, Staufen, Deutschland                    |
| Mikroskope                         | Carl Zeiss, Jena, Deutschland                |
| Mikrowelle                         | Privileg, Hamburg, Deutschland               |
| NanoDrop                           | PeqLab, Erlangen, Deutschland                |
| PCR-Thermocycler PTC-200           | MJ-Research/Biometra, Oldendorf, Deutschland |
| Rotationsgerät Intelli-Mixer RM-L2 | ELMI Ltd., Riga, Litauen                     |
| Rotationsgerät                     | Labinco B.V., Breda, Niederlande             |
| Schüttelgerät Vortexgenie 2        | Scientific Industries, New York, USA         |
| Schüttelgerät Vortexer MS3         | IKA, Staufen, Deutschland                    |
| Schüttelgerät REAX 2000            | Heidolph Instruments, Schwabach, Deutschland |
| Spectrophotometer                  | Perkin Elmer, Überlingen, Deutschland        |
| Sterilbank                         | Heraeus, Hanau, Deutschland                  |



|                              |  |
|------------------------------|--|
| TapeStation                  | Agilent Technologies, Waldbronn, Deutschland |
| Thermomixer                  | Eppendorf, Hamburg, Deutschland              |
| TissueLyser                  | Qiagen, Hilden, Deutschland                  |
| Typhoon 9200                 | Molecular Dynamics, Krefeld, Deutschland     |
| Ultraschall-Homogenisator    | Covaris, Woburn, USA                         |
| Vakuumkonzentrator Speed Vac | Christ, Osterode, Deutschland                |
| Vakuumpumpe                  | KNF Neuberger, Freiburg, Deutschland         |
| Wasseraufbereitungsanlage    | Millipore, Eschborn, Deutschland             |
| <u>Zentrifugen:</u>          |  |
| Centrifuge 5804R             | Eppendorf, Hamburg, Deutschland              |
| Heraeus Biofuge fresco       | Thermo Fisher Scientific, Waltham, USA       |
| Heraeus Biofuge pico         | Thermo Fisher Scientific, Waltham, USA       |
| Capsulefuge PMC-060          | TOMY, Tokyo, Japan                           |
| Mini-Zentrifuge              | Carl Roth, Karlsruhe, Deutschland            |

### 2.3 Verbrauchsmaterialien

|                                       |  |
|---------------------------------------|--|
| Falcon Tubes 50 ml                    | Greiner Bio-One, Frickenhausen, Deutschland  |
| Microtube AFA Fiber Pre-Slit Snap-Cap | Covaris, Woburn, USA                         |
| Optical Tube mit Deckel               | Agilent Technologies, Waldbronn, Deutschland |
| Pipettenspitzen 50 – 1.000 µl blau    | VWR, Darmstadt, Deutschland                  |
| Reaktionsgefäß DNA Low Binding 0,5 ml | Sarstedt, Nümbrecht, Deutschland             |
| Reaktionsgefäß DNA Low Binding 1,5 ml | Sarstedt, Nümbrecht, Deutschland             |
| Reaktionsgefäß 2 ml                   | Sarstedt, Nümbrecht, Deutschland             |
| Reaktionsgefäße, Safe-Lock, 1,5ml     | VWR, Darmstadt, Deutschland                  |
| Serologische Pipetten 5 ml            | Sarstedt, Nümbrecht, Deutschland             |
| Serologische Pipetten 10 ml           | Sarstedt, Nümbrecht, Deutschland             |

## 2.4 Chemikalien

Alle Chemikalien wurden, soweit nicht anders angegeben, von Sigma, Deisendorf oder Merk, Darmstadt bezogen.

## 2.5 Enzyme, Nukleotide und molekularbiologische Kits

|  |   |
|--|---|
| DNeasy Blood & Tissue Kit                      | Qiagen, Hilden, Deutschland   |
| DNA Ladder 1 kb plus                           | Thermo Fisher Scientific, Waltham, MA, USA                          |
| dATPs  | NEB, Frankfurt, Deutschland   |
| dNTPs  | Roche, Basel, Schweiz   |
| EpiMark® Methylated DNA Enrichment Kit         | NEB, Frankfurt, Deutschland   |
| Ethidium bromid                                | Sigma-Aldrich, München, Deutschland                                 |
| Fermentas DNA loading dye (6x)                 | Thermo Fisher Scientific, Waltham, MA, USA                          |
| Glycogen                                       | Ambion/Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA |
| Klenow (3'→5' exo)                             | Enzymatics, Beverly, USA  |
| Klenow Fragment                                | Enzymatics, Beverly, USA  |
| LabChip XT DNA Chips                           | Perkin Elmer, Waltham, MA, USA                                      |
| MinElute PCR Purification Kit                  | Qiagen, Hilden, Deutschland   |
| NEXTflex™ DNA Barcodes                         | Bioo Scientific, Austin, TX, USA                                    |
| Nuclease-freies Wasser                         | Gibco/Thermo Fisher Scientific, Waltham, MA, USA                    |
| Phusion Hot Start High-Fidelity DNA Polymerase | Thermo Fisher Scientific, Waltham, MA, USA                          |
| PMSF (Phenylmethylsulfonylfluorid)             | Sigma-Aldrich, München, Deutschland                                 |
| Proteinase K                                   | Roche, Basel, Schweiz   |
| QIAquick PCR Purification Kit                  | Qiagen, Hilden, Deutschland   |
| QIAamp DNA Micro Kit                           | Qiagen, Hilden, Deutschland   |
| S-adenosylmethionine (SAM)                     | NEB, Frankfurt, Deutschland   |
| SssI CpG Methyltransferase                     | NEB, Frankfurt, Deutschland   |

|  |                          |
|--|--------------------------|
| T <sub>4</sub> DNA Ligase                | Enzymatics, Beverly, USA |
| T <sub>4</sub> DNA Polymerase            | Enzymatics, Beverly, USA |
| T <sub>4</sub> DNA Polynucleotide Kinase | Enzymatics, Beverly, USA |

## 2.6 Größenstandards

Die 1 kB Plus DNA Ladder wurde von Thermo Fisher Scientific, Waltham, USA bezogen.

## 2.7 Datenbanken und Software

Zur bioinformatischen Aufbereitung der Rohdaten waren verschiedene Programme nötig. Ein Großteil der Datenaufbereitung und -verarbeitung wurde mit Hilfe des Statistikprogramms *R statistics* (Version 3.6.1; <https://cran.r-project.org/>) durchgeführt. Für einige speziellere Fragestellungen wurden jedoch auch weitere Programme genutzt: *Bowtie2* (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) (123) und *samtools* (<http://samtools.sourceforge.net/>) (124) wurden zum Anpassen und Speichern der Sequenzen verwendet. *HOMER* Next Generation Sequencing Analysis (<http://homer.ucsd.edu/homer/ngs/>) (125) ist eine Software zum Zusammenstellen und Analysieren von Sequenzierungsdaten. Differentielle Methylierung zwischen Primärtumoren und Metastasen wurde mittels *DESeq2* analysiert (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>) (126). Analysen zur Funktion der relevanten Gene wurden mittels *GREAT* (Version 4.0.4, <http://great.stanford.edu/public/html/>) (127) und *Metascape* (<http://metascape.org/gp/index.html>) (128) durchgeführt.

## **3 Methoden**

### **3.1 Arbeiten mit Mäusen**

Alle Arbeiten an Mäusen wurden am Lehrstuhl für Experimentelle Medizin und Therapieverfahren durchgeführt.

#### **3.1.1 Züchtung, Transplantation und Operation der Mäuse**

Die Züchtung, Transplantation und Operation der Mäuse wurden in der Zeit vom März 2011 bis Dezember 2012 von der Arbeitsgruppe von Prof. Klein durchgeführt, die das zentrale Mausprojekt der Forschergruppe 2127 koordiniert. In dieser Zeit wurden 174 Mäuse, die für diese Arbeit relevant sind, mit einer noch nicht entarteten Brustdrüse einer drei bis fünf Wochen alten BalbNeuT-Maus in das ausgeräumte Fettpolster der vierten Brustdrüse einer Maus des Wildtyps BALB/c transplantiert. Der daraufhin in der vierten Brustdrüse entstehende Primärtumor wurde in einer Operation reseziert, wenn er eine Größe von 5-10 mm erreicht hatte. Dies war in der Regel nach 25 (+/- 8) Wochen der Fall. Der Tumor wurde als Kryomaterial gelagert.

#### **3.1.2 Tötung und Dissektion der Mäuse**

Nach der Operation wurden diese Mäuse nach weiteren 17 (+/- 5; im Median 16,5) Wochen getötet. Bei der Dissektion wurden Lunge, Leber, Milz und Knochenmark entnommen und als Kryomaterial gelagert. Dies geschah zwischen März 2011 und Dezember 2012.

Um Material von nicht transplantierten BALB/c-Mäusen als Kontrolle zu erhalten, wurden im März und April 2015 Gewebeentnahmen von je einer Wildtyp BALB/c-Maus durchgeführt. Die Mäuse wurden dafür durch Genickbruch getötet. Anschließend wurde die vierte Brustdrüse beidseits entnommen, um nicht tumoröses Kontrollgewebe zu gewinnen. Das gesunde Lungengewebe wurde im Rahmen der Vorarbeiten der Forschergruppe von Prof. Christoph Klein gewonnen.

## 3.2 Verarbeitung der Gewebeproben

### 3.2.1 Verarbeitung von Tumoren und Metastasen

Die Tumoren und Metastasen aus den Jahren 2011/2012 wurden nach der Sektion in flüssigem Stickstoff schockgefroren und bei -80 °C aufbewahrt. Die Proben von Tumoren und Metastasen wurden von der Arbeitsgruppe von Prof. Christoph Klein bis zur Extraktion der genomischen DNA (gDNA) weiterverarbeitet.

### 3.2.2 Gewebeverdau des Kontrollgewebes

Das Kontrollgewebe der vierten Brustdrüse beider BALB/c-Mäuse wurde direkt nach der Entnahme von mir verdaut, um eine Einzelzellsuspension zu erhalten. In Tabelle 3-1 und 3-2 sind die Zusammensetzungen von Basalmedium und Verdaumedium dargestellt.

**Tabelle 3-1 Basalmedium für den Verdau von Brustdrüsengewebe**

| <i>Komponente</i>        | <i>Volumen</i>         |
|--------------------------|------------------------|
| DMEM                     | 250 ml                 |
| Ham's F12                | 250 ml                 |
| Hepes (1M)               | 5ml (10 nM)            |
| Pen/Strep                | 5 ml                   |
| BSA                      | 1 % (5 g)              |
| Insulin                  | 500 µl (10 µg/ml)      |
| <b><i>Endvolumen</i></b> | <b><i>510,5 ml</i></b> |

**Tabelle 3-2: Verdaumedium für den Verdau von Brustdrüsengewebe**

| <i>Komponente</i>        | <i>Volumen</i>        | <i>Endkonzentration</i> |
|--------------------------|-----------------------|-------------------------|
| Basalmedium              | 10 ml                 |                         |
| Kollagenase              | 100 µl                | 0,33 mg/ml              |
| Hyaluronidase            | 100 µl                | 100 µg/ml               |
| DNase I                  | 400 µl                | 0,1 mg/ml               |
| <b><i>Endvolumen</i></b> | <b><i>10,6 ml</i></b> |                         |

In eine Petrischale mit 10 cm Durchmesser wurden 0,5-1 ml des Verdaumediums gegeben und die Gewebeprobe hinzugefügt. Mit zwei Skalpellen wurde das Gewebe zerkleinert und mit Zugabe des restlichen Verdaumediums auf- und abpipettiert. Daraufhin wurde das Gewebe 20-40 Minuten bei 37°C inkubiert. Nach

10 und nach 20 Minuten der Inkubationszeit erfolgte jeweils erneutes Pipettieren, um die Vereinzelnung der Zellen zu gewährleisten. In einem 50 ml-Proberöhrchen wurde die Zellsuspension und eine äquivalente Menge PBS (*phosphate-buffered saline*), mit der die Petrischale ausgespült wurde, abzentrifugiert (300xg, Raumtemperatur, 10 min). Der Überstand wurde verworfen und das Pellet durch leichtes Schütteln mit 1 ml Trypsin/EDTA (37°C) resuspendiert und 5 Minuten bei Raumtemperatur inkubiert. Durch Zugabe von 5 ml Basalmedium mit 10% FCS (*fetal calf serum*) wurde der Verdau inaktiviert und die erhaltene Lösung wurde zehnmal auf- und abpipettiert. Die Zellsuspension wurde nun durch ein 100 µm-Zellsieb in ein neues 50 ml-Röhrchen filtriert. Die auf dem Filter verbleibenden Zellen wurden mit dem Stempel einer 5 ml-Spritze vorsichtig zerrieben. Das Röhrchen wurde mit 10 ml PBS ausgespült und diese Flüssigkeit wurde ebenfalls durch das Zellsieb pipettiert. Die Zellsuspension wurde für 5-8 Minuten bei Raumtemperatur und 300xg zentrifugiert. Bis auf einen Rest von ca. 1 ml wurde der Überstand verworfen, das Pellet in der restlichen Flüssigkeit resuspendiert und das Volumen genau bestimmt. Die erhaltene Zellsuspension wurde für die Isolation von genomischer DNA (s. 3.3.1), ChIP (s. 3.3.6) und ATAC (s. 3.3.8) verwendet.

### **3.3 Allgemeine molekularbiologische Methoden**

#### **3.3.1 Isolation von genomischer DNA aus Brustdrüsengewebe**

Genomische DNA wurde nach Anweisung des Animal Tissue Spin-Column Protokolls (DNeasy Blood & Tissue Kit) gewonnen. Zunächst wurde das Gewebe in ATL-Puffer gegeben und mit dem TissueLyser (Qiagen) für 1x20s bei 15 Hz zerkleinert. Nach Hinzufügen von Proteinase K wurde das Gewebe über Nacht bei 56°C und 1050 rpm im Thermomixer inkubiert.

#### **3.3.2 Isolation von genomischer DNA aus Zellen**

Genomische DNA wurde nach Anweisung des Animal Blood or Cell Spin-Column Protokolls (DNeasy Blood & Tissue Kit) hergestellt.

#### **3.3.3 Quantifizierung der DNA**

Die genaue DNA-Konzentration wurde mithilfe des Nanodrop ND 1000 (PeqLab) Spektrophotometers bestimmt.

### 3.3.4 *In vitro* Methylierung der DNA

Genomische DNA, gewonnen aus dem Kontrollgewebe der dritten Brustdrüse einer BALB/c-Maus, wurde komplett methyliert, um die maximal mögliche Methylierung des Mausgenoms zu erfassen. Die Methylierung erfolgte durch die CpG-Methyltransferase *M. SssI* (New England Biolabs®), die alle Cytosinreste (C<sup>5</sup>) der doppelsträngigen Dinukleotid-Erkennungssequenz 5'...CG...3' methyliert (94). S-Adenosylmethionin (SAM) (New England Biolabs®) dient dabei als Methylgruppendorator. Hierfür wurden 5-6 µg gDNA für zwei Stunden bei 37°C mit der CpG-Methyltransferase inkubiert. Dabei wurde die in Tabelle 3-3 dargestellte Reaktionszusammensetzung genutzt.

**Tabelle 3-3 Reaktionszusammensetzung für *in vitro* DNA-Methylierung**

| <i>Komponente</i>                  | <i>Volumen</i>       |
|------------------------------------|----------------------|
| DNA                                | 5-6 µg               |
| NEBuffer 2                         | 10 µl                |
| SAM                                | 0,5 µl               |
| <i>SssI</i>                        | 5 µl                 |
| H <sub>2</sub> O <sub>bidest</sub> | ad 100 µl            |
| <b><i>Endvolumen</i></b>           | <b><i>100 µl</i></b> |

Nach zwei Stunden wurden erneut 0,75 µl SAM und 1 µl *SssI* hinzugegeben und der Ansatz weitere zwei Stunden bei 37°C inkubiert. Anschließend wurde *SssI* für 20 Minuten bei 65°C inaktiviert. Die Aufreinigung erfolgte nach Angaben des Cleanup of Genomic DNA-Protokolls (QIAamp DNA Micro Kit, Qiagen).

### 3.3.5 Methyl-CpG-Immunopräzipitation (MCIp)

Die Methyl-CpG-Immunopräzipitation dient dazu, methylierte DNA aus fragmentierter genomischer DNA zu isolieren. Dazu werden hydrophile magnetische Protein A-Beads verwendet, auf deren Oberfläche ein Komplex aus dem Fc-Fragment des humanen IgG<sub>1</sub> und dem humanen MBD2-Protein (methyl-CpG binding domain 2 Protein) gebunden ist. Da das Fc-Fragment ein Dimer mit zwei Bindungsstellen ist und je Protein A-Molekül zwei Fc-Fragmente binden, werden vier MBD2-Domänen je Protein A-Molekül exponiert. MBD2 gehört zu einer Familie von Kernproteinen mit methyl-CpG binding domain (MBD) und bindet spezifisch methylierte DNA (95).

Zunächst wurde die Integrität der genomischen DNA mittels TapeStation (Agilent) überprüft. Anschließend wurde die gDNA mit dem Ultraschall-Homogenisator (Covaris®) fragmentiert (s. 3.3.5.1). Die Methyl-CpG-Immunopräzipitation wurde mit dem EpiMark® Methylated DNA Enrichment Kit (New England Biolabs®) nach Anweisung des Herstellers durchgeführt.

### 3.3.5.1 Fragmentierung der genomischen DNA

Die genomische DNA wurde mittels Sonifikation am Ultraschall-Homogenisator (Covaris®) nach Angaben des Herstellers fragmentiert (Tab. 3-4). Dabei wird Schallenergie in einem Brennpunkt gebündelt und es entstehen dort unzählige Bläschen. Nachdem keine weitere Schallenergie mehr zugeführt wird, fallen die Bläschen in sich zusammen und erzeugen lokal eine schnelle Bewegung der gelösten DNA. Die entstandenen Fragmente haben jeweils eine Größe von ca. 300 Basenpaaren und wurden mittels TapeStation® kontrolliert. Es wurden 200 ng genomische DNA in 130 µl gibco-H<sub>2</sub>O mit nachfolgenden Einstellungen in microTUBEs (Covaris®) sonifiziert.

**Tabelle 3-4 Parameter für die Sonifikation der DNA mit dem Ultraschall-Homogenisator (Covaris®)**

| Zielbereich der Basenpaare (Peak) | 300 bp |
|-----------------------------------|--------|
| Duty Cycle                        | 10%    |
| Intensity                         | 5      |
| Cycles per burst                  | 200    |
| Frequency sweeping                | 180 s  |
| Temperatur                        | 7°C    |
| Wasserlevel                       | 12     |

### 3.3.5.2 Binden von MDB2-Fc an die Protein A-Magnetbeads

Um die Protein A-Magnetbeads zu resuspendieren, wurden sie 30 Minuten bei Raumtemperatur rotiert. Nach Hinzufügen von 1 µl MBD2-Fc und 11 µl 1X Bind/Wash Buffer zu 10 µl Magnetbeads, wurde das Gemisch weitere 15 Minuten bei Raumtemperatur rotiert und danach zwei Mal mit 1.000 µl 1X Bind/Wash Buffer gewaschen.



### 3.3.5.3 Hinzufügen fragmentierter DNA zu MDB2-Fc/Protein A-Magnetbeads

Die fragmentierte DNA wurde nach folgendem Ansatz an MDB2-Fc/Protein A-Magnetbeads (New England Biolabs®) gebunden, indem sie eine Stunde im Rotor bei Raumtemperatur inkubiert wurden (Tab. 3-5).

**Tabelle 3-5 Reaktionszusammensetzung für die Bindung von fragmentierter DNA an MBD2-Fc-Magnetbeads**

| <i>Komponente</i>                  | <i>Volumen</i>       |
|------------------------------------|----------------------|
| 5 x Bind/Wash Buffer               | 20 µl                |
| Fragmentierte DNA                  | 70 µl (≈ 200 ng DNA) |
| MBD2-Fc/Protein A-Magnetbeads      | 10 µl                |
| H <sub>2</sub> O <sub>bidest</sub> | ad 100 µl            |
| <b>Endvolumen</b>                  | <b>100 µl</b>        |

Es folgen drei Waschschrte mit 1.000 µl 500 mM-Bind/Wash Buffer, um die ungebundene DNA zu entfernen.

### 3.3.5.4 Elution angereicherter, stark methylierter DNA

Durch die Zugabe von 40 µl DNase-freiem H<sub>2</sub>O und 15-minütiger Inkubation im Thermomixer bei 65°C und 400 Umdrehungen pro Minute (*rounds per minute*; rpm) wird die gebundene DNA eluiert. Das Eluent enthält die angereicherte, methyl-CpG enthaltende DNA und wurde bei -20°C aufbewahrt.

### 3.3.6 Chromatin-Immunopräzipitation (ChIP)

Mit der Chromatin-Immunopräzipitation (ChIP) lassen sich Protein-DNA-Interaktionen untersuchen. Zunächst werden durch Formaldehyd kovalente Quervernetzungen zwischen der DNA und den Proteinen (Chromatin) hergestellt. Anschließend wird die quervernetzte DNA-Protein-Verbindung durch Sonifizieren zerkleinert und spezifische Antikörper werden an die zu untersuchenden Proteine gebunden. Die DNA-Protein-Verbindungen mit gebundenem Antikörper werden gefällt. Danach werden die DNA-Protein-Verbindungen aufgelöst und die Basensequenz der hierbei erhaltenen DNA wird untersucht. (96) Für die Chromatin-Immunopräzipitation werden Einzelzellen benötigt, hierfür wurde die aus dem Gewebeverdau der gesunden Brustdrüse erhaltene Zellsuspension verwendet. Die Zellen wurden bei 1.300 rpm (300xg) für 10 Minuten bei 4°C zentrifugiert und anschließend in 37°C warmem RPMI-Kulturmedium inklusive 10% FCS

resuspendiert, so dass man eine Zellkonzentration von 1 Millionen Zellen je Milliliter erhält.

### **3.3.6.1 Quervernetzung mit Formaldehyd**

Die Quervernetzung zwischen DNA und Chromatin wurde durch Zugabe von 1%igem Formaldehyd zur Zellsuspension erreicht. Während einer 10-minütigen Inkubationszeit bei RT wurde das Proberöhrchen mehrmals auf den Kopf gestellt, um ein Absetzen der Zellen zu vermeiden. Die Reaktion wurde durch die Beigabe von Glycin (2,625 M; 1/20 des aktuellen Volumens der Suspension) gestoppt und 10 Minuten bei 4°C und mit 1.300 rpm (300xg) zentrifugiert. Der Überstand wurde verworfen und das Pellet vorsichtig resuspendiert. Das Pellet wurde zwei Mal mit 25 ml eiskaltem PBS/PMSF/NaBut gewaschen, jeweils mit 300xg bei 4°C für 8 Minuten zentrifugiert und der Überstand verworfen. Zuletzt wurde das Pellet wieder vorsichtig resuspendiert. Schockfrieren des Rückstands in flüssigem Stickstoff und Aufbewahrung bei -80°C ist zu diesem Zeitpunkt möglich.

### **3.3.6.2 Zell- und Nukleuslyse**

Um das quervernetzte Chromatin zu gewinnen, wurden die fixierten Zellen und Zellkerne lysiert. Hierfür wurden sie in 250 µl Suspensionspuffer pro 10 Millionen Zellen suspendiert und durch Zugabe von 250 µl Zellysepuffer pro 10 Millionen Zellen 10 Minuten auf Eis lysiert. Nach einer Zentrifugation von 5 Minuten bei 700xg und 4°C, wurde der Überstand verworfen und der Rückstand mit 400 µl Nukleuslysepuffer pro 10 Millionen Zellen resuspendiert.

### **3.3.6.3 Fragmentierung des Chromatins**

Das im vorhergehenden Schritt zugänglich gemachte Chromatin wurde nun durch Sonifizierung im Ultraschall-Homogenisator (Covaris®) fragmentiert. Je nach Volumen wurden entweder Chromatin Proberöhrchen (Volumen 300 µl bis 3 ml) oder Mikrotubes (Volumen bis 130 µl) und unterschiedliche Einstellungen für die Sonifizierung verwendet (s. Tabellen 3-6 und 3-7).

**Tabelle 3-6** Parameter für die Sonifikation des Chromatins mit dem Ultraschall-Homogenisator (Covaris®) für ein Volumen von 300 µl bis 3 ml

| <i>Parameter</i>   | <i>Einstellungen</i> |
|--------------------|----------------------|
| Fragmentgröße      | 200 bp               |
| Duty Cycle         | 20 %                 |
| Intensity          | 8                    |
| Cycles per burst   | 200                  |
| Frequency sweeping | 2.400 s              |
| Wassertemperatur   | 7°C                  |
| Wasserlevel        | 15                   |

**Tabelle 3-7** Parameter für die Sonifikation des Chromatins mit dem Ultraschall-Homogenisator (Covaris®) für ein Volumen bis 130 µl

| <i>Parameter</i>   | <i>Einstellungen</i> |
|--------------------|----------------------|
| Fragmentgröße      | 200 bp               |
| Duty Cycle         | 2%                   |
| Intensity          | 5                    |
| Cycles per burst   | 100                  |
| Frequency sweeping | 480 s                |
| Wassertemperatur   | 7°C                  |
| Wasserlevel        | 12                   |

Um das Lysat von den Zelltrümmern zu trennen, wurde 5 Minuten bei 4°C und mit 13.000 rpm zentrifugiert und der Überstand in ein neues Proberöhrchen überführt. Das Lysat wurde anschließend mit Zellkernlysepuffer vermischt und davon 20 µl mit 2 µl Proteinase K (20 µg/µl) bei 65°C über Nacht inkubiert, um die Quervernetzung zwischen DNA und Chromatin rückgängig zu machen. Am nächsten Tag wurde 1 µl RNase (10 µg/µl) hinzugegeben und weitere 30 Minuten bei 37°C inkubiert. Die Probe wurde mit 13.000 rpm für 5 Minuten bei 4°C zentrifugiert und der Überstand in ein neues Proberöhrchen pipettiert. Anschließend wurde die Probe mit dem QiaQuick PCR DNA Purification Kit® (Qiagen, Hilden, Deutschland) nach Anleitung des Herstellers aufgereinigt und mit 30 µl Elutionspuffer eluiert. Mittels Agarose-Gelelektrophorese (2%iges Agarosegel, 50 bp-Leiter) wurde die Größe der

fragmentierten DNA bestimmt. Es wurden 700 ng DNA auf ein 2%iges Agarosegel aufgetragen und gemeinsam mit 3 µl einer 50 bp DNA Leiter 30 Minuten bei 100 V aufgetrennt.

#### 3.3.6.4 Immunopräzipitation

Für jede Immunopräzipitation (IP) wurden 80 µl des gereinigten Lysats eingesetzt. Jedes Aliquot eines Chromatinansatzes wurde zunächst einer Vorreinigung (*Pre-clearing*) unterzogen, um alle Proteine zu eliminieren, die unspezifisch an Komponenten der IP binden. Dazu wurden zunächst 50 µl Sepharose CL-4B Beads drei Mal pro IP mit TE Puffer (pH 8,0) gewaschen und anschließend jeweils eine Minute bei 3.000 rpm zentrifugiert, um das Ethanol zu entfernen. Das Proberöhrchen mit den Beads wurde dann bis zum Erreichen des ursprünglichen Volumens vor dem Waschen mit Dilutionspuffer (DB, ohne Inhibitoren) aufgefüllt. Es wurden jeweils 25 µl Rinderalbumin (BSA, 20%) und 4 µl Glykogen pro 1 ml Beads hinzugegeben und das Gemisch wurde zwei Stunden bei Raumtemperatur rotiert. Zum Pre-clearing wurden nun jeweils 80 µl Lysat (entspricht 2 Millionen Zellen) mit dem 1,5-fachem Volumen (120 µl) an DB inklusive Inhibitoren und 50 µl inkubierten Beads je IP vermischt und bei 4°C für zwei Stunden mit 6 rpm rotiert. Das vorgereinigte Lysat wurde mit 13.000 rpm für 5 Minuten bei 4°C zentrifugiert. Die 200 µl Überstand pro IP wurden in neue DNA-Low Binding-Proberöhrchen überführt, die Antikörper (2,5 µg) wurden hinzugefügt und über Nacht bei 4°C und mit 6 rpm Rotation inkubiert. Je IP wurden außerdem 50 µl Protein A Sepharose Beads benötigt, die ebenfalls zuerst drei Mal mit TE-Puffer gewaschen, jeweils mit 3.000 rpm für eine Minute abzentrifugiert und anschließend mit DB inklusive Inhibitoren auf das ursprüngliche Volumen aufgefüllt wurde. Es wurde 0,4 µl Glykogen pro 100 µl Beads und 2,5 µl BSA (20%) pro 100 µl Beads hinzugegeben und bei 4°C und 6 rpm über Nacht rotiert.

Nach dieser Inkubation wurden je IP 50 µl der Protein A Sepharose Beads zur Antikörper-DNA-Mischung hinzugefügt und für zwei Stunden mit 6 rpm (4°C) rotiert, um die Beads an die Antikörper-DNA-Komplexe zu binden. Anschließend wurden die Proben bei 2.800 rpm (500xg) und 4°C für 5 Minuten zentrifugiert und der Überstand verworfen. Es folgten jeweils mehrere Waschschrte, um unspezifisch gebundenes Chromatin zu entfernen. Die Bead-Komplexe wurden mit je 400 µl der drei Waschpuffer (WB I-III) eluiert, jeweils 5 Minuten bei Raumtemperatur und 6 rpm

rotiert, für 2 Minuten bei 4°C mit 500g abzentrifugiert. Der Überstand wurde verworfen. Im letzten Waschschrift wurden die Proben drei Mal mit 400 µl TE-Puffer (1x) resuspendiert und analog zu den vorangegangenen Waschschriften abzentrifugiert. Um an Beads und Antikörper gebundene DNA-Komplexe zu eluieren, wurden die Proben mit jeweils 110 µl Elutionspuffer (EB) für 20 Minuten inkubiert und dabei alle 5 Minuten geschüttelt oder durchgehend bei 1 rpm rotiert. Nach einer zweiminütigen Zentrifugation mit 500xg bei Raumtemperatur wurden 100 µl des Überstands in ein neues 1,5 ml-Proberöhrchen überführt. Die Beads wurden erneut mit 100 µl EB für 10 Minuten nach demselben Muster inkubiert und abzentrifugiert. 100 µl des Überstandes wurden zum initialen 100 µl Eluat im neuen 1,5 ml-Proberöhrchen hinzugefügt. Um die Quervernetzung zwischen DNA und Chromatin rückgängig zu machen, wurde nun 10 µl NaCl (5M) zu jeder Probe hinzugefügt und bei 65°C über Nacht inkubiert.

### **3.3.6.5 Aufreinigung des Chromatins**

Zur Aufreinigung des Chromatins wurden Proben und Input mit 7 µl RNase (0,33 µg/µl) für eine Stunde bei 37°C inkubiert. Dann wurde 5 µl Proteinase K (20 µg/µl) hinzugefügt und für eine weitere Stunde bei 37°C inkubiert. Die weitere Aufreinigung erfolgte mit dem QiaQuick PCR Purification Kit® (Qiagen, Hilden, Deutschland). Die Proben wurden mit dem fünffachen Volumen (1 ml) PB-Puffer (ohne pH-Indikator) für 30 Minuten bei Raumtemperatur inkubiert. Die DNA wurde an die Säulen gebunden, indem zwei Mal 750 µl der Probe auf die Säule gegeben und je 30 Sekunden bei 11.000 rpm zentrifugiert wurde. Dann wurden 750 µl Waschpuffer PE auf die Säule gegeben und 30 Sekunden bei 13.000 rpm zentrifugiert. Die Säule wurde mit einem neuen Auffanggefäß für 2 Minuten bei 13.000 rpm trocken zentrifugiert und mit 53 µl auf 50°C erwärmten EB-Puffer eluiert. Nach der Zugabe von EB wurde für zwei bis fünf Minuten bei Raumtemperatur inkubiert und für eine Minute bei 11.000 rpm abzentrifugiert. 40 µl des erhaltenen Eluats wurden für die Library Präparation (s. 3.3.7) verwendet, um anschließend mittels Next Generation Sequencing analysiert zu werden.

*Benötigte Reagenzien und Materialien:*

(genannte Volumen-/Gewichts-% oder Stoffmengenkonzentrationen beziehen sich auf die Endkonzentration, gelagert wurden die Reagenzien laut Anweisung des Herstellers oder bei Raumtemperatur, falls nicht anders angegeben)

|                              |         |   |
|------------------------------|---------|---|
| Formaldehyd                  | 16%     | Formaldehyde<br>(Thermo Scientific, #28906)   |
| Glycin                       | 2,625 M | 9,85 g Glycine<br>(Roth, #3908.2)<br>→ Ad 50 ml nuclease-freies ddH <sub>2</sub> O  |
| BSA                          | 20%     | 2 g Bovine serum albumin<br>(Sigma-Aldrich, #A4503)<br>→ Ad 10 ml nuclease-freies ddH <sub>2</sub> O  |
| Glykogen                     | 5 mg/ml | Glycogen<br>(Thermo Fisher Scientific, #AM9510)   |
| HEPES/KOH (pH 7,9)           | 1 M     | 238,3 g 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid<br>(Sigma-Aldrich, #H3357)<br>→ Mit KOH an pH 7,9 angleichen<br>→ Ad 1 l nuclease-free ddH <sub>2</sub> O  |
| KCl                          | 2 M     | 149,2 g Kaliumchlorid<br>(AppliChem, #1039)<br>→ Ad 1 l nuclease-free ddH <sub>2</sub> O  |
| EDTA/NaOH (pH 8)             | 0.5 M   | 186,1 g Ethylenediaminetetraacetic acid disodium salt dihydrate<br>(Sigma-Aldrich, #E5134)<br>→ in 500 ml nuclease-freiem ddH <sub>2</sub> O auflösen<br>→ mit NaOH an pH 8,0 angleichen<br>→ Ad 1 l nuclease-freies ddH <sub>2</sub> O |
| PMSF                         | 100 mM  | 1,74 g Phenylmethanesulfonyl fluoride<br>(Sigma-Aldrich, #P7626)<br>→ Ad 100 ml Isopropanol (p.a.)<br>(Merck Millipore, #1096342511)  |
| NaBut                        |         | Sodium butyrate<br>(Sigma-Aldrich, #B5887)  |
| Protease Inhibitor Mix (50x) |         | EDTA-free Protease Inhibitor Cocktail<br>(Sigma-Aldrich, #04693132001)  |
| NP-40                        | 10%     | Nonidet P40 Substitute<br>(Sigma-Aldrich, #11332473001)   |
| NaCl                         | 5 M     | 29,2 g Sodium chloride<br>(Merck Millipore, #106404)<br>→ Ad 100 ml nuclease-freies ddH <sub>2</sub> O  |

|   |                        |   |
|---|------------------------|---|
| Tris/HCl (pH 7,4)   | 1 M                    | 121,1 g Tris ultrapure<br>(Affymetrix (USB), #75825)<br>→ mit HCl an pH 7,4 angleichen<br>→ Ad 1 l nuclease-freies ddH <sub>2</sub> O   |
| TE (1 x, pH 8,0)  | 10 mM<br>1 mM          | 10 ml Tris/HCl (1 M, pH 8,0)<br>2 ml EDTA/NaOH (0,5 M, pH 8,0)<br>→ Ad 1 l nuclease-freies ddH <sub>2</sub> O                           |
| SDS   | 20%                    | 20 g Sodium dodecyl sulfate<br>(Sigma-Aldrich, #L4390)<br>→ Ad 100 ml nuclease-freies ddH <sub>2</sub> O                                |
| Empigen BB  | 30%                    | 3 ml Empigen BB detergent (100%)<br>(Sigma-Aldrich, #45165)<br>7 ml nuclease-freies ddH <sub>2</sub> O                                  |
| Triton X-100  | 10%                    | 1 ml Triton X-100 (100%)<br>(Sigma-Aldrich, #T8787)<br>9 ml nuclease-freies ddH <sub>2</sub> O  |
| LiCl  | 2.5 M                  | 10,6 g Lithium chloride<br>(Merck Millipore, #438002)<br>→ Ad 100 ml nuclease-freies ddH <sub>2</sub> O                                 |
| Deoxycholate  | 10%                    | 1g Sodium deoxycholate<br>(Sigma-Aldrich, #D6750)<br>→ Ad 10 ml nuclease-freies ddH <sub>2</sub> O                                      |
| NaHCO <sub>3</sub>  | 1 M                    | 84 g Sodium hydrogen carbonate<br>(Merck Millipore, #1063290500)<br>→ Ad 1 l nuclease-freies ddH <sub>2</sub> O                         |
| RNAse A   | 10 µg/µl               | RNAse A<br>(Qiagen, #1006810)   |
| Proteinase K  | 20 µg/µl               | Proteinase K<br>(Sigma-Aldrich, #03115879001)   |
| Sepharose Cl-4B Beads   | 50 µl/IP               | Sepharose Cl-4B Beads<br>(Sigma-Aldrich, #CL4B200)  |
| Cell Buffer Mix (CBM)   | 10 mM<br>85 mM<br>1 mM | 1 ml HEPES/KOH (1 M, pH 7,9)<br>4,25 ml KCl (2 M)<br>200 µl EDTA/NaOH (500 mM, pH 8,0)<br>→ Ad 97 ml nuclease-freies ddH <sub>2</sub> O |
| <i>Direkt vor Gebrauch müssen folgende Inhibitoren pro ml CBM hinzugefügt werden:</i> |                        |   |
|   | 1 mM                   | 10 µl PMSF (100 mM)<br>20 µl Protease Inhibitor Mix (50x)   |
| Suspension Buffer   |                        | 900 µl CBM inkl. Inhibitoren<br>100 µl nuclease-freies ddH <sub>2</sub> O   |

|   |                                       |  |
|---|---------------------------------------|--|
| Cell Lysis Buffer   | 1%                                    | 900 µl CBM inkl. Inhibitoren<br>100 µl NP-40 (10%)   |
| Nuclear Lysis Buffer  | 50 mM<br>1%<br>0,5%<br>10 mM          | 5 ml Tris/HCl (1 M, pH 7,4)<br>5 ml SDS (20%)<br>1,43 ml Empigen BB (30%)<br>2 ml EDTA/NaOH (500 mM, pH 8,0)<br>➔ Ad 97 ml nuclease-freies ddH <sub>2</sub> O                            |
| <i>Direkt vor Gebrauch müssen folgenden Inhibitoren pro ml Nuclear Lysis Buffer hinzugefügt werden:</i> |                                       |  |
|   | 1 mM                                  | 10 µl PMSF (100 mM)<br>20 µl Protease Inhibitor Mix (50x)  |
| Dilution Buffer (DB)  | 20 mM<br>100 mM<br>2 mM<br>0,5%       | 2 ml Tris/HCl (1 M, pH 7,4)<br>2 ml NaCl (5 M)<br>400 µl EDTA/NaOH (500 mM, pH 8,0)<br>5 ml Triton X-100 (10%)<br>➔ Ad 97 ml nuclease-freies ddH <sub>2</sub> O                          |
| <i>Direkt vor Gebrauch müssen folgenden Inhibitoren pro ml Dilution Buffer hinzugefügt werden:</i>      |                                       |  |
|   | 1 mM                                  | 10 µl PMSF (100 mM)<br>20 µl Protease Inhibitor Mix (50x)  |
| Wash Buffer<br>WBI  | 20 mM<br>150 mM<br>0,1%<br>1%<br>2 mM | 2 ml Tris/HCl (1 M, pH 7,4)<br>3 ml NaCl (5 M)<br>500 µl SDS (20%)<br>10 ml Triton X-100 (10%)<br>400 µl EDTA/NaOH (500 mM, pH 8,0)<br>➔ Ad 100 ml nuclease-freies ddH <sub>2</sub> O    |
| WBII  | 20 mM<br>500 mM<br>1%<br>2 mM         | 2 ml Tris/HCl (1 M, pH 7,4)<br>10 ml NaCl (5 M)<br>10 ml Triton X-100 (10%)<br>400 µl EDTA/NaOH (500 mM, pH 8,0)<br>➔ Ad 100 ml nuclease-freies ddH <sub>2</sub> O                       |
| WBIII<br>(vor Licht schützen!)  | 10 mM<br>250 mM<br>1%<br>1%<br>1 mM   | 1 ml Tris/HCl (1 M, pH 7,4)<br>10 ml LiCl (2,5 M)<br>10 ml NP-40 (10%)<br>10 ml Deoxycholat (10%)<br>200 µl EDTA/NaOH (500 mM, pH 8,0)<br>➔ Ad 100 ml nuclease-freies ddH <sub>2</sub> O |
| Elution Buffer  | 0.1 M<br>1%                           | 500 µl NaHCO <sub>3</sub> (1 M)<br>250 µl SDS (20%)<br>➔ Ad 5 ml nuclease-freies ddH <sub>2</sub> O  |



### 3.3.7 Library Präparation für Next Generation Sequencing

Es wurden Next Generation Sequencing (NGS) Libraries erstellt, um diese auf der Illumina Plattform (HiSeq 1000) des Kompetenzzentrums Fluoreszente Bioanalytik (KFB) der Universität Regensburg zu analysieren. Für die Library Präparation wurden durchgehend DNA-Low Binding-Proberöhrchen von Sarstedt verwendet. Alle Enzyme wurden von Enzymatics bezogen, bis auf die Phusion Hot Start High-Fidelity DNA Polymerase® (Thermo Scientific). Es wurden NEXTflex™ DNA Barcodes (Bioo Scientific) als Adapter genutzt.

Die fragmentierte DNA besitzt an ihren Enden Überhänge eines DNA-Strangs, sog. *sticky ends*, die durch die Sonifikation entstanden sind. Im ersten Schritt mussten diese Enden zu *blunt ends*, also glatten oder stumpfen Enden umgewandelt werden (s. Tab. 3-8). Hierzu benutzt man die T<sub>4</sub>-DNA-Polymerase und Klenow Fragment, die beide 3'→5'-Exonuklease- und 5'→3'-Polymeraseaktivität besitzen. Allerdings ist die 3'→5'-Exonukleaseaktivität der T<sub>4</sub>-DNA-Polymerase größer als die des Klenow Fragments, weshalb sie dazu benutzt wird, die 3'-Überhänge zu entfernen. Umgekehrt ist die 5'→3'-Polymeraseaktivität des Klenow Fragment größer, daher erfüllt es die Aufgabe, die 5'-Enden aufzufüllen (129).

**Tabelle 3-8 Reaktionszusammensetzung für End Repair der DNA**

| <i>Komponente</i>                       | <i>Volumen</i>      |
|---|---------------------|
| H <sub>2</sub> O <sub>bidest</sub>      | 3,34 µl             |
| T <sub>4</sub> DNA Ligase Buffer (10 x) | 5 µl                |
| dNTP Mix (je 10 mM)                     | 1 µl                |
| T <sub>4</sub> DNA Polymerase           | 0,30 µl             |
| Klenow Fragment                         | 0,06 µl             |
| T <sub>4</sub> Polynucleotide Kinase    | 0,30 µl             |
| IP DNA Probe                            | 40 µl               |
| <b><i>Endvolumen</i></b>                | <b><i>50 µl</i></b> |

Dieser Ansatz wurde 30 Minuten bei 20°C im PCR-Thermocycler inkubiert. Die anschließende Aufreinigung wurde nach einem erweiterten Protokoll des QIAquick MinElute PCR Purification Kit® (Qiagen) durchgeführt. Dazu wurden jeweils 250 µl PB-Puffer zu den Proben hinzugefügt, das Gemisch aus Probe und Puffer auf eine MinElute-Säule gegeben und bei 11.000 rpm 30 Sekunden zentrifugiert. Der Durchfluss wurde erneut auf die Säule pipettiert, ebenso zentrifugiert und dann

verworfen. Die Säule wurde zwei Mal mit 750 µl ethanolhaltigem PE-Puffer gewaschen und jeweils bei 13.000 rpm 45 Sekunden zentrifugiert. Danach wurde die Säule in einem neuen Abfallgefäß weitere 2 Minuten bei 13.000 rpm trocken zentrifugiert und anschließend in ein neues 1,5 ml-DNA lowbind-Proberöhrchen überführt. Der verbleibende Tropfen PE-Puffer wurde vom Gummiring in der Säule mit der Pipette entfernt und die Säule wurde eine Minute luftgetrocknet. Es wurde zwei Mal mit 10,25 µl EB-Puffer eluiert, der zuvor auf 50°C erwärmt worden war. Der EB-Puffer wurde dabei mittig auf den Filter pipettiert und dann 5 Minuten inkubiert. Jeweils eine Minute Zentrifugation bei 11.000 rpm folgten nach der Zugabe von EB-Puffer und man erhielt 18 µl Eluat.

Im nächsten Schritt wird ein Adenin an die 3'-Enden der *blunt ends* der Fragmente angefügt. Das katalysierende Enzym in dieser Reaktion ist das Klenow-(3'-5' exo-)-Fragment, das 5'→3'-Polymeraseaktivität besitzt, jedoch nicht die 3'→5'- und 5'→3'-Exonukleaseaktivität der DNA Polymerase I (s. Tab. 3-9).

**Tabelle 3-9 Reaktionszusammensetzung für die Adenylierung am 3'-Ende**

| <i>Komponente</i>            | <i>Volumen</i>   |
|------------------------------|------------------|
| Blue Buffer (10 x)           | 2 µl             |
| dATP (10 mM)                 | 0,40 µl (195 µM) |
| Klenow Fragment (3'-5' exo-) | 0,20 µl          |
| IP Probe (nach end repair)   | 18 µl            |
| <b>Endvolumen</b>            | <b>20,6 µl</b>   |

Nach einer Inkubationszeit von 30 Minuten bei 37 °C begann die Ligation der Adapter. Die Adapter tragen an ihrem 3'-Ende ein einzelnes Thymin, somit sind sie komplementär zum DNA-Fragment und lassen sich mit der T<sub>4</sub>-DNA-Ligase ligieren. Sie dienen dazu, dass sich später die DNA-Fragmente an die Next Generation Sequencing-Durchflusszelle hybridisieren lassen. Zunächst wurden die Adapter 1:2 verdünnt mit 1 x T<sub>4</sub> DNA Ligase Puffer (10 x T<sub>4</sub> DNA Ligase Puffer 1:10 verdünnt mit H<sub>2</sub>O<sub>bidest</sub>). Zu der Probe von einem Volumen von 20,6 µl wurden 22,6 µl Rapid Ligation Buffer (2x) hinzugefügt und auf Eis miteinander vermischt. Dann wurde 1 µl des 1:2 verdünnten Adapters und anschließend 1 µl der T<sub>4</sub> DNA Ligase dazugegeben und erneut auf Eis gemischt. Bei 30°C erfolgte die Ligation für 10 Minuten im Thermocycler. Um auf ein Gesamtvolumen von 50 µl zu kommen, wurden im Anschluss an die Ligation 4,8 µl H<sub>2</sub>O<sub>bidest</sub> zur Probe gegeben.

Es folgte eine Aufreinigung mit Agencourt AMPure XP-Magnetbeads® (Beckman Coulter Life Sciences) im Verhältnis 1:1,1, um überschüssige Adapter zu entfernen. Von den magnetischen Beads wurden 55 µl zu den 50 µl der Probe hinzugegeben. Unverzüglich wurde mittels Vortexer gemischt, die Reaktionsgefäße invertiert und erneut gevortext. Nach einer Inkubationszeit von 5 Minuten, wurden die Reaktionsgefäße in einen magnetischen Partikel-Konzentrator gestellt. Es wurde 2 Minuten gewartet, bis sich alle Magnetbeads an der Wand des Reaktionsgefäßes konzentriert hatten, dann wurde der Überstand abgenommen und verworfen. Nach zwei 30-sekündigen Waschschritten mit 500 µl Ethanol (70%, frisch zubereitet), wurden die Magnetbeads für 2 Minuten bei 37°C im Thermocycler getrocknet und anschließend mit 50 µl H<sub>2</sub>O<sub>bidest</sub> für 2 Minuten eluiert. Die Reaktionsgefäße wurden für eine Sekunde bei 13.000xg zentrifugiert, und das Eluat in ein neues Reaktionsgefäß überführt, um eine zweite Aufreinigung mit soeben beschriebener Abfolge durchzuführen. Als einzige Abweichung wurden am Ende der zweiten Aufreinigung die MCIP-Proben mit 25 µl und die CHIP-Proben mit 37,5 µl EB-Puffer eluiert.

Die mit den Adaptern modifizierten DNA-Fragmente wurden nun mittels einer PCR amplifiziert. Bei der Reaktionszusammensetzung und dem Ablauf der PCR wurde erneut zwischen MCIP- und CHIP-Proben unterschieden (Tab. 3-10 – 3-13).

**Tabelle 3-10 Reaktionszusammensetzung für die PCR zur Anreicherung der CHIP-Probe**

| <i>Komponente</i>                | <i>Volumen</i> | <i>Endkonzentration</i> |
|----------------------------------|----------------|-------------------------|
| Phusion HF Reaction Buffer (5 x) | 10 µl          | 1 x                     |
| BiooPrimer1 (100 µM)             | 0,25 µl        | 0,5 µM                  |
| BiooPrimer2 (100 µM)             | 0,25 µl        | 0,5 µM                  |
| dNTP Mix (10 mM)                 | 1,50 µl        | je 300 µM               |
| Phusion DNA Polymerase           | 0,50 µl        | n.a.                    |
| Adapter-gebundene CHIP-Probe     | 37,50 µl       | n.a.                    |
| <b>Endvolumen</b>                | <b>50 µl</b>   |                         |

**Tabelle 3-11 Reaktionszusammensetzung für die PCR zur Anreicherung der MClp-Probe**

| <i>Komponente</i>                | <i>Volumen</i> | <i>Endkonzentration</i> |
|----------------------------------|----------------|-------------------------|
| Phusion HF Reaction Buffer (5 x) | 10 µl          | 1 x                     |
| Betain (5 M)                     | 13 µl          | 1,3 M                   |
| BiooPrimer1 (100 µM)             | 0,25 µl        | 0,5 µM                  |
| BiooPrimer2 (100 µM)             | 0,25 µl        | 0,5 µM                  |
| dNTP Mix (10 mM)                 | 1,0 µl         | je 200 µM               |
| Phusion DNA Polymerase           | 0,50 µl        | n.a.                    |
| Adapter-gebundene MClp-Probe     | 25 µl          | n.a.                    |
| <b>Endvolumen</b>                | <b>50 µl</b>   |                         |

**Tabelle 3-12 PCR-Protokoll zur Anreicherung der ChIP-Proben (Teil 1)**

| <i>Programmschritt</i> | <i>Temperatur</i> | <i>Zeit</i> | <i>Zyklen</i> |
|------------------------|-------------------|-------------|---------------|
| Initiale Denaturierung | 98 °C             | 30 s        | 1 x           |
| Denaturierung          | 98 °C             | 10 s        | 4 x           |
| Annealing              | 65 °C             | 30 s        |               |
| DNA-Synthese           | 72 °C             | 30 s        |               |
| Finale DNA-Synthese    | 72 °C             | 5 min       | 1 x           |
| Abkühlung              | 4 °C              | ∞           |               |

**Tabelle 3-13 PCR-Protokoll zur Anreicherung der MClp-Proben (Teil 1)**

| <i>Programmschritt</i> | <i>Temperatur</i> | <i>Zeit</i> | <i>Zyklen</i> |
|------------------------|-------------------|-------------|---------------|
| Initiale Denaturierung | 98 °C             | 2 min       | 1 x           |
| Denaturierung          | 98 °C             | 30 s        | 4 x           |
| Annealing              | 65 °C             | 30 s        |               |
| DNA-Synthese           | 72 °C             | 45 s        |               |
| Finale DNA-Synthese    | 72 °C             | 5 min       | 1 x           |
| Abkühlung              | 4 °C              | ∞           |               |

Nach der PCR wurde eine weitere Aufreinigung mit Agencourt AMPure XP-Magnetbeads® (Beckman Coulter Life Sciences) wie zuvor beschrieben durchgeführt, dieses Mal im Verhältnis 1:1,8. Somit wurden zu der Probe mit einem Volumen von 50 µl 90 µl Magnetbeads hinzugefügt. Eluiert wurde bei diesem Schritt mit 10 µl H<sub>2</sub>O<sub>bidest.</sub> Das 10 µl Eluat wurde in ein neues Reaktionsgefäß überführt und mittels LabChip XT® (Caliper) einer Selektion nach Größe unterzogen. Mit dem LabChip XT DNA 750 - Chip wurden DNA-Fragmente einer Größe von 275 bp ± 15

% (234-316 bp) gewonnen. Input-DNA wurde im Bereich 250 bp  $\pm$  15% (213-287 bp) zurückerhalten. Das gewonnene Volumen wurde mit H<sub>2</sub>O<sub>bidest</sub> bis 37,5  $\mu$ l für ChIP-Proben und bis 25  $\mu$ l für MChIP-Proben aufgefüllt, um den zweiten Teil der PCR durchzuführen.

Die Reaktionszusammensetzung für den zweiten Teil der PCR entsprach jener des ersten Teils (s. Tab. 3-10 und 3-11), jedoch wurde die Anzahl der Zyklen auf 12 erhöht (Tab. 3-14 und 3-15).

**Tabelle 3-14 PCR-Protokoll zur Anreicherung der ChIP-Proben (Teil 2)**

| <i>Programmschritt</i> | <i>Temperatur</i> | <i>Zeit</i> | <i>Zyklen</i> |
|------------------------|-------------------|-------------|---------------|
| Initiale Denaturierung | 98 °C             | 30 s        | 1 x           |
| Denaturierung          | 98 °C             | 10 s        | 12 x          |
| Annealing              | 65 °C             | 30 s        |               |
| DNA-Synthese           | 72 °C             | 30 s        |               |
| Finale DNA-Synthese    | 72 °C             | 5 min       | 1 x           |
| Abkühlung              | 4 °C              | $\infty$    |               |

**Tabelle 3-15 PCR-Protokoll zur Anreicherung der MChIP-Proben (Teil 2)**

| <i>Programmschritt</i> | <i>Temperatur</i> | <i>Zeit</i> | <i>Zyklen</i> |
|------------------------|-------------------|-------------|---------------|
| Initiale Denaturierung | 98 °C             | 2 min       | 1 x           |
| Denaturierung          | 98 °C             | 30 s        | 12 x          |
| Annealing              | 65 °C             | 30 s        |               |
| DNA-Synthese           | 72 °C             | 45 s        |               |
| Finale DNA-Synthese    | 72 °C             | 5 min       | 1 x           |
| Abkühlung              | 4 °C              | $\infty$    |               |

Im Anschluss an den zweiten Teil der PCR wurden die DNA-Fragmente abschließend erneut zwei Mal mit Agencourt AMPure XP-Magnetbeads<sup>®</sup> wie oben beschrieben aufgereinigt. Das Verhältnis Magnetbeads zu Probe war in diesem Fall 1:1,1. Eluiert wurde zunächst mit 50  $\mu$ l H<sub>2</sub>O<sub>bidest</sub> und anschließend mit 16  $\mu$ l EB-Puffer.

Die Analyse von 1  $\mu$ l des Eluats auf der TapeStation<sup>®</sup> lieferte den Nachweis einer erfolgreichen Library. Die Library-Proben wurden im Kompetenzzentrum Fluoreszenz Bioanalytik (KFB) der Universität Regensburg analysiert, um NGS-Daten zu erhalten.

### 3.3.8 Assay for Transposase-Accessible Chromatin (ATAC)

Der Assay for Transposase-Accessible Chromatin (ATAC) ist eine Methode, um die genomweite Zugänglichkeit des Chromatins zu kartieren. Hierfür wird eine hyperaktive T5 Transposase verwendet, die in die zugänglichen Chromatinregionen Adapter für die anschließende Sequenzierung einfügt (97).

Für ATAC werden 50.000 Zellen in Einzelzellsuspension benötigt, die aus dem Gewebeverdaulich der vierten Brustdrüse gewonnen wurden (s. 3.3.1). Um lebendige Zellen von Zelltrümmern zu trennen, wurde eine Dichtegradientenzentrifugation mit Ficoll, einem Saccharose-Epichlorhydrin-Copolymer, durchgeführt. Hierfür wurden 50 Millionen Zellen in 15 ml PBS-Puffer auf 20 ml Ficoll (bei Raumtemperatur) gegeben und bei 700xg bei 20°C für 30 Minuten zentrifugiert. Anschließend wurde die Interphase abgenommen, einmalig mit PBS gewaschen und dann in ein kleines Volumen PBS aufgenommen. Nach Bestimmung der Zellzahl wurden je 50.000 Zellen pro Ansatz in ein 1,5 ml-Probenröhrchen für ATAC überführt.

#### 3.3.8.1 Isolation der Zellkerne

Die Zellen wurden bei 300xg bei 4°C für 8 Minuten zentrifugiert, mit 50 µl kaltem PBS gewaschen und in 50 µl kaltem 3x-Lysepuffer aufgenommen. Die Proben wurden dann 5 Minuten auf Eis inkubiert, um die Zellen zu lysieren und so die Zellkerne zu isolieren. Anschließend erfolgte eine Zentrifugation der Zellkerne bei 500xg bei 4°C, sowie mit geringem Antrieb und geringer Bremse. Die Proben wurden mit je 50 µl Transposase-Puffer (1x) gewaschen und in 50 µl Mastermix (Tab. 3-16) aufgenommen.

**Tabelle 3-16 Reaktionszusammensetzung für den Mastermix zur Zellkernisolation für eine Probe**

| <i>Komponente</i>                  | <i>Volumen</i> |
|------------------------------------|----------------|
| H <sub>2</sub> O <sub>bidest</sub> | 22,5 µl        |
| Tagment DNA Puffer (TD, 2 x)       | 25 µl          |
| Tagment DNA Enzym (TDE1)           | 2,5 µl         |
| <b>Endvolumen</b>                  | <b>50 µl</b>   |

Es folgte eine Inkubation für 30 Minuten bei 37°C und eine Aufreinigung mit dem MinElute PCR Purification Kit® (Qiagen, Hilden, Deutschland) nach Anleitung des Herstellers. Eluiert wurde jeweils in 12 µl EB-Puffer. Nach diesem Schritt konnten die Proben bei -20°C aufbewahrt werden.

### 3.3.8.2 Amplifikation und Library Präparation

Um die für die PCR benötigte Zykluszahl zu bestimmen, wurde eine quantitative PCR (qPCR) durchgeführt. Es wurde je 1 µl der Proben zu 9 µl Mastermix (Tab. 3-17; 3-18) hinzugefügt.

**Tabelle 3-17 Reaktionszusammensetzung für den qPCR-Mastermix für eine Probe**

| Komponente                         | Volumen     |
|------------------------------------|-------------|
| Sybr Green (100 x)                 | 0,1 µl      |
| Primer S                           | 0,5 µl      |
| Primer AS                          | 0,5 µl      |
| Betaine                            | 2,6 µl      |
| H <sub>2</sub> O <sub>bidest</sub> | 2,95 µl     |
| dNTP (10 mM)                       | 0,3 µl      |
| Phusion Buffer (5 x)               | 2 µl        |
| Phusion                            | 0,05 µl     |
| <b>Endvolumen</b>                  | <b>9 µl</b> |

**Tabelle 3-18 qPCR-Protokoll zur Anreicherung der ATAC-Proben**

| Programmschritt        | Temperatur | Zeit   | Zyklen |
|------------------------|------------|--|--------|
| Initiales Schmelzen    | 72 °C      | 5 min  | 1 x    |
| Initiale Denaturierung | 98 °C      | 30 s   | 1 x    |
| Denaturierung          | 98 °C      | 10 s   | 25 x   |
| Annealing              | 63 °C      | 30 s   |        |
| DNA-Synthese           | 72 °C      | 1 min  |        |
| Schmelzkurve           | 60 → 95 °C | 95 °C für 15 s, dann in<br>0,5°C Schritten abwärts | 1 x    |
| Abkühlung              | 4 °C       | ∞  |        |

Es wurde diejenige Zykluszahl für die anschließende PCR ausgewählt, bei der die Probe noch keine Sättigung erreicht hatte und der CT (*threshold cycle*)  $\frac{1}{4}$  der maximalen Intensität entsprach. Für die PCR wurden 5 µl der Probe verwendet und je 2,5 µl der jeweiligen Primerpaare, sowie 39,75 µl PCR-Mastermix (Tab. 3-19) hinzugefügt.

**Tabelle 3-19 Reaktionszusammensetzung des PCR-Mastermix für eine Probe**

| <i>Komponente</i>                  | <i>Volumen</i>  |
|------------------------------------|-----------------|
| H <sub>2</sub> O <sub>bidest</sub> | 15 µl           |
| HF Phusion Reaction Buffer (5 x)   | 10 µl           |
| dNTP                               | 1,5 µl          |
| Betaine (5 M)                      | 13 µl           |
| Phusion DNA Polymerase             | 0,25 µl         |
| <b>Endvolumen</b>                  | <b>39,75 µl</b> |

In Tabelle 3-20 ist das PCR-Programm für die Anreicherung der ATAC-Proben dargestellt, welches mit der in der qPCR bestimmten Zykluszahl durchgeführt wurde.

**Tabelle 3-20 PCR-Protokoll zur Anreicherung der ATAC-Proben**

| <i>Programmschritt</i> | <i>Temperatur</i> | <i>Zeit</i> | <i>Zyklen</i>                |
|------------------------|-------------------|-------------|------------------------------|
| Initiales Schmelzen    | 72 °C             | 3 min       | 1 x                          |
| Initiale Denaturierung | 98 °C             | 30 s        | 1 x                          |
| Denaturierung          | 98 °C             | 10 s        | In qPCR bestimmte Zykluszahl |
| Annealing              | 63 °C             | 30 s        |                              |
| DNA-Synthese           | 72 °C             | 3 min       |                              |
| Finale DNA-Synthese    | 72 °C             | 1 min       | 1 x                          |
| Abkühlung              | 4 °C              | ∞           |                              |

Im Anschluss an die PCR wurde die Probe zwei Mal mit AMPure PX<sup>®</sup> Beads aufgereinigt. Für die erste Reinigung des PCR-Produkts wurden die Beads in einem Verhältnis von 1,8:1 zur Probe eingesetzt. Diese wurden für 10 Minuten bei RT mit der Probe inkubiert und anschließend für 3 Minuten an den magnetischen Ständer gestellt. Der Überstand wurde verworfen und es wurde zwei Mal mit 200 µl Ethanol (70%) gewaschen und 5 bis 10 Minuten bei Raumtemperatur getrocknet. Anschließend wurden die Beads mit 40 µl Wasser eluiert und 5 Minuten bei Raumtemperatur inkubiert. Nach Zentrifugation und weiteren 3 Minuten am Magnetständer, wurde der Überstand in ein neues Proberöhrchen überführt. Die zweite Reinigung diente der Größenselektion und die Beads wurden zuerst in einem Verhältnis von 0,55:1 zur Probe eingesetzt und 10 Minuten bei Raumtemperatur inkubiert. Nach 3 Minuten am magnetischen Ständer wurde der Überstand in ein



neues Proberöhrchen überführt und die eingesetzten Beads verworfen. Es wurden neue Beads eingesetzt, nun wieder im Verhältnis 1,8:1 zur Probe. Gewaschen wurde analog zur ersten Reinigung (zwei Mal mit 200 µl Ethanol (70%)) und eluiert wurde mit 15 µl Wasser. Die Qualität des Eluats wurde mittels des High Sensitivity dsDNA Kit<sup>®</sup> an der Tape Station überprüft. Diese Probe konnte bei -20 °C aufbewahrt werden.

*Benötigte Reagenzien und Materialien:*

(genannte Volumen-/Gewichts-% oder Stoffmengenkonzentrationen beziehen sich auf die Endkonzentration, gelagert wurden die Reagenzien laut Anweisung des Herstellers oder bei Raumtemperatur, falls nicht anders angegeben)

|                                     |                        |  |
|-------------------------------------|------------------------|--|
| Ficoll                              |                        | GE Healthcare, Little Chalfont, Buckinghamshire, England   |
| Digitonin                           | 1% (w/v)               | Promega  |
| DNase                               | 35,5 U/µl              | Sigma-Aldrich, München, Deutschland  |
| NP-40                               | 10% (w/v)              | Sigma-Aldrich, München, Deutschland  |
| Tween-20                            | 10% (w/v)              | Sigma-Aldrich, München, Deutschland  |
| Lysepuffer                          |                        | Nextera, Illumina, San Diego, California, USA  |
| Nuclease-freies Wasser              |                        | Gibco/Thermo Fisher Scientific, Waltham, MA, USA   |
| 2x TD (Tagment DNA Puffer)          |                        | Nextera, Illumina, San Diego, California, USA  |
| Sybr Green (100 x)                  |                        | Qiagen, Hilden, Germany  |
| Primer S/AS                         |                        | TruSeq Dual Index Sequencing Primer Box<br>(Illumina, #PE-121-1003)  |
| Nextera DNA Library Preparation Kit |                        | Illumina, San Diego, USA   |
| Betaine (5 M)                       |                        | Sigma-Aldrich, München, Deutschland  |
| HF Phusion Reaction Buffer (5 x)    |                        | Thermo Fisher Scientific, Waltham, USA   |
| Phusion DNA Polymerase              |                        | Thermo Fisher Scientific, Waltham, USA   |
| Agencourt AMPure XP Beads           |                        | Agencourt AMPure XP<br>(Beckman Coulter Life Sciences, #A63881)  |
| Lysepuffer:                         | 10 mM<br>10 mM<br>3 mM | 0,5 ml Tris/HCl (1 M), pH 7.4<br>0,1 ml NaCl (5 M)<br>0,15 ml MgCl <sub>2</sub> (1 M)<br>➔ Ad 50 ml nuclease-freies ddH <sub>2</sub> O |
| 3x Lysepuffer:                      |                        | 970 µl Lysepuffer  |

|                     |       |                      |
|---------------------|-------|----------------------|
|                     | 0,1%  | 10 µl NP-40 (10%)    |
|                     | 0,1%  | 10 µl Tween-20 (10%) |
|                     | 0,01% | 10 µl Digitonin (1%) |
| Transposase-Puffer: |       | 990 µl Lysepuffer    |
|                     | 0,1%  | 10 µl Tween-20 (10%) |

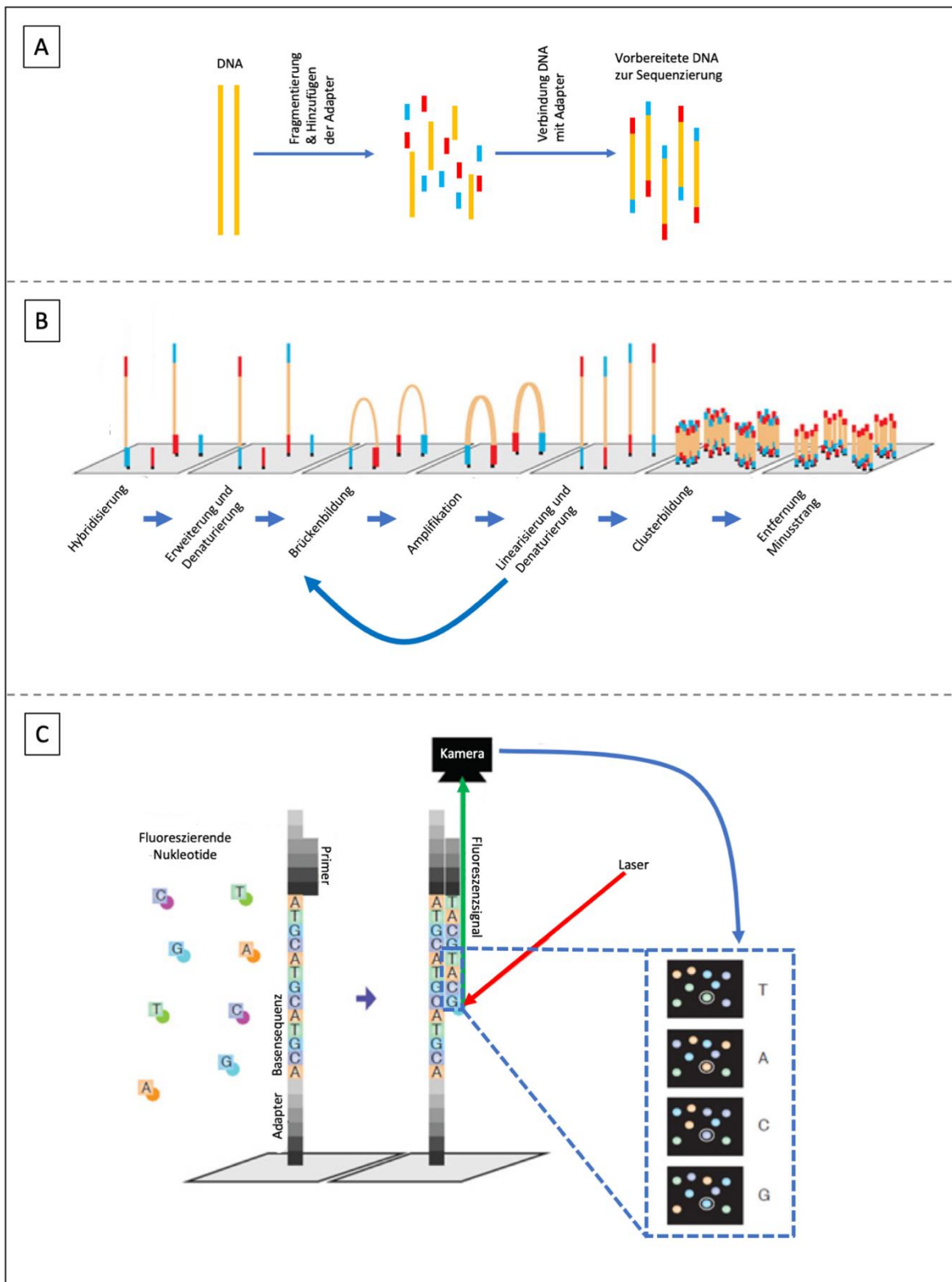
### 3.4 Next Generation Sequencing

Next Generation Sequencing (NGS) ist ein breiter Begriff, der sich auf eine Gruppe von Methoden bezieht, die (1) der Herstellung eines genomischen Templates oder der Prozessierung von genomischer DNA für nachgeschaltete Sequenzierung mit (2) anschließender beinahe simultaner oder massiver paralleler Generierung von Millionen bis Milliarden kurzer Sequenzabschnitte, (3) der Abgleichung („*alignment*“) der Sequenzabschnitte mit einer Referenzsequenz und schließlich der Zusammensetzung der abgeglichenen Sequenzabschnitte und Entdeckung von genetischen und epigenetischen Varianten dient (130).

Für das NGS im Rahmen dieser Arbeit wurde die Plattform Illumina HiSeq 2000® (Illumina, San Diego, Kalifornien, USA) genutzt. Für die Sequenzierung verwendet Illumina eine Amplifizierungsreaktion mittels „Brückenbildung“, welche auf der Oberfläche einer Durchflusszelle stattfindet. Die Oberfläche der Durchflusszelle ist mit Einzelstrang-Oligonukleotiden benetzt, die den in der Herstellung der Library (s. 3.3.7) verwendeten Adaptoren entsprechen. An diese werden willkürlich die Einzelstrang-DNA-Fragmente mit passenden Adaptoren gebunden. „Brücken-Amplifikation“ findet zur Herstellung doppelsträngiger Fragmente statt. Das Priming findet dabei an den zunächst freien Enden der Einzelstrangfragmente statt, welche sich an korrespondierende Oligonukleotide durch Brückenbildung anlagern und bei Zugabe von Nukleotiden und Enzymen amplifiziert werden. Durch lokale Denaturierung und Verlängerung der sogenannten Cluster-Amplifikation an Millionen Stellen der Oberfläche der Durchflusszelle werden DNA-Kopien erstellt, die dichte Cluster bilden (Abb. 4B).

Illuminas „*Sequencing by Synthesis*“ basiert auf der Methode der „*Cyclic Reversible Termination*“ (CRT). Um die Cluster zu sequenzieren wird ein Universalprimer an die Adaptersequenz der DNA-Abschnitte gebunden. Die vier Deoxynukleotide sind jeweils an eine fluoreszierende Farbe gekoppelt und durch eine reversible Terminatorgruppe modifiziert. Anschließend findet in einem Reaktionszyklus die

Verlängerung um ein komplementäres, fluoreszenzmarkiertes Nukleotid durch die Polymerase statt, da die DNA-Synthese danach durch die blockierende Gruppe unterbrochen wird. Nicht eingebaute Nukleotide werden gewaschen. Durch Laseranregung wird das Lichtsignal dieses Nukleotids detektiert, anschließend wird die Terminatorgruppe und die Fluoreszenzmarkierung abgespalten und der Vorgang wiederholt. Somit entstehen vierfarbige Bilder zur Entschlüsselung der Basensequenz. Ein Cluster wird als eine Base erkannt, da es identische DNA-Sequenzen enthält. Zum Schluss folgt ein Abgleichen und Vergleichen der Sequenzabschnitte mit einem Referenzgenom, um Mutationen festzustellen (Abb. 4C).



**Abbildung 4: Darstellung der Sequenzierung per Next Generation Sequencing.** A) Vorbereitung der DNA zur Amplifizierung per Fragmentierung und Bindung von Adaptoren an beiden Enden. B) Darstellung der sogenannten „Brückenamplifikation“ in einer Durchflusszelle. Durch diese Amplifikation entstehen schlussendlich in jedem Kanal der Durchflusszelle mehrere Millionen dichte Cluster aus doppelsträngiger DNA. C) Sequenzierung der DNA. Hierfür werden mit fluoreszierenden Substanzen markierte Nukleotide durch die DNA-Polymerase eingebaut und Basenpaare werden durch Laseranregung und Fluoreszenzdetektion identifiziert. Modifiziert nach Chaitankar, 2016 (131).

### 3.5 Bioinformatische Auswertung

#### 3.5.1 DNA-Methylierung durch MCIp

Durch die Sequenzierung der Methyl-CpG-Immunopräzipitation Libraries (MCIp-seq) konnten die Daten für eine weitere Auswertung gewonnen werden.

##### 3.5.1.1 Datensatz

Der für die nachfolgende Analyse verwendete Datensatz wurde aus 88 Proben generiert. Davon waren vier (durch *SssI*-Behandlung) vollständig methylierte Proben aus normalem Brustgewebe, zwei Kontrollproben aus normalem Brustgewebe, zwei Kontrollproben aus normalem Lungengewebe und je 40 Proben aus Primärtumoren- und Metastasenproben. Letztere konnten teilweise einander zugeordnet werden (Tab. 3-21).

**Tabelle 3-21 Anzahl zugeordneter Primärtumoren und Metastasen**

| Anzahl an Mäusen | Anzahl an Metastasen |
|------------------|----------------------|
| 16               | 0                    |
| 14               | 1 (14*1 = 14 Proben) |
| 4                | 2 (4*2 = 8 Proben)   |
| 6                | 3 (6*3 = 18 Proben)  |

Daraus ist ersichtlich, dass 16 Mäuse nur einen Primärtumor und keine Metastasen, vier Mäuse aber zum Beispiel einen Primärtumor und zwei Metastasen aufwiesen. Das ergibt insgesamt 12 Proben (4\*1 Primärtumor + 4\*2 Metastasen = 12 Proben) in unserem Datensatz. Wenn Mäusen in diesem Datensatz keine Metastasen zum Primärtumor zugeordnet wurden, bedeutet das jedoch nicht, dass sie keine Metastasen entwickelt haben.

##### 3.5.1.2 Datenmapping

Die MCIp-seq-Daten wurden mithilfe von Bowtie2 und der voreingestellten Option „--sensitive-local“ auf das GRCm38 Maus-Referenzgenom (mm10) abgebildet. Die daraus hervorgegangenen SAM-Dateien wurden in BAM-Dateien konvertiert und anschließend mit *samtools* sortiert und kartiert.

##### 3.5.1.3 Bioinformatische Aufbereitung der Daten

Alle Regionen, die in unseren Proben potenziell methyliert werden können, wurden mithilfe der Software *HOMER* aus den vier vollständig methylierten Proben (durch

SssI-Behandlung) abgeleitet. Diese vier Proben wurden zusammengefasst und nur Regionen, die in allen vier Proben methyliert waren, wurden beibehalten. Weiterhin wurden Regionen, die Artefakte in Next-Generation Sequencing-Experimenten produzierten, solche die sich in bekannten repetitiven DNA-Abschnitten (mindestens 75% repetitive DNA) befanden und Regionen mit mangelhafter Kartierbarkeit aus den nachfolgenden Analysen ausgeschlossen. Des Weiteren wurden nur Sequenzabschnitte mit einer mittleren Abbildungsqualität (*Mappability Score* von 0,8) analysiert.

Die jetzt noch vorhandenen Regionen bildeten die Grundlage der folgenden Analysen als Kontrollen im Sinne einer maximalen Methylierung. Die durch NGS gewonnenen Sequenzabschnitte aller restlichen Proben wurden diesen zuvor definierten Regionen zugeordnet. Hierbei existieren pro Region jeweils mehrere Sequenzabschnitte, die methyliert sein können. Die Anzahl der Sequenzabschnitte pro Region wurde zunächst normalisiert, indem sie mit einer virtuellen Referenzprobe verglichen wurde, die aus dem geometrischen Mittelwert aller ermittelten Sequenzabschnitte gebildet wurde.

#### **3.5.1.4 Analyse der Methylierungsmuster**

Um die Übereinstimmung der Methylierungsmuster von Primärtumoren und Metastasen beurteilen zu können, wurden verschiedene Methoden verwendet, welche die Proben jeweils bezüglich ihrer Ähnlichkeit im Methylierungsmuster einordnen: die multidimensionale Skalierung (MDS), die Hauptkomponentenanalyse (Principal Component Analysis; PCA) und die *T-distributed Stochastic Neighbor Embedding* (t-SNE). MDS-Plots basieren auf einer Principal Component Analyse, mithilfe welcher die Ähnlichkeit der Methylierung aller Proben in einem zweidimensionalen Raum visualisiert werden kann. Je ähnlicher sich die entsprechenden Proben in ihrem Methylierungsmuster sind, desto näher liegen die einzelnen Punkte also zusammen. Diese Analysen wurden mithilfe von *R statistics* (Version 3.6.1) durchgeführt.

#### **3.5.1.5 Regionsbasierte Analyse**

Die in dieser Arbeit gewonnenen Methylierungsdaten wurden für die regionsbasierte Analyse am Lehrstuhl für Statistische Bioinformatik der Universität Regensburg im Rahmen des kooperierenden bioinformatischen Projekts der Forschergruppe FOR2127 weiter analysiert. Unterschiedliche Methylierung zwischen

Primärtumoren und Metastasen wurde mit dem Paket *DESeq2* aus dem Statistikprogramm R (Version 3.4.3) ausgewertet (126). Alle Regionen, die in den Metastasen-Proben signifikant hyper- oder hypomethyliert waren, wurden dann bezüglich ihrer funktionellen Annotation untersucht. Dazu wurden die Regionen in das Internetprogramm *GREAT (Genomic Regions Enrichment Annotation Tool)* eingelesen, welches die Regionen bestimmten Genen zuordnet. Um nur die für den hier verwendeten Versuchsaufbau relevanten Gene zu ermitteln, wurde als Hintergrund der Analyse die Gesamtheit der methylierbaren DNA-Regionen verwendet (via *SssI*-Behandlung der DNA ermittelt). Anschließend wurde eine Literaturrecherche durchgeführt, um festzustellen, welche der hypo- oder hypermethylierten Gene bereits mit Mammakarzinom und DNA-Methylierung in Verbindung gebracht wurden. Hierbei wurden Genlisten zum Thema Brustkrebs und Methylierung durchsucht (24,132,133). In einer weiteren Recherche auf Pubmed (Suchbegriffe: [Genname] + methylation; [Genname] + breast cancer; [Genname] + methylation + breast cancer; [Genname] + methylation + breast cancer + metastasis) wurden die aus den Genlisten identifizierten Gene weiter aufgeschlüsselt und evaluiert, ob die Gene bereits mit Methylierung, Brustkrebs oder Metastasierung assoziiert waren.

Außerdem wurden alle hyper- oder hypomethylierten Gene mithilfe des Programms *Metascape* (128) bezüglich ihrer Funktionalität untersucht. In *Metascape* sind alle bekannten Gene des Genoms eingelesen. Hierzu wurde die den *SssI*-Regionen zugeordneten Gene als Hintergrund verwendet. Für jede Genliste wurde eine Anreicherungsanalyse bezüglich Pathways und Prozessen durchgeführt. *Metascape* verwendet *GO Biological Processes*, *KEGG Pathway*, *Reactome Gene Sets* und *CORUM* als ontologische Quellen. Gruppen von funktionell ähnlichen Genen werden mit *Metascape* in Clustern dargestellt. Für diese Cluster wurden diejenigen Gruppen ausgewählt, in denen mindestens drei Gene einen hochsignifikanten Unterschied zwischen Primärtumoren und Metastasen zeigten ( $p < 0,01$ ). Außerdem sollten die Gruppen einen Anreicherungsfaktor von  $> 1,5$  aufweisen. Der Anreicherungsfaktor ist das Verhältnis zwischen den festgestellten Signalen und den Signalen, die man per Zufall erwarten würde. Die statistisch signifikanteste Funktionsbezeichnung innerhalb eines Clusters wurde jeweils ausgewählt, um das Cluster zu repräsentieren.

## 4 Ergebnisse

Ziel dieser Arbeit war es, den Beitrag epigenetischer Mechanismen zur Metastasierung von Brustkrebs in einem Mausmodell zu untersuchen. Dabei wurden spezifische epigenetische Veränderungen, insbesondere die DNA-Methylierung, und deren Ausprägung bei Primärtumoren und Metastasen untersucht.

### 4.1 Übersicht über Proben und durchgeführte Experimente

Wie in den Methoden dargestellt, wurden für die nachfolgenden Experimente Proben aus 40 Primärtumoren und 40 teils gepaarten Lungenmetastasen, sowie Gewebeproben aus zwei gesunden Brustdrüsen und zwei gesunden Lungen erhoben. Zusätzlich haben wir vier Gewebeproben aus gesunden Brustdrüsen mittels SssI behandelt, um eine vollständig methylierte Referenzprobe zu erhalten. Zum besseren Verständnis der Ergebnisse zeigt Tabelle 4-1 die angewendeten Experimente pro Gewebe. Im Nachfolgenden werden lediglich die Ergebnisse zu den MCIp-Daten dargestellt, die ChIP- und ATAC-Experimente, die mit den gesunden Gewebeproben durchgeführt wurden, sind nicht Teil dieser Auswertungen.

**Tabelle 4-1 Probenübersicht und durchgeführte Experimente pro Gewebe**

| <i>Gewebe</i>                                  | <i>Experimente</i>   |
|--|--|
| 4x gesundes Brustgewebe                        | 4x DNA-Komplettmethylierung mittels SssI<br>als Referenz<br>ChIP, ATAC |
| 2x gesundes Brustgewebe                        | MCIp   |
| 2x gesundes Lungengewebe                       | MCIp   |
| <b>40x Gewebe aus Mamma-<br/>Primärtumoren</b> | MCIp   |
| 40x Gewebe aus Lungenmetastasen                | MCIp   |

Die kryokonservierte genomische DNA (gDNA) der Primärtumor- und Metastasenproben wurde für die Methylierungsanalysen mittels MCIp verwendet. Hierfür standen 40 Metastasen und 40 Primärtumoren von insgesamt 40 verschiedenen transplantierten Mäusen zur Verfügung. Von den Primärtumoren hatten 24 Proben gepaarte Metastasen (1-3 Metastasen pro Primärtumor, Tabelle



4-2, siehe auch 3.5.1.1). Von den weiteren 16 Mäusen gab es jeweils nur den Primärtumor.

**Tabelle 4-2 Probenübersicht Primärtumoren und Metastasen**

| <i>Maus-ID</i> | <i>Probennr.</i> | <i>Gewebe</i>     |
|----------------|------------------|-------------------|
| 3145           | 1                | Lungenmetastase 2 |
| 3145           | 53               | Primärtumor       |
| 3145           | 54               | Lungenmetastase 1 |
| 3315           | 8                | Lungenmetastase 2 |
| 3315           | 9                | Lungenmetastase 3 |
| 3315           | 61               | Primärtumor       |
| 3315           | 62               | Lungenmetastase 1 |
| 3500           | 48               | Primärtumor       |
| 3511           | 63               | Primärtumor       |
| 3511           | 64               | Lungenmetastase   |
| 3519           | 37               | Primärtumor       |
| 3522           | 10               | Lungenmetastase 2 |
| 3522           | 65               | Primärtumor       |
| 3522           | 66               | Lungenmetastase 1 |
| 3524           | 19               | Primärtumor       |
| 3524           | 20               | Lungenmetastase   |
| 3525           | 21               | Primärtumor       |
| 3525           | 22               | Lungenmetastase 1 |
| 3525           | 23               | Lungenmetastase 2 |
| 3525           | 24               | Lungenmetastase 3 |
| 3555           | 38               | Primärtumor       |
| 3569           | 25               | Primärtumor       |
| 3569           | 26               | Lungenmetastase   |
| 3576           | 27               | Primärtumor       |
| 3576           | 28               | Lungenmetastase   |
| 3599           | 39               | Primärtumor       |
| 3612           | 46               | Primärtumor       |
| 3615           | 47               | Primärtumor       |
| 3620           | 6                | Lungenmetastase 2 |
| 3620           | 7                | Lungenmetastase 3 |
| 3620           | 59               | Primärtumor       |
| 3620           | 60               | Lungenmetastase 1 |
| 3627           | 79               | Primärtumor       |
| 3627           | 80               | Lungenmetastase   |
| 3631           | 77               | Primärtumor       |
| 3631           | 78               | Lungenmetastase   |
| 3635           | 67               | Primärtumor       |
| 3635           | 68               | Lungenmetastase   |

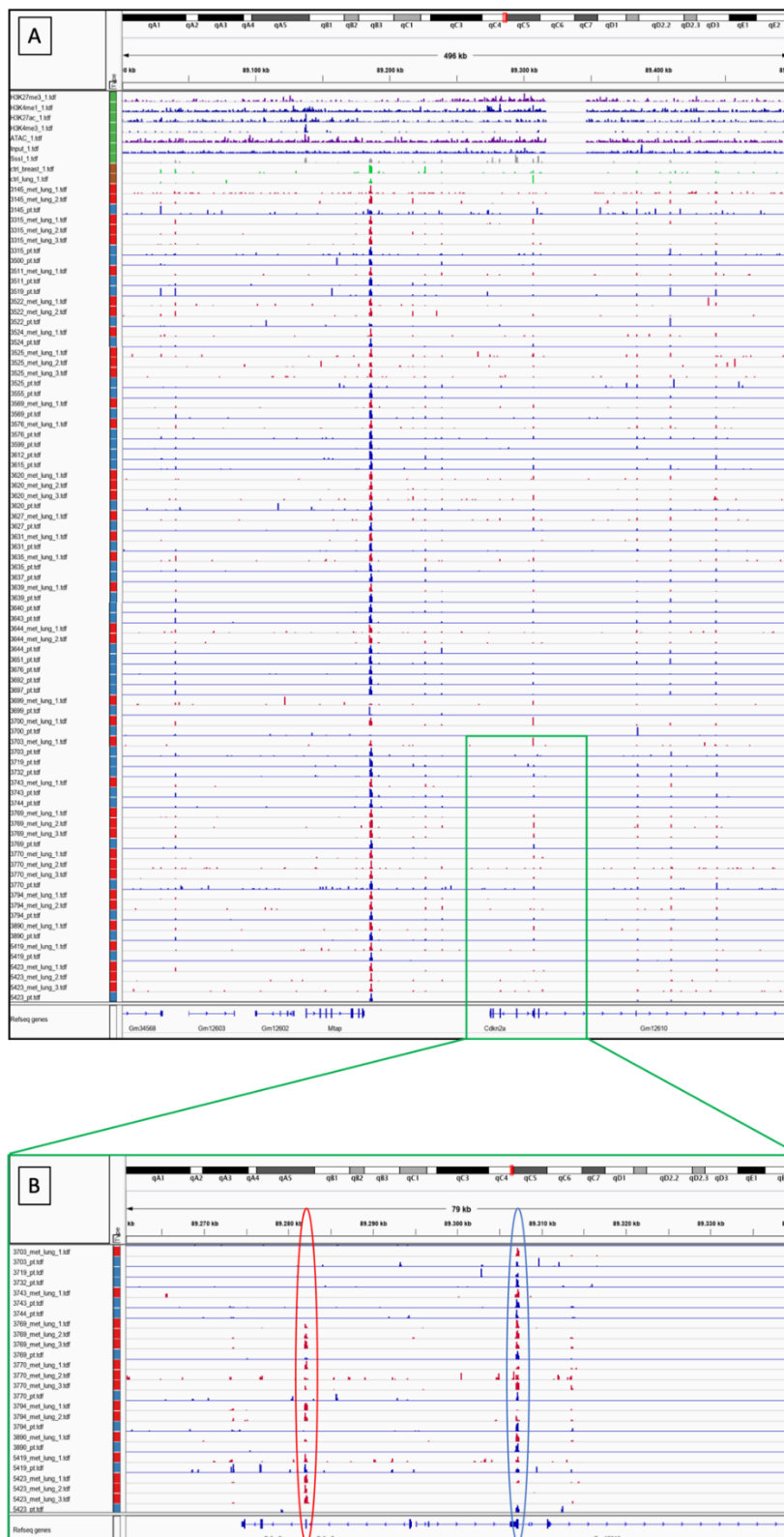
|      |    |                   |
|------|----|-------------------|
| 3637 | 52 | Primärtumor       |
| 3639 | 29 | Primärtumor       |
| 3639 | 30 | Lungenmetastase   |
| 3640 | 41 | Primärtumor       |
| 3643 | 40 | Primärtumor       |
| 3644 | 11 | Lungenmetastase 2 |
| 3644 | 73 | Primärtumor       |
| 3644 | 74 | Lungenmetastase 1 |
| 3651 | 42 | Primärtumor       |
| 3676 | 44 | Primärtumor       |
| 3692 | 43 | Primärtumor       |
| 3697 | 51 | Primärtumor       |
| 3699 | 71 | Primärtumor       |
| 3699 | 72 | Lungenmetastase   |
| 3700 | 69 | Primärtumor       |
| 3700 | 70 | Lungenmetastase 1 |
| 3703 | 31 | Primärtumor       |
| 3703 | 32 | Lungenmetastase   |
| 3719 | 45 | Primärtumor       |
| 3732 | 49 | Primärtumor       |
| 3743 | 33 | Primärtumor       |
| 3743 | 34 | Lungenmetastase   |
| 3744 | 50 | Primärtumor       |
| 3769 | 2  | Lungenmetastase 2 |
| 3769 | 3  | Lungenmetastase 3 |
| 3769 | 55 | Primärtumor       |
| 3769 | 56 | Lungenmetastase 1 |
| 3770 | 4  | Lungenmetastase 2 |
| 3770 | 5  | Lungenmetastase 3 |
| 3770 | 57 | Primärtumor       |
| 3770 | 58 | Lungenmetastase 1 |
| 3794 | 12 | Lungenmetastase 2 |
| 3794 | 75 | Primärtumor       |
| 3794 | 76 | Lungenmetastase 1 |
| 3890 | 35 | Primärtumor       |
| 3890 | 36 | Lungenmetastase   |
| 5419 | 13 | Primärtumor       |
| 5419 | 14 | Lungenmetastase   |
| 5423 | 15 | Primärtumor       |
| 5423 | 16 | Lungenmetastase 1 |
| 5423 | 17 | Lungenmetastase 2 |
| 5423 | 18 | Lungenmetastase 3 |

## 4.2 Übersicht über die MCIp-Daten

Der Auswertung wurde eine vollständig methylierte Referenzprobe zugrunde gelegt, welche mittels SssI-Behandlung erzeugt wurde. Die gDNA der jeweils 40 Primärtumor- und Metastasenproben und die der vier Kontrollproben (je zwei Mal gesundes Brust- und Lungengewebe) wurde mittels Methyl-CpG-Immunopräzipitation (MCIp) angereichert und anschließend sequenziert, um die Methylierungsmuster der Proben zu analysieren. Damit wurde untersucht, ob es unterschiedliche Muster in den verschiedenen Geweben gibt.

Wie in den Methoden beschrieben, wurden die Sequenzierdaten mithilfe der Programme Bowtie2 und HOMER mit dem Maus-Referenzgenom abgeglichen, sortiert, kartiert und die angereicherten Regionen (sogenannte *Peaks*) identifiziert. Nach der Qualitätskontrolle (Ausschluss von Regionen mit Artefakten, schlechter Abbildungsqualität oder mangelhafter Kartierbarkeit, sowie von Regionen aus Abschnitten repetitiver DNA) blieben schließlich 31.573 angereicherte Regionen in der zusammengefassten Referenzprobe mit SssI-Behandlung. In diesen vordefinierten Regionen wurden die *Peaks* der 40 Primärtumoren (PT) und der 40 Metastasen (MET) gezählt.

Mithilfe des IGV-Browsers wurden die Methylierungsmuster unserer Proben über das gesamte Mausgenom dargestellt (Abb. 5A). Dabei zeigte sich, dass manche Stellen im Genom bei Metastasen im Vergleich zu Primärtumoren vermehrt methyliert sind, oder Methylierung an Stellen im Genom vorhanden ist, die im Primärtumor noch nicht methyliert sind (Abb. 5B). Aufgrund der rein deskriptiven Darstellung der Ergebnisse in diesem Ansatz wurden anschließend über verschiedene bioinformatische Auswertungen die Unterschiede zwischen Metastasen und Primärtumoren noch genauer dargestellt. Ziel dieser weiteren Analyse war es, herauszufinden, inwiefern sich die Proben aus verschiedenen Geweben bezüglich ihres Methylierungsmusters ähneln beziehungsweise unterscheiden.



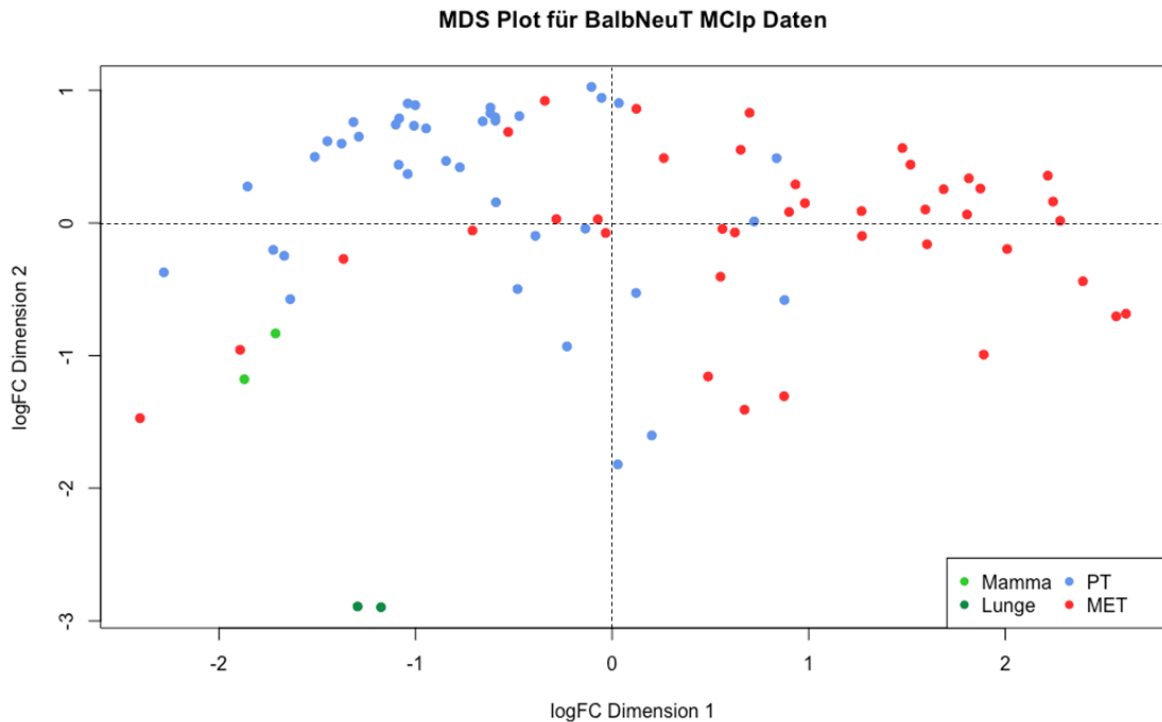
**Abbildung 5:** Darstellung der MCIP-Daten der BalbNeU Mäuse im IGV-Browser. Abb. 5A: Übersicht über mehrere Proben (q-Arm Chromosom 4; chr4:89,000,710-89,500,856). Farbkodierung: PT blau, MET rot, Histonmodifizierungen und ATAC-Daten violett, komplettmethylierte Proben grau, Kontrollproben aus Mammagewebe grün. Abb. 5B: Methylierungs-Peaks einiger Proben im CDKN2A-Gen (vergrößerter Ausschnitt (grün) aus Abb. 5A mit Hervorhebung zweier Methylierungsregionen: In MET hypermethylierte Region (rot umrandet); allgemein hohe Methylierungsdichte (blau umrandet). Farbkodierung der Methylierungspeaks: PT blau, MET rot.

### 4.3 Clustering der BalbNeuT-Proben

Die Methylierungsmuster aller 88 Proben wurden mit verschiedenen Methoden des Clusterings genauer analysiert. Hierbei war das Ziel darzustellen, inwiefern sich die untersuchten Proben bezüglich ihres globalen Methylierungsmusters ähneln oder unterscheiden. Die drei hierzu verwendeten Methoden sind die multidimensionale Skalierung (MDS), eine Hauptkomponentenanalyse (Principal Component Analysis; PCA) und die *T-distributed Stochastic Neighbor Embedding* (t-SNE).

#### 4.3.1 Multidimensionale Skalierung (MDS)

Zunächst wurde eine multidimensionale Skalierung (MDS) durchgeführt und die Ergebnisse in einem entsprechenden MDS-Plot dargestellt. In Abbildung 6 sind die ersten zwei Dimensionen der MDS des Gen-Methylierungsmusters auf einer logarithmischen Skala abgebildet. Proben, die Ähnlichkeiten hinsichtlich der Methylierung aufweisen, haben in dieser Darstellung eine geringe Distanz zueinander. Hierbei lässt sich erkennen, dass die meisten Metastasenproben in der ersten Dimension einen positiven Wert haben, während sowohl die Primärtumoren als auch die vier Proben aus gesundem Gewebe auf Dimension 1 tendenziell negative Werte aufweisen. Die zweite Dimension scheint vor allem zwischen Lungen- und Brustgewebe zu differenzieren, da lediglich die Lungenproben einen stark negativen Wert aufweisen, während Proben aus Primärtumorgewebe, gesunder Brust und metastasiertem Brustkrebs im Vergleich deutlich näher am neutralen Nullwert liegen. Dies zeigt, dass die Methode valide Ergebnisse liefert, da im MDS-Plot gesundes Lungengewebe von allen pathologischen Geweben differenziert werden kann. Zweitens wird deutlich, dass sich Metastasenproben von den restlichen Proben unterscheiden, da sie im Vergleich eher positive Werte in der ersten Dimension aufweisen. Es gibt aber einige Metastasenproben, die näher am Methylierungsmuster der Primärtumoren liegen. Insgesamt zeigt sich also ein variables Bild des Methylierungsmusters im MDS-Plot.

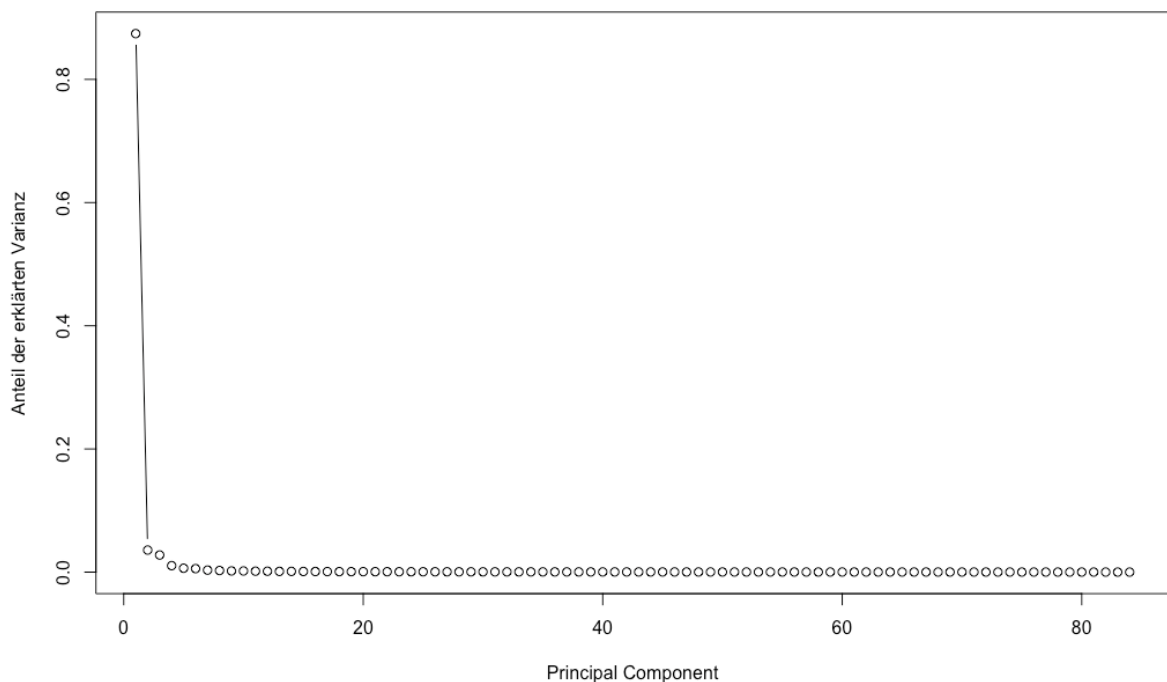


**Abbildung 6: Multidimensionale Skalierung, MDS-Plot für die Methylierungsdaten (MClp).** Darstellung der Proben: 40 Primärtumoren (PT), 40 Metastasen (MET), zwei Proben aus normalem Brustgewebe (Mamma) und zwei Proben aus normalem Lungengewebe (Lunge). LogFC = logarithmierter Fold Change.

#### 4.3.2 Principal Component Analyse (PCA)

Als weitere Methode, um aus den Daten einen Clusterplot des Methylierungsmusters zu generieren und Muster in dem hochdimensionalen Datensatz zu erkennen, wurde eine Principal Component Analyse (PCA), auch Hauptkomponentenanalyse genannt, durchgeführt. Eine PCA ist eine Unterform der multidimensionalen Skalierung. Um die vieldimensionalen Daten der Methylierungsmuster zu vereinfachen und zu veranschaulichen, wird die Vielzahl an Variablen durch eine Transformation in Hauptkomponenten angenähert (134,135). Diese Komponenten werden in einem iterativen Prozess ermittelt, wobei die erste Komponente jeweils die größte Varianz des Datensatzes erklärt, die zweite Komponente die zweitgrößte Varianz, und so weiter (siehe Abb. 7). Das Verfahren kommt der Erzeugung einer idealen Linie durch die Daten nahe, welche die maximal mögliche Varianz erklären soll. Diese Annäherung wird in mehreren Schritten noch optimiert, bis eine ideale Variante gefunden wird. Anschließend wird das gleiche Verfahren mit einer zweiten Dimension durchgeführt, bis die gesamte Varianz erklärt ist. Dadurch wird ermöglicht, die komplexen Daten mit einer geringeren

Anzahl an Dimensionen darzustellen. Ein weiterer Vorteil der PCA besteht darin, dass es trotz Reduktion der Dimensionen kaum zum Datenverlust kommt, da alle Informationen in den aussagekräftigen Komponenten integriert werden. In Abbildung 7 ist zu sehen, dass beinahe 90 % der Varianz der gesamten Methylierungsdaten durch die erste Hauptkomponente erklärt werden. Die nächsten beiden Komponenten erklären weitere 3,6 bzw. 2,8 % der Varianz, während die restlichen Komponenten nach der iterativen Analyse vernachlässigbar wenig Information beitragen.



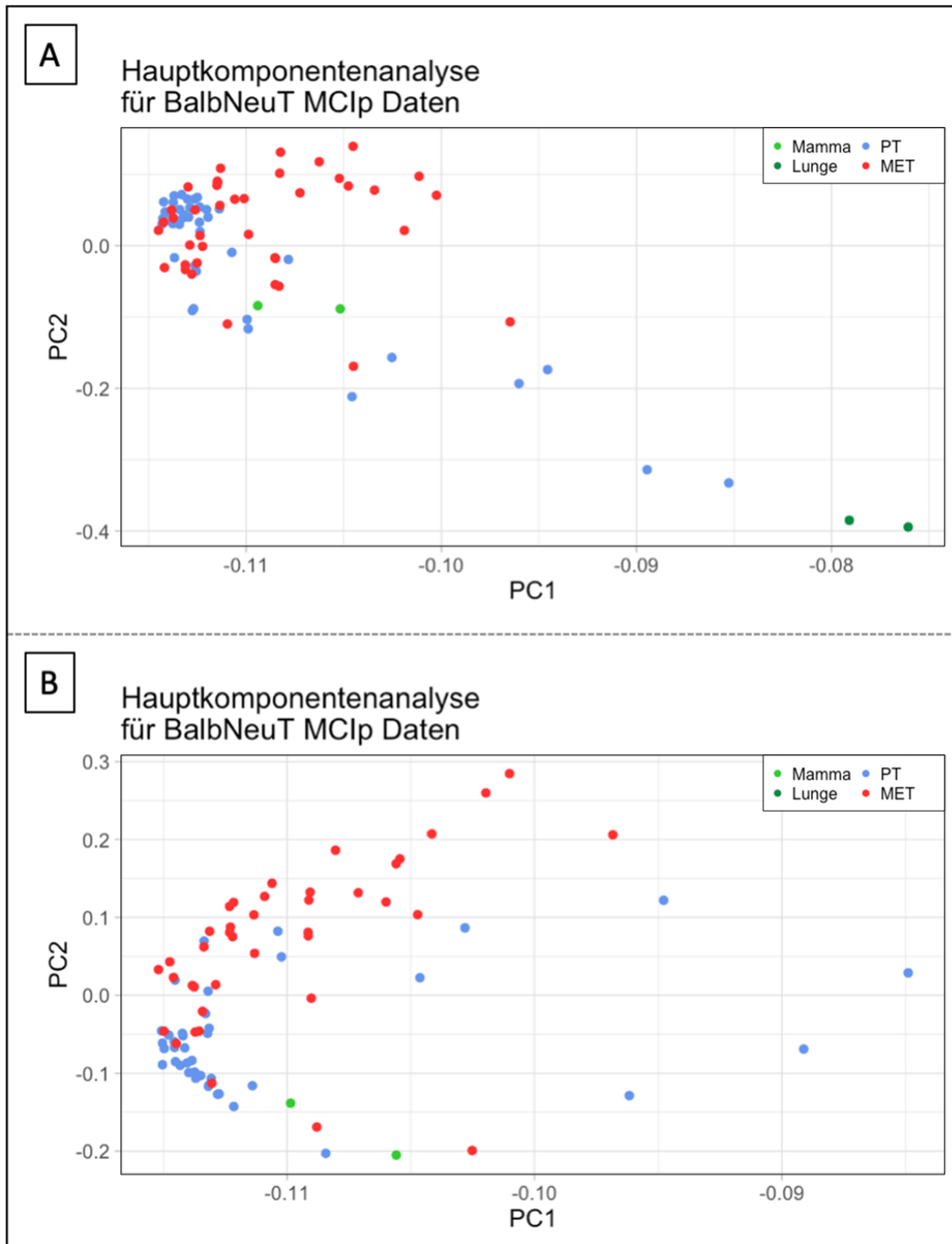
**Abbildung 7: Principal Component Analyse (PCA) der Methylierungsdaten.** Die erste Hauptkomponente erklärt 90% der Varianz der Daten, die zweite und dritte Hauptkomponente erklären < 5% der Varianz.

Die ersten beiden Komponenten können dann zweidimensional abgebildet werden (Abb. 8A). Sowohl die erste als auch die zweite Komponente (entsprechend den Achsen PC1 und PC2 in Abb. 8A) erfassen die wesentlichen Unterschiede der Gen-Methylierung zwischen gesundem Lungengewebe und den restlichen Gewebearten. Allerdings lassen sich auch Aussagen über die Unterschiede zwischen Primärtumoren- und Metastasenproben treffen: Ein Teil der Primärtumoren bildet ein dichtes Cluster (blaue Punkte in Abb. 8A), was darauf hinweist, dass ein ähnliches Methylierungsmuster vorliegt. Es gibt allerdings auch einige Primärtumoren, welche nicht direkt im blauen Cluster liegen. Dahingegen

liegen die Metastasenproben weniger nah zusammen und weisen somit eine größere Variabilität ihrer Methylierungsmuster auf. So scheinen einige Metastasen ein ähnliches Muster wie die Primärtumoren zu haben, andere unterscheiden sich deutlicher von den Primärtumoren.

Interessanterweise zeigt sich in dieser Darstellung, dass das Methylierungsmuster von Genen in gesunden Mammae keinen großen Unterschied zu jenen aus Primärtumoren oder Metastasen aufweist. Um die Beeinflussung der Daten durch die Proben des gesunden Lungengewebes zu eliminieren, wurde die PCA ohne gesundes Lungengewebe wiederholt. Dadurch erfolgte eine deutlichere Gruppierung des Großteils der Primärtumoren (Abb. 8B). Einige Primärtumorproben liegen jedoch weiter entfernt vom Hauptcluster als alle Metastasen. Die Methylierungsmuster der Metastasen zeigen dabei insgesamt eine größere Varianz, wobei einige auch den Proben im Primärtumorcluster ähnlich sind. Interessanterweise erscheint es hier so, als ob die zweite Hauptkomponente (PC2) relativ gut zwischen der Methylierung der Primärtumoren und Metastasen unterscheiden kann: 75,6% der negativ auf PC2 ladenden Proben sind Primärtumoren (31 von 41). Gleichzeitig laden 77,5% der Primärtumoren auf PC2 negativ (31 von 40). Demgegenüber sind positiv ladende Proben eher den Metastasen zuzuordnen (32 von 43 Proben = 74,4%). 80% der Metastasen laden auf PC2 positiv (32 von 40).



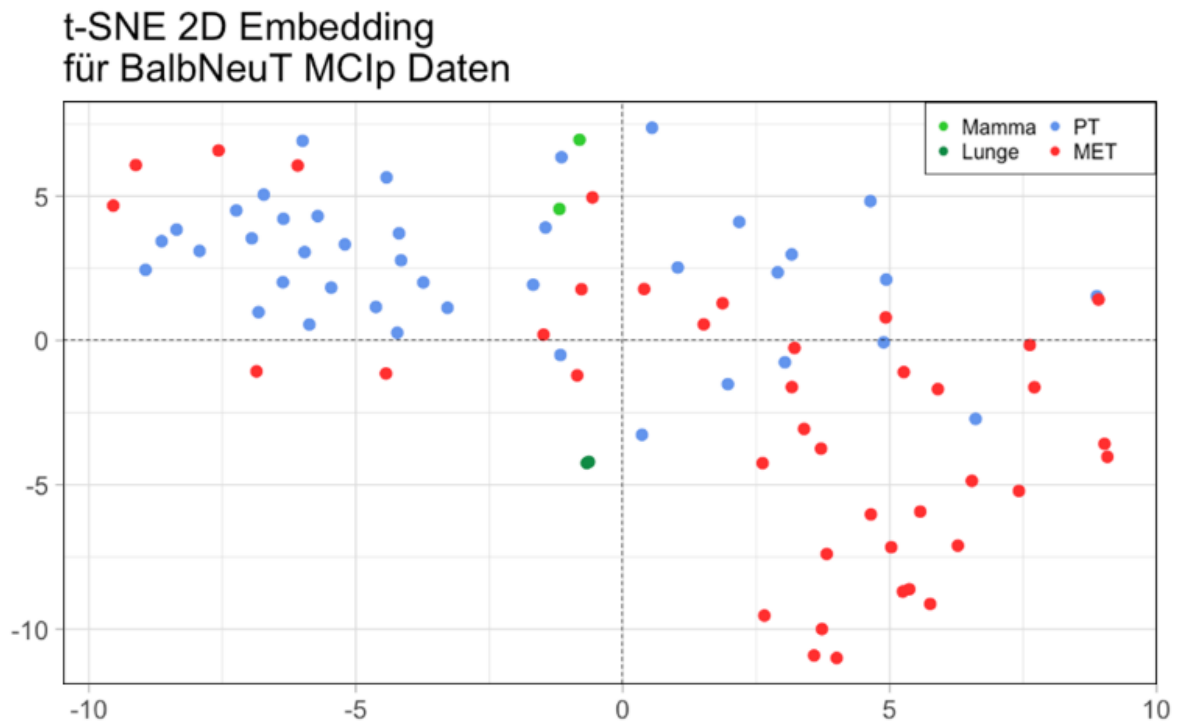


**Abbildung 8: Principal Component Analysis (PCA) Plot der Methylierungsdaten (MCIp).** PC1=Principal Component 1, PC2= Principal Component 2. Abb. 8A: PCA inklusive der Proben aus gesundem Lungengewebe. Abb. 8B: PCA nach Ausschluss der Proben aus gesundem Lungengewebe. Primärtumorproben (PT) blau, Metastasen (MET) rot, gesunde Mammaproben (Mamma) hellgrün, gesunde Lungenproben (Lunge) dunkelgrün.

### 4.3.3 T-distributed Stochastic Neighbor Embedding (t-SNE)

Zuletzt kam die sogenannte *T-distributed Stochastic Neighbor Embedding* (t-SNE) zum Einsatz. Im Gegensatz zu den linearen Methoden von MDS und PCA wird dabei eine nicht-lineare Methode zur Dimensionalitätsreduktion angewandt. Einer t-SNE liegt ein Algorithmus für maschinelles Lernen zugrunde, der die Einbettung eines hochdimensionalen Datensatzes in einen zwei- oder dreidimensionalen Raum erlaubt, ohne dass das ursprüngliche Clustering verloren geht. Dabei wird für jeden Punkt die Wahrscheinlichkeit, dass er neben dem jeweils anderen Punkt liegt, errechnet. Je näher die Punkte in dieser Darstellung beieinander liegen, desto wahrscheinlicher ist es also, dass sie sich ähneln (136).

Auch in dieser Darstellung (Abb. 9) lässt sich eine grobe Aufteilung von Primärtumoren und Metastasen in zwei Clustern feststellen. So liegen die Methylierungsmuster von Metastasen vorwiegend im positiven Bereich der x-Achse und jene aus Primärtumoren überwiegend im negativen Bereich der x-Achse. Es kommt jedoch auch hier nicht zu einer eindeutigen Trennung, denn es finden sich jeweils einige Proben, die in ihrem Methylierungsmuster eher der anderen Gruppe ähneln. Es gibt außerdem einen Bereich zwischen den beiden Clustern, in dem die Primärtumoren und Metastasen nicht klar abgrenzbar sind und in welchen auch die gesunden Kontrollen von Brust- und Lungengewebe fallen. So ist auch mithilfe dieser Methode keine klare Abgrenzung der Proben bezüglich ihrer Methylierungsmuster erkennbar.



**Abbildung 9: T-distributed Stochastic Neighbor Embedding (t-SNE) Plot für die MCIp-Daten.** Primärtumorproben blau, Metastasenproben rot, gesunde Mammaproben hellgrün, gesunde Lungenproben dunkelgrün.

#### 4.4 Auswertung der Regionen

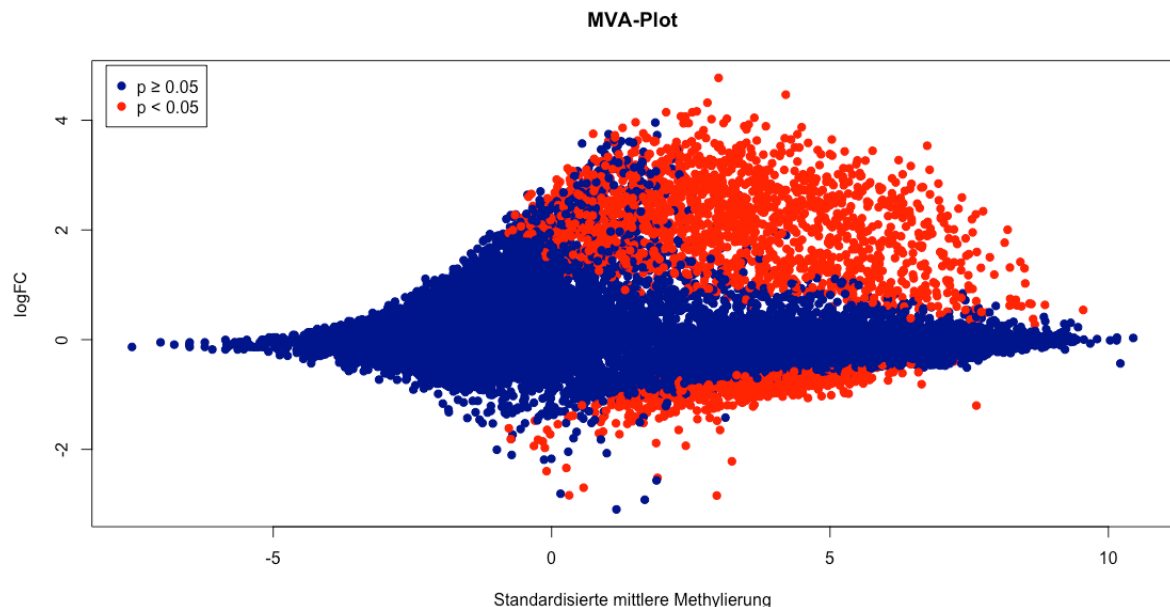
Da sich in den Clusteranalysen andeutete, dass die Primärtumoren sich in Bezug auf ihr Methylierungsmuster von den Metastasen unterscheiden, wurde im nächsten Schritt untersucht, an welchen Regionen im Genom diese Unterschiede besonders deutlich hervortreten. Hierbei ging es vor allem um die Frage, welche und wie viele Regionen eine unterschiedliche Methylierung in Primärtumoren und Metastasen zeigten. Nach einer Qualitätskontrolle der Methylierungsdaten konnten 29.224 Regionen zur Analyse verwendet werden. Von diesen zeigten 1.745 Regionen eine verstärkte Methylierung in den Metastasen und 1.068 Regionen waren in den Metastasen im Vergleich zu Proben der Primärtumoren weniger stark methyliert.

##### 4.4.1 Mean-versus-Average-Plot (MVA-Plot)

Diese Ergebnisse lassen sich sehr gut im sogenannten Mean-versus-Average-Plot (MVA-Plot) darstellen. Dieser zeigt das Verhältnis zwischen der durchschnittlichen Methylierung einer genomischen Region (x-Achse: Standardisierte mittlere Methylierung, berechnet aus dem logarithmierten Mittelwert aller Proben) und dem logarithmierten Fold Change (logFC) der einzelnen Regionen, welcher die

differenzielle Methylierung von Metastasen und Primärtumoren anzeigt (Abb. 10). Ein  $\logFC > 0$  zeigt Regionen, in denen die Metastasen-Proben hypermethyliert sind und ein  $\logFC < 0$  bildet die Regionen ab, die in den Metastasen hypomethyliert sind. Regionen, in denen die Veränderung der Methylierung nach Korrektur für multiples Testen mittels Benjamini-Hochberg-Korrektur signifikant ( $p < 0.05$ ) war, sind rot markiert.

Dabei lässt sich die oben bereits in Zahlen ausgeführte Tatsache beobachten, dass bei Metastasen mehr Regionen hyper- als hypomethyliert sind. Im MVA-Plot sind dementsprechend bei der Mehrzahl der roten Punkte ein  $\logFC > 0$  zu finden. Regionen mit geringer CpG-Dichte (Standardisierte mittlere Methylierung  $< 0$ ) scheinen generell keine signifikanten Methylierungsunterschiede zwischen Primärtumoren und Metastasen aufzuweisen, was sich durch das fehlende Methylierungspotenzial bei geringer CpG-Dichte erklären lässt. Außerdem muss bei der Interpretation dieser Daten beachtet werden, dass Programme wie *DESeq2* darauf ausgelegt sind, stets Änderungen in beide Richtungen nachzuweisen. Da die Hypomethylierung in den Metastasen nur geringer ausgeprägt ist als die Hypermethylierung, kann es sein, dass aufgrund der bioinformatischen Auswertung nur die wenigen Ausreißer der hypomethylierten Regionen die tatsächlich vorhandenen Unterschiede darstellen.

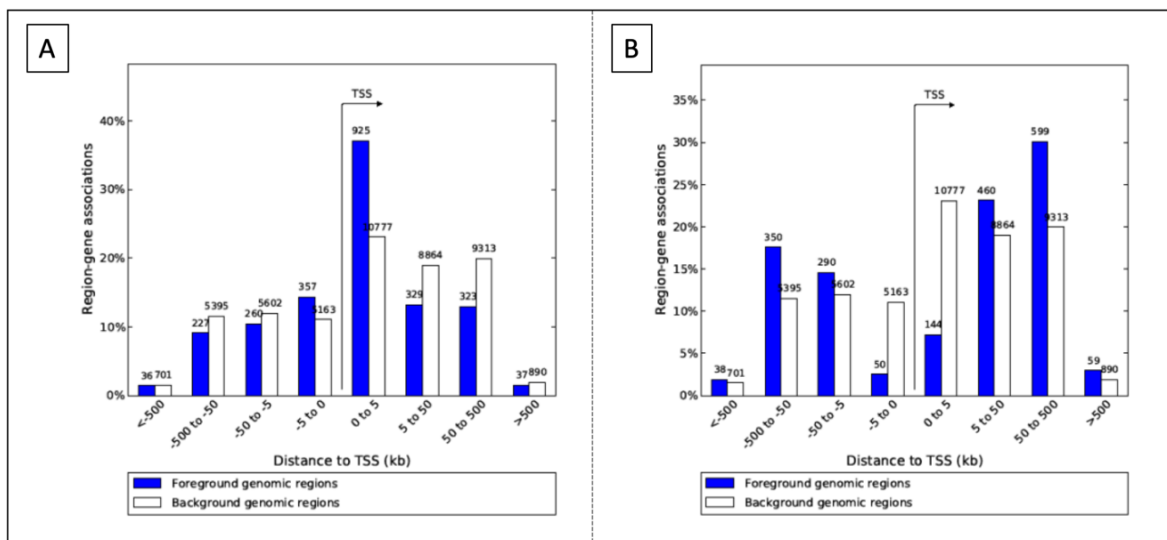


**Abbildung 10: Mean-versus-Average (MVA) Plot für die MChp Daten der BalbNeuT Mäuse für 29.224 Regionen.** Regionen ohne signifikante Veränderung der Methylierung: blau, Regionen mit signifikanter Zu- oder Abnahme der Methylierung: rot (adjustiertes  $p < 0.05$ ).  $\logFC$  = logarithmierter Fold Change.  $\logFC > 0$ : Region in Metastasen hypermethyliert,  $\logFC < 0$ : Region in Metastasen hypomethyliert.

#### 4.4.2 Annotation mittels Genomic Regions Enrichment of Annotations Tool (GREAT)

Die ermittelten Regionen mit signifikanten Methylierungsunterschieden sollten nun definierten Genen und im Anschluss ihren biologischen Funktionen zugeordnet werden. Zur Annotation von nicht-kodierenden Regionen, wie in den hier analysierten Methylierungsdaten, eignet sich das Analyse-Tool *GREAT* (Genomic Regions Enrichment of Annotations Tool), da es nicht nur proximale, sondern auch distale regulatorische Genregionen berücksichtigt (127). Als Hintergrundregionen wurden diejenigen Regionen gewählt, die durch *SssI*-Behandlung der DNA unserer Zellen ermittelt wurden.

In einer ersten Übersichtsgraphik ließ sich die Lage der Regionen im Vergleich zum Transkriptionsstartpunkt der entsprechenden Gene darstellen (s. Abb. 11). Interessanterweise lässt sich im hier analysierten Datensatz feststellen, dass Hypermethylierung in Metastasengewebe häufiger in Promotorregionen stattfindet (Abb. 11A), während Hypomethylierung von Metastasengewebe in den meisten Fällen in den Introns und Exons stattzufinden scheint (Abb. 11B). Dies korreliert mit der Annahme, dass die Genexpression allgemein durch eine vermehrte Methylierung im Promotorbereich oder eine verminderte Methylierung in Introns und Exons verändert wird (137,138).



**Abbildung 11: Abstand in Kilobasen (kb) der signifikant hypermethylierten (A) und hypomethylierten (B) Regionen zum nächsten Transkriptionsstartpunkt (transcription start site, TSS).** Als Hintergrund (Background genomic regions, weiß) wurden die durch *SssI*-Behandlung ermittelten Genregionen verwendet. Die hyper- und hypomethylierten Regionen sind als *Foreground genomic regions* (blau) dargestellt.

#### 4.4.2.1 Hypermethylierte Gene

Aus der Annotation mittels *GREAT* wurde eine Liste der Gene generiert, die mit hypermethylierten Regionen assoziiert sind. Den 1.745 hypermethylierten Regionen der MChp-Experimente wurden 1.912 assoziierte Gene zugeordnet. Von diesen waren 83 Gene bereits bekanntermaßen mit Brustkrebs und/oder Methylierung assoziiert. Tabelle 4-3 zeigt, ob die Gene mit Brustkrebs, Methylierung und/oder Metastasierung assoziiert sind.

Von den 83 hypermethylierten Genen war bei 81 eine Assoziation mit Methylierung im Allgemeinen bekannt, 71 Gene wurden bereits in Zusammenhang mit der Entstehung von Brustkrebs beschrieben und 51 Gene fanden in der Literatur zur Methylierung von Brustkrebs Erwähnung. Zu den 16 folgenden Genen war bereits bekannt, dass die Methylierung eine Rolle bei der Metastasierung von Brustkrebs spielt: *CDKN1C*, *CDKN2A*, *FOXF2*, *HIC1*, *HOXA5*, *HOXA9*, *HOXB13*, *IRF4*, *MGMT*, *NKX2-2*, *PAX6*, *PGR*, *PITX2*, *RUNX3*, *SFRP* und *WT1*. Diese Gene wurden im Folgenden in den 80 hier analysierten Gewebeproben genauer betrachtet. In Abb. 12 ist für jedes dieser 16 Gene die Region mit dem jeweils größten Unterschied zwischen Primärtumoren und Metastasen in Hinblick auf ihre Methylierung dargestellt. Man kann erkennen, dass auch in unserem Mausmodell alle entsprechenden Regionen in den Metastasen deutlich hypermethyliert waren (genaue Ergebnisse siehe Tabelle 4-3).

Hypermethylierung im Gen *CDKN2A* (*cyclin dependent kinase inhibitor 2A*), ist bereits mehrfach als Faktor bei der malignen Progression von Brustkrebs beschrieben worden (139–141) und ist mit Proliferation und Aggressivität von Brustkrebs assoziiert (142,143). *CDKN2A* besteht aus 4 Exons, wobei es mindestens 3 Spleißvarianten gibt. In unserer Studie zeigte sich die Region 4\_1451 in den Metastasen signifikant hypermethyliert im Vergleich zu den Primärtumoren [Mittelwert (Standardabweichung) des  $\text{Log}_2$  Fold Change: 3,06 (0,46); adj.  $P \leq 7,18 \cdot 10^{-09}$ ]. Diese hypermethylierte Region umfasst das gesamte Exon 2.

Tabelle 4-3 *Hypermethylierte Gene*

| Gensymbol      | Genname   | Involviert in |                               |                                |  | logFC (SE)<br>adj. P-Wert <sup>#</sup> |
|----------------|---|---------------|-------------------------------|--------------------------------|--|--|
|                |   | Methylierung  | Entwicklung von<br>Brustkrebs | Methylierung bei<br>Brustkrebs | Methylierung bei<br>der Metastasierung<br>von Brustkrebs |  |
| <b>ALOX5</b>   | Arachidonat 5-Lipoxygenase                                  | Ja (167)      | Ja (36)                       | unbekannt                      | unbekannt  | 2,81 (0,53)<br>9,38*10 <sup>-6</sup>   |
| <b>ATOH1</b>   | atonal bHLH transcription factor 1                          | Ja (9)        | Ja (3)                        | unbekannt                      | unbekannt  | 2,70 (0,37)<br>2,95*10 <sup>-10</sup>  |
| <b>BDNF</b>    | brain derived neurotrophic factor                           | Ja (2046)     | Ja (90)                       | Ja (4)                         | unbekannt  | 2,32 (0,57)<br>0,0016                  |
| <b>CACNA1H</b> | calcium voltage-gated channel<br>subunit alpha1 H           | Ja (13)       | Ja (7)                        | unbekannt                      | unbekannt  | 1,79 (0,57)<br>0,024                   |
| <b>CDKN1C</b>  | cyclin dependent kinase inhibitor<br>1C                     | Ja (199)      | Ja (36)                       | Ja (8)                         | Ja (1)   | 0,80 (0,18)<br>0,0015                  |
| <b>CDKN2A</b>  | cyclin dependent kinase inhibitor<br>2A                     | Ja (2627)     | Ja (618)                      | Ja (181)                       | Ja (22)  | 3,06 (0,46)<br>7,18*10 <sup>-9</sup>   |
| <b>CLIP4</b>   | CAP-Gly domain containing linker<br>protein family member 4 | Ja (9)        | unbekannt                     | unbekannt                      | unbekannt  | 2,49 (0,48)<br>1,69*10 <sup>-5</sup>   |
| <b>DBX1</b>    | developing brain homeobox 1                                 | Ja (2)        | Ja (1)                        | Ja (1)                         | unbekannt  | 2,33 (0,35)<br>1,89*10 <sup>-8</sup>   |
| <b>DDX25</b>   | DEAD-box helicase 25  | unbekannt     | unbekannt                     | unbekannt                      | unbekannt  | 2,23 (0,44)<br>2,57*10 <sup>-5</sup>   |
| <b>DLL4</b>    | delta like canonical Notch ligand 4                         | Ja (10)       | Ja (48)                       | Ja (1)                         | unbekannt  | 1,50 (0,23)<br>5,35*10 <sup>-8</sup>   |
| <b>DMRTA2</b>  | DMRT like family A2   | unbekannt     | unbekannt                     | unbekannt                      | unbekannt  | 2,39 (0,35)<br>4,77*10 <sup>-9</sup>   |
| <b>DOK5</b>    | docking protein 5   | Ja (3)        | Ja (1)                        | unbekannt                      | unbekannt  | 3,13 (0,61)<br>2,15*10 <sup>-5</sup>   |
| <b>EDNRB</b>   | endothelin receptor type B                                  | Ja (222)      | Ja (30)                       | Ja (2)                         | unbekannt  | 2,55 (0,36)<br>1,61*10 <sup>-9</sup>   |
| <b>FBRSL1</b>  | fibrosin like 1   | unbekannt     | unbekannt                     | unbekannt                      | unbekannt  | 2,09 (0,40)<br>1,56*10 <sup>-5</sup>   |
| <b>FGF12</b>   | fibroblast growth factor 12                                 | Ja (9)        | Ja (3)                        | Ja (1)                         | unbekannt  | 2,06 (0,26)<br>3,81*10 <sup>-11</sup>  |

|               |   |          |           |           |           |                                       |
|---------------|---|----------|-----------|-----------|-----------|---------------------------------------|
| <b>FOXA2</b>  | forkhead box A2   | Ja (61)  | Ja (26)   | Ja (5)    | unbekannt | 2,45 (0,35)<br>2,30*10 <sup>-9</sup>  |
| <b>FOXF2</b>  | forkhead box F2   | Ja (7)   | Ja (18)   | Ja (2)    | Ja (1)    | 3,33 (0,60)<br>3,15*10 <sup>-6</sup>  |
| <b>GABRA4</b> | gamma-aminobutyric acid type A receptor subunit alpha4  | Ja (9)   | Ja (1)    | Ja (1)    | unbekannt | 3,82 (0,54)<br>9,02*10 <sup>-10</sup> |
| <b>GAD1</b>   | glutamate decarboxylase 1                               | Ja (52)  | unbekannt | unbekannt | unbekannt | 1,94 (0,47)<br>0,0015                 |
| <b>GATA6</b>  | GATA binding protein 6                                  | Ja (49)  | Ja (7)    | unbekannt | unbekannt | 1,78 (0,22)<br>1,14*10 <sup>-11</sup> |
| <b>GBX2</b>   | gastrulation brain homeobox 2                           | Ja (11)  | Ja (2)    | unbekannt | unbekannt | 2,41 (0,38)<br>1,06*10 <sup>-7</sup>  |
| <b>HAND1</b>  | heart and neural crest derivatives expressed 1          | Ja (32)  | unbekannt | unbekannt | unbekannt | 0,77 (0,19)<br>0,005                  |
| <b>HIC1</b>   | HIC ZBTB transcriptional repressor 1                    | Ja (198) | Ja (43)   | Ja (23)   | Ja (5)    | 3,33 (0,60)<br>3,15*10 <sup>-6</sup>  |
| <b>HLX</b>    | H2.0 like homeobox                                      | Ja (10)  | Ja (1)    | unbekannt | unbekannt | 2,84 (0,51)<br>2,24*10 <sup>-6</sup>  |
| <b>HOXA5</b>  | homeobox A5   | Ja (76)  | Ja (36)   | Ja (12)   | Ja (1)    | 1,36 (0,37)<br>0,0066                 |
| <b>HOXA6</b>  | homeobox A6   | Ja (8)   | Ja (1)    | unbekannt | unbekannt | 1,36 (0,37)<br>0,0066                 |
| <b>HOXA9</b>  | homeobox A9   | Ja (176) | Ja (21)   | Ja (11)   | Ja (2)    | 2,45 (0,28)<br>2,86*10 <sup>-13</sup> |
| <b>HOXB13</b> | homeobox B13  | Ja (26)  | Ja (65)   | Ja (5)    | Ja (1)    | 3,63 (0,53)<br>2,29*10 <sup>-9</sup>  |
| <b>HOXD11</b> | homeobox D11  | Ja (6)   | Ja (2)    | Ja (1)    | unbekannt | 1,50 (0,20)<br>1,70*10 <sup>-10</sup> |
| <b>HOXD3</b>  | homeobox D3   | Ja (24)  | Ja (5)    | unbekannt | unbekannt | 2,85 (0,31)<br>7,60*10 <sup>-15</sup> |
| <b>HOXD9</b>  | homeobox D9   | Ja (9)   | Ja (4)    | Ja (1)    | unbekannt | 2,24 (0,28)<br>9,85*10 <sup>-12</sup> |
| <b>INA</b>    | internexin neuronal intermediate filament protein alpha | Ja (4)   | unbekannt | unbekannt | unbekannt | 2,84 (0,38)<br>9,93*10 <sup>-11</sup> |
| <b>IRF4</b>   | interferon regulatory factor 4                          | Ja (48)  | Ja (13)   | Ja (5)    | Ja (1)    | 3,07 (0,77)<br>0,0018                 |

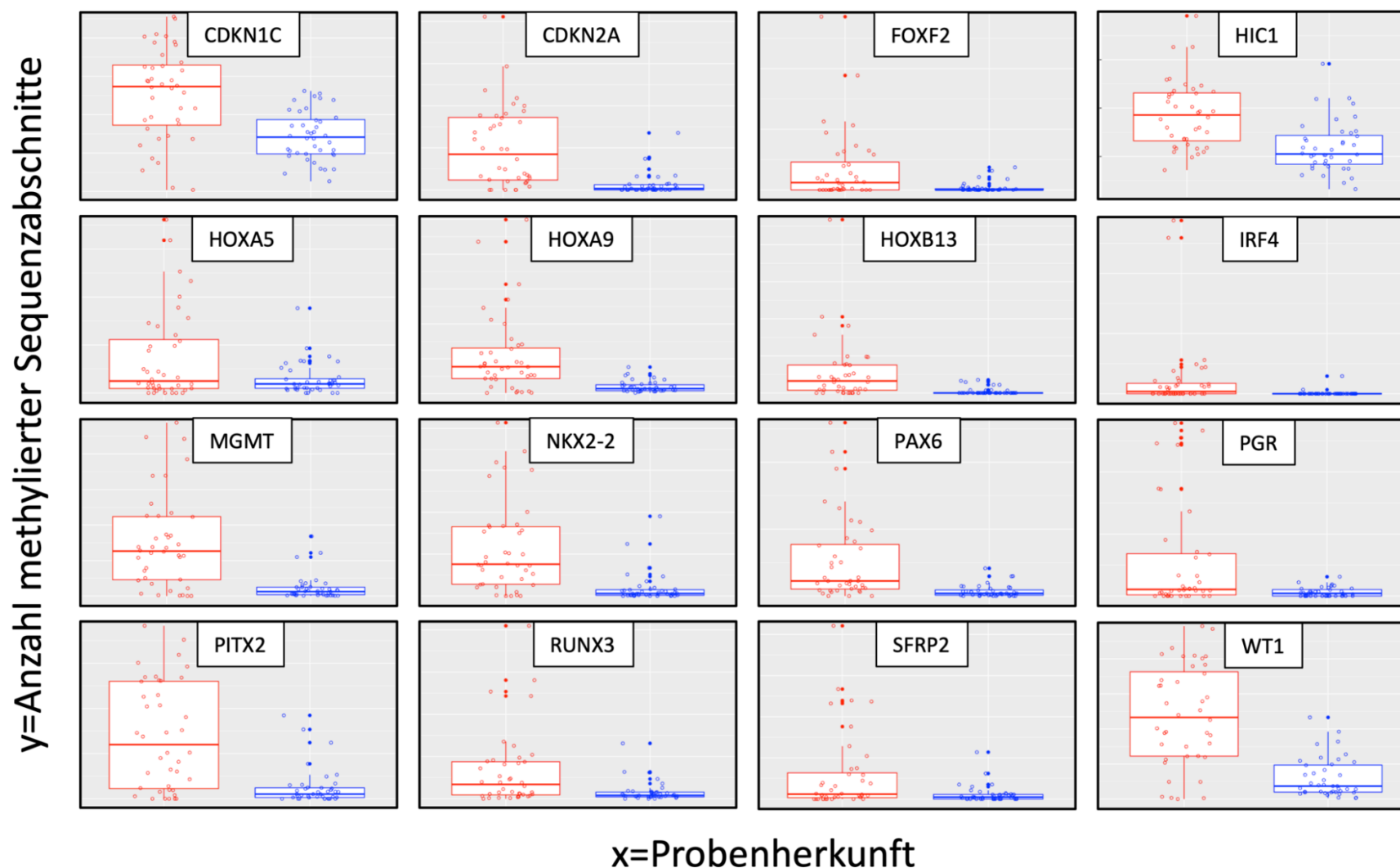


|                |  |           |           |           |           |                                       |
|----------------|--|-----------|-----------|-----------|-----------|---------------------------------------|
| <b>ITGA4</b>   | integrin subunit alpha 4                         | Ja (26)   | Ja (4)    | Ja (1)    | unbekannt | 2,65 (0,48)<br>4,32*10 <sup>-6</sup>  |
| <b>LHX2</b>    | LIM homeobox 2                                   | Ja (19)   | Ja (8)    | Ja (5)    | unbekannt | 3,87 (0,44)<br>4,17*10 <sup>-14</sup> |
| <b>MGMT</b>    | O-6-methylguanine-DNA<br>methyltransferase       | Ja (2930) | Ja (125)  | Ja (73)   | Ja (8)    | 2,75 (0,38)<br>3,88*10 <sup>-10</sup> |
| <b>MNX1</b>    | motor neuron and pancreas<br>homeobox 1          | Ja (5)    | Ja (5)    | unbekannt | unbekannt | 3,36 (0,54)<br>9,54*10 <sup>-8</sup>  |
| <b>MSC</b>     | musculin/activated B-cell factor 1<br>(ABF-1)    | Ja (3)    | unbekannt | unbekannt | unbekannt | 3,05 (0,60)<br>2,65*10 <sup>-5</sup>  |
| <b>NEFM</b>    | neurofilament medium                             | Ja (11)   | Ja (3)    | Ja (1)    | unbekannt | 2,34 (0,34)<br>3,49*10 <sup>-9</sup>  |
| <b>NEUROD1</b> | neuronal differentiation 1                       | Ja (37)   | Ja (7)    | Ja (5)    | unbekannt | 2,67 (0,51)<br>1,22*10 <sup>-5</sup>  |
| <b>NID2</b>    | nidogen 2  | Ja (32)   | Ja (2)    | Ja (1)    | unbekannt | 1,80 (0,42)<br>0,0010                 |
| <b>NKX2-1</b>  | NK2 homeobox 1                                   | Ja (21)   | Ja (3)    | unbekannt | unbekannt | 2,77 (0,33)<br>1,14*10 <sup>-12</sup> |
| <b>NKX2-2</b>  | NK2 homeobox 2                                   | Ja (17)   | Ja (2)    | Ja (2)    | Ja (1)    | 2,62 (0,40)<br>2,51*10 <sup>-8</sup>  |
| <b>NPY</b>     | neuropeptide Y                                   | Ja (315)  | Ja (56)   | Ja (4)    | unbekannt | 3,10 (0,79)<br>0,0022                 |
| <b>NR2E1</b>   | nuclear receptor subfamily 2<br>group E member 1 | Ja (6)    | Ja (5)    | unbekannt | unbekannt | 3,20 (0,41)<br>2,07*10 <sup>-11</sup> |
| <b>NR2F2</b>   | nuclear receptor subfamily 2<br>group F member 2 | Ja (22)   | Ja (13)   | Ja (2)    | unbekannt | 1,88 (0,34)<br>4,05*10 <sup>-6</sup>  |
| <b>NR4A3</b>   | nuclear receptor subfamily 4<br>group A member 3 | Ja (13)   | Ja (13)   | unbekannt | unbekannt | 0,91 (0,18)<br>1,4*10 <sup>-4</sup>   |
| <b>NRN1</b>    | neuritin 1                                       | Ja (9)    | Ja (2)    | Ja (1)    | unbekannt | 1,85 (0,32)<br>1,22*10 <sup>-6</sup>  |
| <b>NXPH1</b>   | neurexophilin 1                                  | Ja (5)    | Ja (19)   | Ja (1)    | unbekannt | 2,36 (0,41)<br>1,26*10 <sup>-6</sup>  |
| <b>OSR1</b>    | odd-skipped related transcription<br>factor 1    | Ja (13)   | Ja (4)    | unbekannt | unbekannt | 2,78 (0,65)<br>6,2*10 <sup>-4</sup>   |
| <b>OTX2</b>    | orthodenticle homeobox 2                         | Ja (35)   | Ja (1)    | Ja (1)    | unbekannt | 2,89 (0,40)<br>5,39*10 <sup>-10</sup> |

|                |  |          |           |           |           |                                      |
|----------------|--|----------|-----------|-----------|-----------|--------------------------------------|
| <b>PAX6</b>    | paired box 6   | Ja (121) | Ja (23)   | Ja (13)   | Ja (3)    | 2,84 (0,38)<br>$1,16 \cdot 10^{-10}$ |
| <b>PAX7</b>    | paired box 7   | Ja (45)  | Ja (7)    | unbekannt | unbekannt | 3,14 (0,44)<br>$1,03 \cdot 10^{-9}$  |
| <b>PAX9</b>    | paired box 9   | Ja (12)  | Ja (5)    | unbekannt | unbekannt | 2,77 (0,33)<br>$1,14 \cdot 10^{-12}$ |
| <b>PCDH8</b>   | protocadherin 8  | Ja (20)  | Ja (2)    | Ja (2)    | unbekannt | 1,85 (0,29)<br>$1,14 \cdot 10^{-7}$  |
| <b>PCDHGC5</b> | protocadherin gamma subfamily C, 5                         | Ja (1)   | unbekannt | unbekannt | unbekannt | 1,43 (0,44)<br>0,023                 |
| <b>PDGFRA</b>  | platelet derived growth factor receptor alpha              | Ja (98)  | Ja (41)   | Ja (2)    | unbekannt | 2,05 (0,45)<br>$3,07 \cdot 10^{-4}$  |
| <b>PGR</b>     | progesterone receptor                                      | Ja (108) | Ja (2127) | Ja (47)   | Ja (4)    | 3,19 (0,50)<br>$2,89 \cdot 10^{-8}$  |
| <b>PHOX2A</b>  | paired like homeobox 2A                                    | Ja (5)   | unbekannt | unbekannt | unbekannt | 1,05 (0,24)<br>$8,25 \cdot 10^{-4}$  |
| <b>PHOX2B</b>  | paired like homeobox 2B                                    | Ja (16)  | unbekannt | unbekannt | unbekannt | 1,67 (0,25)<br>$2,76 \cdot 10^{-8}$  |
| <b>PITX2</b>   | paired like homeodomain 2                                  | Ja (82)  | Ja (34)   | Ja (27)   | Ja (7)    | 2,51 (0,41)<br>$2,55 \cdot 10^{-7}$  |
| <b>POMC</b>    | proopiomelanocortin  | Ja (914) | Ja (212)  | Ja (3)    | unbekannt | 1,46 (0,29)<br>$7,66 \cdot 10^{-5}$  |
| <b>POU4F1</b>  | POU class 4 homeobox 1                                     | Ja (5)   | Ja (4)    | Ja (1)    | unbekannt | 2,55 (0,36)<br>$1,61 \cdot 10^{-9}$  |
| <b>POU4F2</b>  | POU class 4 homeobox 2                                     | Ja (13)  | Ja (12)   | Ja (2)    | unbekannt | 1,99 (0,34)<br>$1,24 \cdot 10^{-6}$  |
| <b>POU4F3</b>  | POU class 4 homeobox 3                                     | Ja (7)   | Ja (1)    | unbekannt | unbekannt | 2,44 (0,37)<br>$2,12 \cdot 10^{-8}$  |
| <b>PRKCB</b>   | protein kinase C beta                                      | Ja (16)  | Ja (5)    | Ja (1)    | unbekannt | 2,16 (0,32)<br>$8,09 \cdot 10^{-9}$  |
| <b>PUS3</b>    | pseudouridine synthase 3                                   | Ja (4)   | Ja (1)    | unbekannt | unbekannt | 2,23 (0,44)<br>$2,57 \cdot 10^{-5}$  |
| <b>RASGRF2</b> | Ras protein specific guanine nucleotide releasing factor 2 | Ja (14)  | Ja (5)    | Ja (1)    | unbekannt | 3,34 (0,47)<br>$4,80 \cdot 10^{-10}$ |
| <b>RASL10A</b> | RAS like family 10 member A                                | Ja (3)   | unbekannt | unbekannt | unbekannt | 0,80 (0,16)<br>$1,61 \cdot 10^{-4}$  |

|              |   |          |           |           |           |                                       |
|--------------|---|----------|-----------|-----------|-----------|---------------------------------------|
| <b>RUNX3</b> | RUNX family transcription factor 3          | Ja (470) | Ja (73)   | Ja (35)   | Ja (5)    | 2,02 (0,36)<br>4,42*10 <sup>-6</sup>  |
| <b>SALL1</b> | spalt like transcription factor 1           | Ja (14)  | Ja (6)    | Ja (2)    | unbekannt | 2,13 (0,37)<br>1,58*10 <sup>-6</sup>  |
| <b>SALL3</b> | spalt like transcription factor 3           | Ja (25)  | Ja (3)    | Ja (2)    | unbekannt | 2,34 (0,35)<br>1,56*10 <sup>-8</sup>  |
| <b>SFRP2</b> | secreted frizzled related protein 2         | Ja (244) | Ja (37)   | Ja (14)   | Ja (1)    | 2,43 (0,51)<br>9,39*10 <sup>-5</sup>  |
| <b>SIM1</b>  | SIM bHLH transcription factor 1             | Ja (16)  | Ja (4)    | Ja (3)    | unbekannt | 2,65 (0,37)<br>6,61*10 <sup>-10</sup> |
| <b>SIX1</b>  | SIX homeobox 1                              | Ja (19)  | Ja (45)   | unbekannt | unbekannt | 3,27 (0,56)<br>5,82*10 <sup>-7</sup>  |
| <b>SIX6</b>  | SIX homeobox 6                              | Ja (4)   | Ja (3)    | Ja (1)    | unbekannt | 2,42 (0,34)<br>1,53*10 <sup>-9</sup>  |
| <b>SSTR1</b> | somatostatin receptor 1                     | Ja (18)  | Ja (18)   | unbekannt | unbekannt | 2,72 (0,33)<br>2,91*10 <sup>-12</sup> |
| <b>TBR1</b>  | T-box brain transcription factor 1          | Ja (22)  | unbekannt | unbekannt | unbekannt | 3,35 (0,43)<br>1,23*10 <sup>-11</sup> |
| <b>TCF3</b>  | transcription factor 3                      | Ja (42)  | Ja (17)   | Ja (1)    | unbekannt | 2,43 (0,45)<br>5,66*10 <sup>-6</sup>  |
| <b>TLX1</b>  | T cell leukemia homeobox 1                  | Ja (15)  | Ja (2)    | Ja (1)    | unbekannt | 2,02 (0,47)<br>8,5*10 <sup>-4</sup>   |
| <b>TOX2</b>  | TOX high mobility group box family member 2 | Ja (4)   | Ja (1)    | Ja (1)    | unbekannt | 0,71 (0,21)<br>0,034                  |
| <b>WRN</b>   | WRN RecQ like helicase                      | Ja (66)  | Ja (51)   | Ja (4)    | unbekannt | 3,03 (0,40)<br>5,88*10 <sup>-11</sup> |
| <b>WT1</b>   | WT1 transcription factor                    | Ja (252) | Ja (221)  | Ja (22)   | Ja (1)    | 2,01 (0,29)<br>6,61*10 <sup>-9</sup>  |

Ergebnisse der in dieser Studie hypermethylierten Regionen in Genen, welche bereits mit Brustkrebs, Methylierung und/oder Metastasierung assoziiert waren. In Klammern steht die Anzahl der Treffer bei der Pubmed-Suche nach dem Gen und dem entsprechenden Suchbegriff. # = Statistische Werte der Regionen mit stärkstem Unterschied zwischen Primärtumor und Metastase innerhalb des entsprechenden Gens. SE = Standardfehler (*Standard Error*). *P*-Werte adjustiert für multiples Testen mittels Benjamini-Hochberg-Korrektur.



**Abbildung 12: Anzahl methylierter Sequenzabschnitte in den nach Literaturrecherche identifizierten 16 Genen.** Für jedes Gen ist die Region mit dem stärksten Unterschied in ihrer Methylierung zwischen Metastasen (rot) und Primärtumoren (blau) dargestellt (Median und Interquartilsabstand als Boxplots, leere Kreise: einzelne Proben)

#### 4.4.2.2 Hypomethylierte Gene

Bei den hypomethylierten Regionen wurden diejenigen extrahiert, welche signifikante Unterschiede aufwiesen und bei denen der logarithmierte Fold Change  $> 1$  lag. Anschließend wurde mittels *GREAT* eine Liste der assoziierten Gene generiert. Hierbei wurden den verbliebenen 173 Regionen 280 Gene zugeordnet. Von diesen waren 2 Gene bereits bekanntermaßen mit Brustkrebs und/oder Methylierung assoziiert. Tabelle 4-4 zeigt, ob die Gene mit Brustkrebs, Methylierung und/oder Metastasierung assoziiert sind.

**Tabelle 4-4 Hypomethylierte Gene**

| <i>Gensymbol</i> | <i>Genname</i>   | <i>Involviert in</i> |                                       |  |   | <i>logFC (SE)<br/>P-Wert<sup>#</sup></i> |
|------------------|--|----------------------|---------------------------------------|--|---|--|
|                  |  | <i>Methylierung</i>  | <i>Entwicklung von<br/>Brustkrebs</i> | <i>Methylierung bei<br/>Brustkrebs</i> | <i>Methylierung bei<br/>der Metastasierung<br/>von Brustkrebs</i> |  |
| <b>PBX1</b>      | PBX homeobox 1   | Ja (23)              | Ja (24)                               | unbekannt                              | unbekannt   | 1,67 (0,41)<br>0,0019                    |
| <b>RAD23B</b>    | RAD23 homolog B, nucleotide<br>excision repair protein | Ja (4)               | Ja (9)                                | unbekannt                              | unbekannt   | 2,70 (0,54)<br>3,26*10 <sup>-05</sup>    |

Ergebnisse der in dieser Studie hypomethylierten Regionen in Genen, welche bereits mit Brustkrebs, Methylierung und/oder Metastasierung assoziiert waren. In Klammern steht die Anzahl der Treffer bei der Pubmed-Suche nach dem Gen und dem entsprechenden Suchbegriff. # = Statistische Werte der Regionen mit stärkstem Unterschied zwischen Primärtumor und Metastase innerhalb des entsprechenden Gens. SE = Standardfehler (*Standard Error*). *P*-Werte adjustiert für multiples Testen mittels Benjamini-Hochberg-Korrektur.

### 4.4.3 Darstellung der assoziierten Gene mittels Metascape

Um festzustellen, welche übergeordneten Funktionen die ermittelten Gene erfüllen, wurde die Plattform Metascape verwendet, welche ein Meta-Analyse-Werkzeug zur Analyse von Pathway Enrichment ist.

#### 4.4.3.1 Regionen mit signifikanter Hypermethylierung

Den 1.745 Regionen, die im hier generierten Datensatz hypermethyliert waren, konnten 1.912 Gene zugeordnet werden. Einen Überblick über die angereicherten Funktionen dieser Gene liefert das Balkendiagramm in Abb. 13.

Besonders auffällig ist, dass es sich hauptsächlich um Funktionen handelt, die für die Differenzierung der Zelle (*pattern specification process, cell fate commitment, cell fate determination, epithelial cell differentiation*), sowie für die Embryonal- und Organentwicklung, z.B. des zentralen Nervensystems, des endokrinen und des Verdauungssystems wichtig sind.

Um die Beziehung zwischen den Funktionsbezeichnungen noch besser darstellen zu können, wurde eine Untergruppe der angereicherten Funktionen ausgewählt und als Netzwerk-Plot abgebildet. Ontologische Funktionsbezeichnungen, die ähnliche Gene enthalten, werden hierbei in Kategorien gruppiert (die entsprechenden Knoten des Netzwerk-Plots sind in gleicher Farbe abgebildet). Funktionsbezeichnungen ab einem Kappa Ähnlichkeits-Score  $> 0,3$  werden mit Kanten verbunden. Abb. 14A zeigt die Farbkodierung in Kategorien, Abb. 14B zeigt die entsprechenden *P*-Werte.

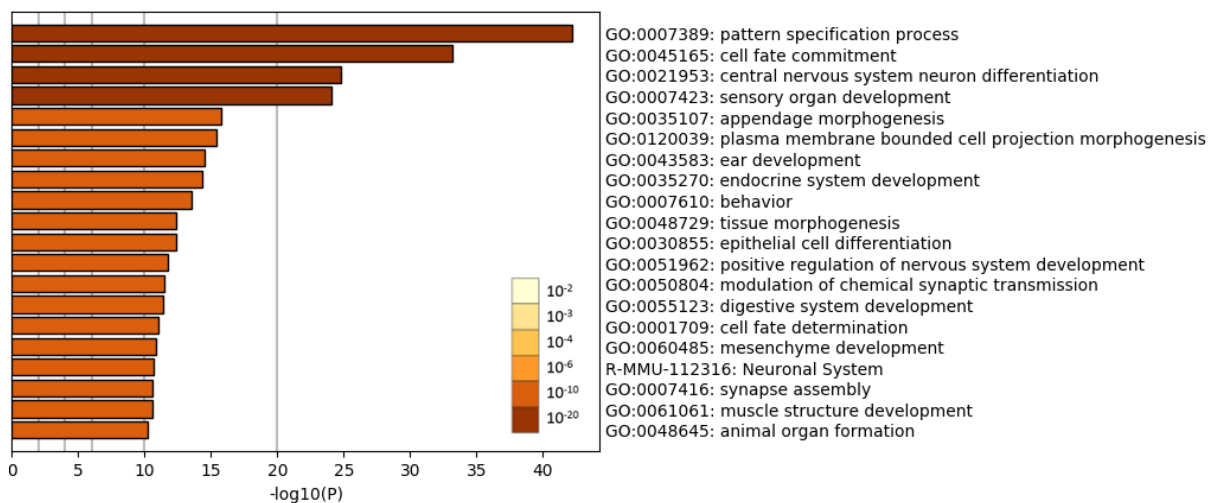
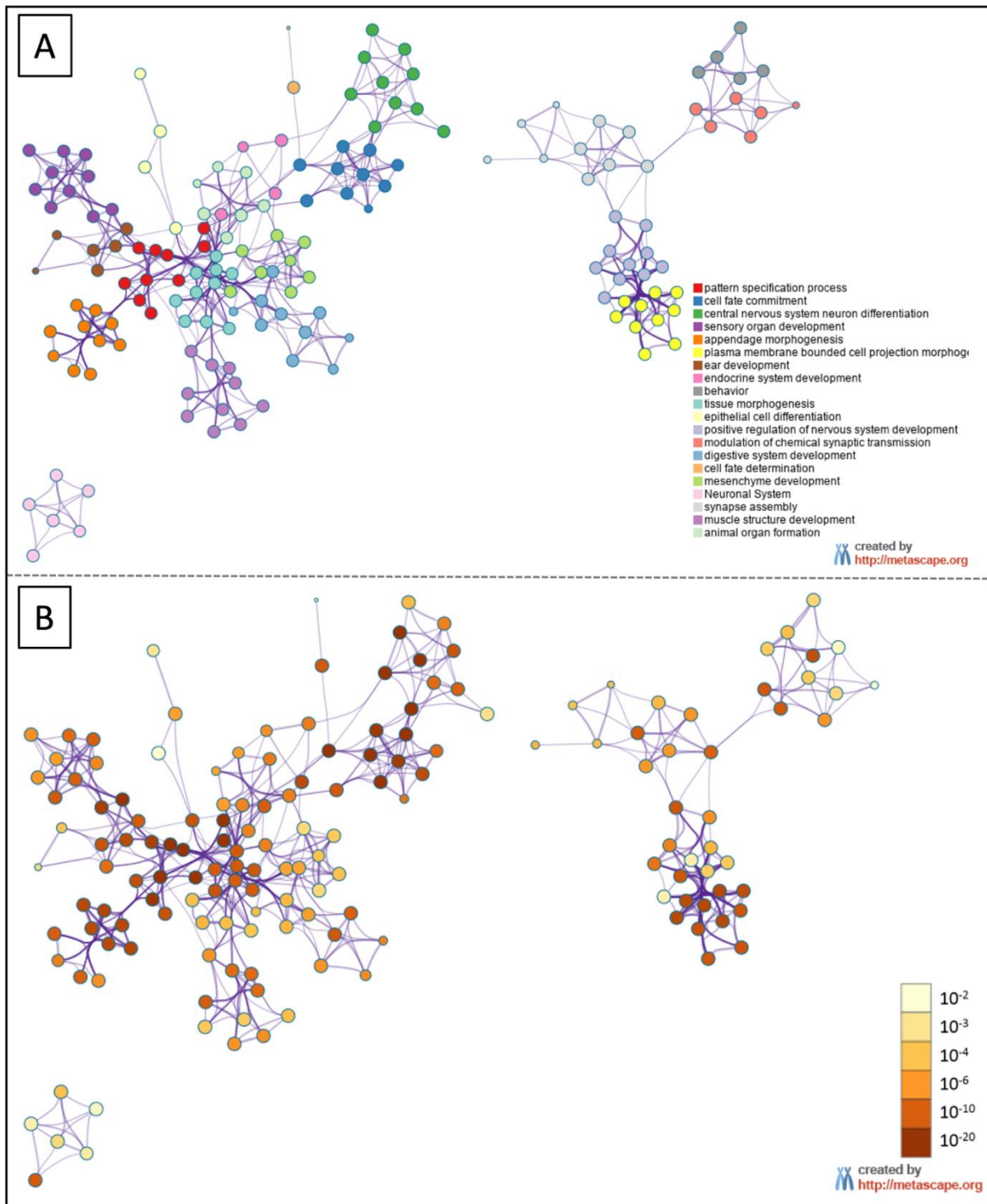


Abbildung 13: Balkendiagramm der Funktionsgruppen, die signifikant in Genen angereichert sind, die mit hypermethylierten Regionen assoziiert sind. Farbkodierung nach *P*-Wert.



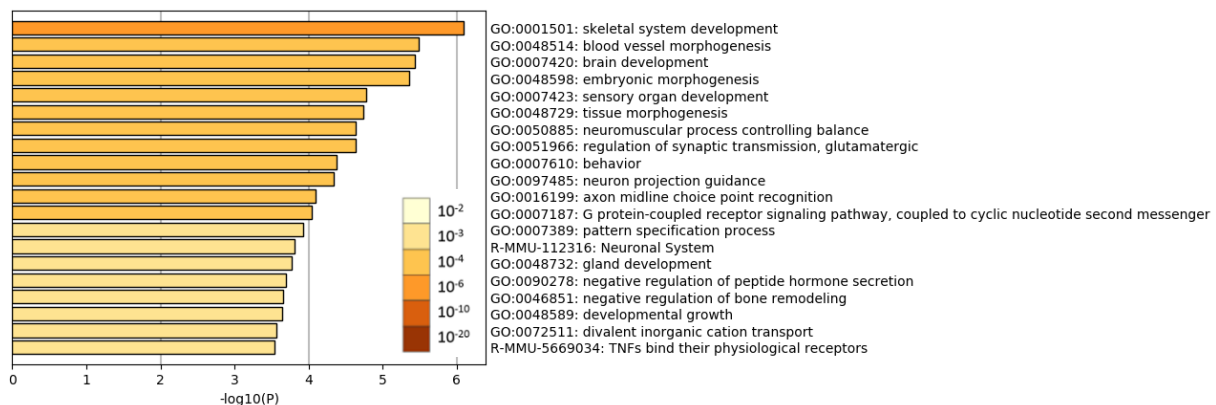
**Abbildung 14: Netzwerk-Plot der Funktionsgruppen, die signifikant in Genen angereichert sind, die mit hypermethylierten Regionen assoziiert sind.** Farbkodierung nach Kategorie (A), in der Knoten, die in dieselbe Kategorie fallen, näher beieinander liegen und nach *P*-Wert (B), bei der Funktionsbezeichnungen, die mehr Gene enthalten, tendenziell einen signifikanteren *P*-Wert haben.



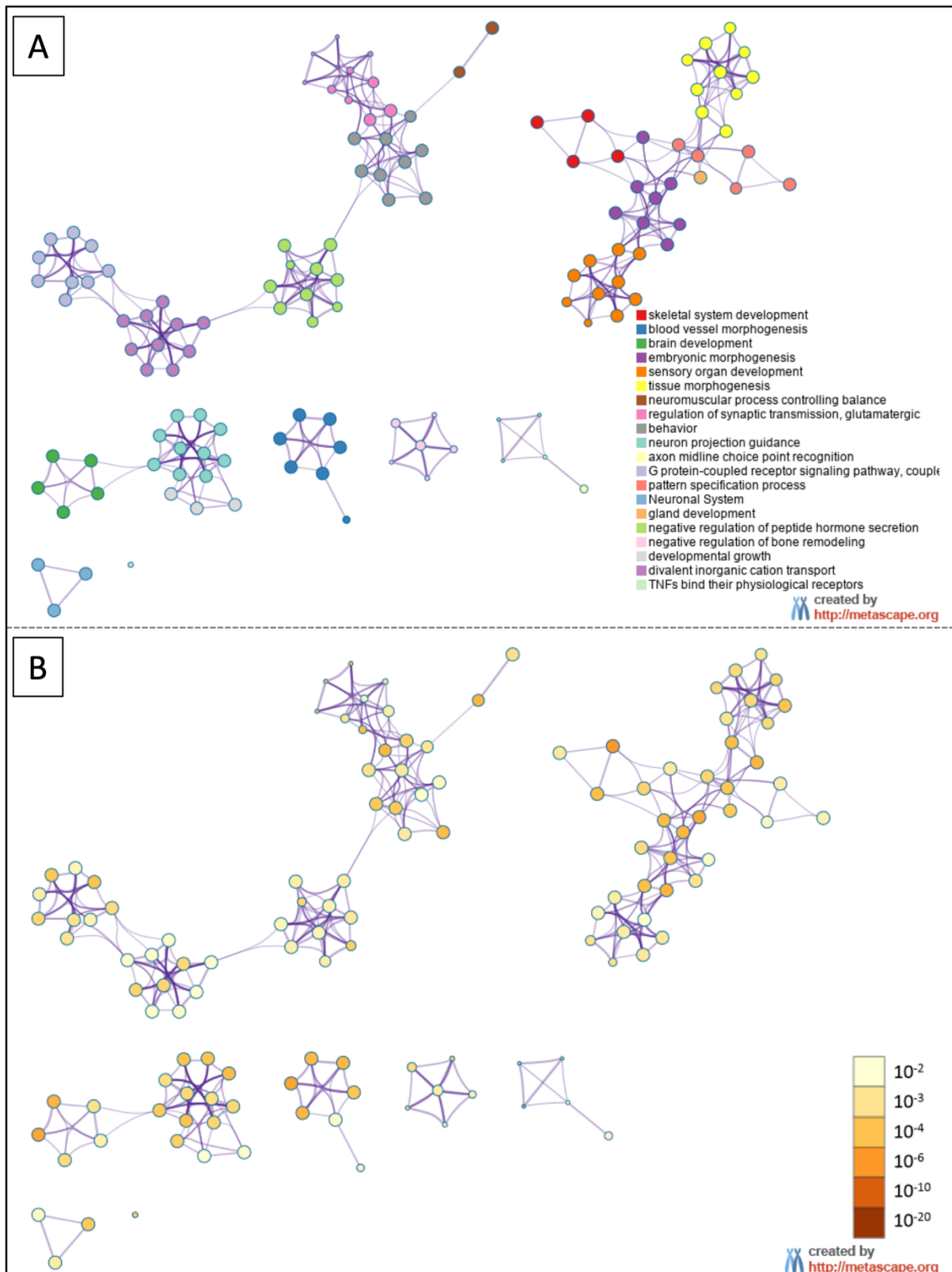
#### 4.4.3.2 Regionen mit signifikanter Hypomethylierung

Den 1.068 Regionen, die in unserem Datensatz hypomethyliert waren, konnten 1.527 Gene zugeordnet werden. Das Balkendiagramm der in Metascape angereicherten Funktionsbezeichnungen der hypomethylierten Regionen (Abb. 15) zeigt deutlich weniger signifikante *P*-Werte im Vergleich zu den hypermethylierten Regionen (vgl. Abb. 13). Ähnlich wie bei den hypermethylierten Regionen sind die mit den hypomethylierten Regionen assoziierte Gene auch solche, die für embryologische Prozesse und Entwicklung zuständig sind. Hier spielen zum Beispiel die Entwicklung des Skelettsystems, des Gehirns und die Entstehung von Drüsen eine Rolle, sowie die Morphogenese von Blutgefäßen und Bindegewebe.

Auch die Funktionskategorien der mit den hypomethylierten Regionen assoziierten Gene lassen sich in einem Netzwerk-Plot darstellen (s. Abb. 16). Im Gegensatz zu Abb. 14 fällt dabei auf, dass mehrere Funktionsbezeichnungen vereinzelte Cluster bilden, was bedeutet, dass die ontologischen Kategorien weniger Gemeinsamkeiten haben.



**Abbildung 15: Balkendiagramm der Funktionsgruppen, die signifikant in Genen angereichert sind, die mit hypermethylierten Regionen assoziiert sind. Farbkodierung nach *P*-Wert.**



**Abbildung 16: Netzwerk-Plot der Funktionsgruppen, die signifikant in Genen angereichert sind, die mit hypomethylierten Regionen assoziiert sind.** Farbkodierung nach Kategorie (A), in der Knoten, die in dieselbe Kategorie fallen, näher beieinander liegen und nach *P*-Wert (B), bei der Funktionsbezeichnungen, die mehr Gene enthalten, tendenziell einen signifikanteren *P*-Wert haben.

## 5 Diskussion

### 5.1 Bewertung der Arbeit

Diese Arbeit beleuchtet die epigenetischen Mechanismen der malignen Progression von Brustkrebs. Wir konnten dabei zeigen, dass sich die DNA-Methylierung zwischen Primärtumoren und Metastasen zwar unterscheidet, sich jedoch keine ganz eindeutigen Cluster an Methylierungsmustern bilden lassen. Dies könnte darauf hindeuten, dass es zu unterschiedlichen Stadien der Tumorprogression zu epigenetischen Veränderungen kommen kann, die mit Metastasierung assoziiert sind. Außerdem konnten wir einige bereits bekannte Gene identifizieren, die bei Metastasen hyper- oder hypomethyliert waren, und vor allem bisher noch unbekannte Regionen aufzeigen, die als neue Zielstrukturen für diagnostische, therapeutische oder prognostische Marker dienen könnten.

Beim Vergleich der Methylierungsmuster zwischen Primärtumoren und Metastasen zeigte sich ein nicht ganz eindeutiges und doch reproduzierbares Bild. So konnten wir mittels PCA, MDS und t-SNE zeigen, dass Metastasen und Primärtumoren tendenziell unterschiedliche Methylierungsmuster aufweisen, eine eindeutige Differenzierung der Proben anhand der epigenetischen Signatur allerdings statistisch nicht möglich ist. Dies wäre gut mit einer frühzeitigen Metastasierung vereinbar. So ist es möglich, dass diejenigen Metastasen, die schon sehr früh während des Tumorwachstums streuen, eine deutlich unterschiedliche Signatur des Epigenoms haben, da sie länger in einem anderen Milieu gewachsen sind. Diejenigen Metastasen, die sich erst relativ spät gebildet haben, dürften dann ein dem Primärtumor noch ähnlicheres Methylierungsmuster aufweisen. Der Zeitpunkt der Metastasierung wurde in unserer Studie allerdings nicht untersucht und daher müsste diese Hypothese in weiteren Studien überprüft werden.

Unsere Arbeit zeigt in einer umfassenden Art und Weise die Veränderung der genomweiten Methylierungsmuster von Primärtumoren und Metastasen bei Brustkrebs in einem Mausmodell und hat einige Stärken: (1) In einem Mausmodell sind die Bedingungen, unter denen die Experimente stattfinden, kontrolliert und bei allen Versuchstieren sehr ähnlich. Außerdem kann durch dieses Modell der konfundierende

Faktor der genetischen Grundausstattung vermindert werden. Zudem konnte durch die Transplantation der noch nicht entarteten Brustdrüse einer BalbNeuT-Maus auf eine BALB/c-Maus die normale Entstehung von Brustkrebs nachgeahmt werden. (2) Diese Arbeit wird durch die Zusammenarbeit mit der Forschergruppe FOR2127 in ein umfassenderes Projekt integriert. Nach dieser Arbeit fortgeführte Projekte werden sich so zum Beispiel der mRNA-Synthese oder den Histonmodifikationen widmen. Eine solch umfangreiche Begutachtung genetischer und epigenetischer Mechanismen erlaubt ein tieferes pathophysiologisches Verständnis der Metastasierung von Brustkrebs. (3) Wir haben nicht nur Tumor- und Metastasengewebe, sondern auch gesundes Mamma- und Lungengewebe als Kontrollen analysiert. Im Gegensatz zur genetischen Information unterscheidet sich die epigenetische Ausstattung von Zellen deutlich zwischen verschiedenen Zelltypen (144). Die Auswertungen des gesunden Kontrollgewebes sind also nötig, um valide Rückschlüsse über die zugrundeliegenden pathophysiologischen Prozesse ziehen zu können.

Unsere Studie hat allerdings auch verschiedene Limitationen: (1) In einem Mausmodell lassen sich nur bedingt Rückschlüsse über die Metastasierung von Brustkrebs beim Menschen ziehen. Epigenetische Veränderungen in einem Tiermodell zu untersuchen ist aber unerlässlich zum pathophysiologischen Verständnis der Metastasierung. Die gefundenen Ergebnisse müssen anschließend in einem translationalen Ansatz im Menschen bestätigt werden. Das BalbNeuT-Mausmodell beruht auf dem mutierten Ratten-HER2/neu Onkogen. Da sich dieses Onkogen jedoch nur bei etwa einem Fünftel der Patientinnen mit Brustkrebs findet, ist es denkbar, dass bei nicht HER2/neu-positivem Brustkrebs andere Mechanismen eine Rolle spielen (145). Allerdings ist das HER2/neu Onkogen einer der pathogenetisch und klinisch bedeutsamsten Marker beim Mammakarzinom, weshalb die vorliegende Arbeit trotzdem eine ausgesprochen hohe klinische Relevanz hat. (2) Des Weiteren besteht die Möglichkeit, dass sich in die genomische DNA der gesammelten Proben DNA von anderen Zelltypen gemischt hat. Besonders bei den sehr kleinen Lungenmetastasen ist eine Verunreinigung durch normales Lungengewebe durchaus denkbar. Unsere Clusterplots zeigen jedoch eine klare Abgrenzung der Proben von normalem Lungengewebe (Negativkontrollen) zu allen anderen Proben (Primärtumoren und Metastasen). Daher kann man, basierend auf diesen statistischen Auswertungen, davon ausgehen, dass die Kontamination nur gering sein sollte. (3) Bei einer niedrigen

CpG-Dichte in einer bestimmten Region können relevante Unterschiede statistisch nur schwierig abgebildet werden. Das heißt, dass es potenziell noch weitere Regionen gibt, in denen sich die Methylierung zwischen Primärtumoren und Metastasen unterscheiden, die wir mit unserer limitierten Anzahl an Proben nicht finden konnten.

Reversible und variable Veränderungen der zellulären Mechanismen spielen eine zentrale Rolle bei der Metastasierung: Die aktive DNA-Demethylierung scheint bei der Entstehung von Metastasen eine Rolle zu spielen (146). Es konnte außerdem gezeigt werden, dass beim Brustkrebs Ähnlichkeiten zwischen gepaarten Primärtumoren und Metastasen bezüglich ihres Methylierungsmusters größer waren als zwischen verschiedenen Metastasen (97,147) und bei Lungenmetastasen gibt es Hinweise, dass die Veränderung der Mikroumgebung einen besonders aggressiven Phänotyp abschwächen kann (148). Die epigenetischen Prozesse, die eine Metastasierung von Brustkrebszellen begünstigen, sind allerdings noch weitestgehend unklar.

Wir konnten in unserer Studie einige Ergebnisse zur bereits bekannten Methylierung von metastasierten Brustkrebszellen replizieren, was die Validität unserer Ergebnisse unterstreicht. So waren zum Beispiel *CDKN2A*, *FOXF2*, *HIC1*, *HOXA5* und *PAX6* in Metastasen im Vergleich zu Primärtumoren hypermethyliert. Die Einordnung dieser exemplarischen Ergebnisse soll im Folgenden dargestellt werden: (1) Für das Gen *CDKN2A* ist bereits eine Hypermethylierung in Brustkrebsgewebe bekannt (139–141) und es ist mit Proliferation und Aggressivität von Brustkrebs assoziiert (142,143). Die Inhibierung des Produkts von *CDKN2A*, P16, vermittelt die Verminderung von Wachstum und Metastasierung bei Brustkrebszellen (143). *CDKN2A* gehört auch zu einem aus sieben Genen bestehenden CpG-Insel-Methylierungspanels, welches das Risiko eines Rückfalls bei Brustkrebs quantifizieren möchte (149). (2) Eine verminderte Expression des Transkriptionsfaktors *FOXF2* ist mit einer früh einsetzenden Metastasierung und einer schlechteren Prognose bei Brustkrebs vom histologischen Grad II und tripelnegativem Brustkrebs assoziiert (150). Die Hypermethylierung des *FOXF2*-Promotors scheint insbesondere bei luminalen und *HER2*-positiven Brustkrebszelllinien mit der malignen Progression einherzugehen (151). Dies passt zu unseren Ergebnissen, da in unserer Studie *FOXF2* ebenfalls hypermethyliert war, wie bei unseren *HER2*/neu-positiven Mäusen zu erwarten war. (3) Die Hypermethylierung des Transkriptionsfaktorgens *HIC1* kann bekanntermaßen zur Progression von

Brustkrebs beitragen (152–154). Die Hypermethylierung von *HIC1* konnte im Serum von circa 40% der Brustkrebspatientinnen ohne und bei 90% der Patientinnen mit metastasiertem Brustkrebs festgestellt werden (155). Die Wiederherstellung der *HIC1*-Expression führt zu reduzierter Zellmigration, -invasion und Metastasierung (156). Auch in unserer Studie zeigte sich *HIC1* hypermethyliert (4). *HOX*-Gene sind Transkriptionsfaktoren, die die Identität von Zellen und Gewebe entlang der anterior-posterioren Körperachse im sich entwickelnden Embryo festlegen (157). *HOXA5* ist bekanntermaßen auch bei Brustkrebs hypermethyliert (158) und kann die Expression von *TP53* (Tumorprotein p53) kontrollieren (159). Bei gemeinsamem Auftreten von Mutationen in *TP53* und Hypermethylierung von *HOXA5* kann dies zur Inaktivierung des *TP53*-Signalwegs führen (158). Der Verlust von *HOXA5* begünstigt die Entdifferenzierung und könnte somit die Tumorprogression fördern (160). Unsere Daten untermauern die These der Beteiligung von Hypermethylierung von *HOXA5* bei der malignen Progression von Brustkrebs. (5) In Brustkrebs-Zelllinien konnte eine Regulierung des Promotors von *PAX6* durch DNA-Methylierung festgestellt werden, was wiederum die Zellmigration und die Expression der pro-metastatischen Proteine MMP2 und MMP9 beeinflusst (112). Die untersuchten Regionen in *PAX6* zeigten sich auch in unserer Studie in den Metastasen im Vergleich zu den Primärtumoren hypermethyliert und bestätigten so die Hypothese der Beteiligung von *PAX6* an der malignen Tumorprogression bei Brustkrebs.

Es zeigten sich aber auch einige Regionen hypermethyliert, welche bisher noch unbekannt waren. Interessanterweise ergaben unsere Daten, dass die in Metastasen hypermethylierten Regionen eher mit Genen assoziiert waren, die für die Organentwicklung und generelle zelluläre Mechanismen zuständig sind. Dies lässt sich unter anderem dadurch erklären, dass in einer Anreicherungsanalyse, wie in dieser Arbeit, vor allem häufig vorkommende Gene Assoziationen aufweisen. Gleichzeitig kann es auch sein, dass die Mechanismen bei der Metastasierung sich auch teilweise mit den Prozessen bei der Organentwicklung und -ausreifung überschneiden (z.B. Migration der Zellen).

Für das bessere Verständnis und eine Einordnung unserer Ergebnisse könnten folgende weitere Experimente sinnvoll sein: Erstens könnte es Sinn machen, weitere

Experimente mit DNMT3-Knockout-Mäusen durchzuführen, um die Beziehung zwischen genetischen und epigenetischen Veränderungen weiter zu beleuchten. Da Metastasen mehr Methylierung aufweisen, stellt sich die Frage, ob Zellen mit vermehrter Methylierung bevorzugt metastasieren oder ob diese Methylierung erst in der metastatischen Nische durch die veränderte Mikroumgebung entsteht. Hierfür könnten in einem nächsten Schritt BalbNeuT-Brustdrüsenzellen transplantiert werden, die je Maus zu 50 % aus Zellen mit DNMT3-Knockout bestehen und zu 50% aus Zellen ohne Knockout. Dies könnte über die Ermittlung des Anteils der Knockout-Zellen in den späteren Tumoren und Metastasen Aufschluss über den Entstehungsort und Entstehungszeitpunkt der DNA-Methylierung geben. Zweitens wäre es hilfreich, die post-transkriptionellen Geschehnisse weiter zu untersuchen. Man könnte zum Beispiel die messenger-RNA und schließlich die Proteinexpression der entsprechend vermehrt oder vermindert exprimierten Gene untersuchen, um herauszufinden, welcher der angestoßenen Prozesse tatsächlich Auswirkungen auf die Proteinsynthese haben. Im Anschluss daran müssten die entsprechenden Regionen dann durch gezieltes (Bisulfit-)Sequenzieren validiert werden, um eine höhere Auflösung auf Allelebene zu erhalten. Und drittens ist es unerlässlich, auch andere epigenetische Veränderungen zu begutachten, wie zum Beispiel Chromatinveränderungen und Histonmodifikationen. Insbesondere Histonmodifikationen wurden bereits als potenzielle Marker für das Therapieansprechen bei Brustkrebs beschrieben (161). So gibt es konkrete Hinweise, dass die Überexpression von EZH2, welches die Effekte des repressiven trimethylierten Histonmarks H3K27me3 vermittelt, mit vermehrter Aggressivität und schlechterer Prognose von Brustkrebs assoziiert ist (74,162).

Im Bereich der klinisch-translationalen Forschung eröffnet unsere Studie auch weitere Ansatzpunkte. In vielen Bereichen der Medizin ist es bereits zum Standard geworden, genetische und epigenetische Diagnostik der Therapie voranzustellen, um spezifische Untergruppen einer Erkrankung individualisiert therapieren und die Prognose basierend auf epigenetischen Untersuchungen besser einschätzen zu können. Die Vorteile im Vergleich zu genetischen Mutationen sind das erhöhte Vorkommen aberranter Methylierungsmuster, die zuverlässige Erkennung der veränderten Methylierung und die relativ einfachen technischen Verfahren zur Detektion (163). So konnten z.B. die methylierten Gene *APC* und *RASSF1A* bereits als unabhängige Prognoseparameter beim Brustkrebs identifiziert werden (164) und *PRAC2*, *TDRD10*

und *TMEM132C* sowie die Adenylatcyclase *ADCY4* gelten als mögliche zukünftige diagnostische und/oder prognostische Biomarker für Brustkrebs (165,166). Darüber hinaus scheint es möglich zu sein, Subtypen von Brustkrebs durch verschiedene Signaturen auf der Grundlage von Histonmodifikationen einzuteilen und die klinische Prognose einzuschätzen (77).

Substanzen, die epigenetisch modulierend wirken, können aber auch als Therapeutika, zum Beispiel im Bereich der individualisierten Medizin angewendet werden (167). So werden bereits Medikamente eingesetzt, welche die Methylierung beeinflussen, z.B. Azacitidin in der Therapie der AML, wobei noch zahlreiche weitere Substanzen in Erprobung sind (168). Im Bereich weiterer epigenetischer Veränderungen (Histon-Acetylierung und Histon-Methylierung) befinden sich ebenfalls Medikamente in Entwicklung (169), um die Therapie von Krebserkrankungen zunehmend zu personalisieren (167). Bestimmte Histonacetylierungen und -methylierungen scheinen dynamischen Einfluss auf verschiedene Brustkrebstypen auszuüben und bieten damit mögliche Zielstrukturen für eine Therapie (78). Durch die komplexen epigenetischen Signalwege beim Brustkrebs ist die gezielte Therapie allerdings erschwert (170). So konnte gezeigt werden, dass die Verwendung von Aromataseinhibitoren bei Brustkrebspatientinnen zur Rekrutierung von EZH2 führt und damit zu einer Erhöhung des H2K27me3 des Homeobox-Gens *HOXC10* in Brustkrebszellen, was letztendlich zur Methylierung und zum Silencing von *HOXC10* und zur Resistenzentwicklung gegenüber Aromataseinhibitoren führt (171). Daher sind Kombinationstherapien in der Therapie von Brustkrebs unerlässlich (170). Unsere Arbeit liefert also weitere Ansatzpunkte zur Beurteilung der diagnostischen, therapeutischen und prognostischen Bedeutsamkeit der von uns identifizierten Regionen in einem klinisch-translationalen Ansatz.

## **5.2 Zusammenfassung**

Es gibt bereits vielfältige neue Ansätze die prognostische und therapeutische Relevanz der epigenetischen Veränderungen bei Brustkrebs zu untersuchen, welche teilweise bereits erfolgreich in der klinischen Praxis angewendet werden. Unsere Studie ergänzt diese Aspekte um die Erforschung der grundlegenden epigenetischen Mechanismen bei der malignen Progression von Brustkrebs im Mausmodell.



Wir konnten erstens zeigen, dass es Unterschiede im Methylierungsmuster zwischen Primärtumoren und Metastasen gibt, wenngleich die unterschiedlichen Zelltypen nicht ganz eindeutig differenzierbar waren, was auf eine frühe Dissemination der Metastasenzellen hinweisen könnte. Zweitens konnten wir diejenigen genetischen Regionen identifizieren, die in den Metastasen hyper-, bzw. hypomethyliert waren. Dies liefert einen weiteren Baustein im Verständnis der Pathogenese der Metastasierung von Brustkrebs, benötigt aber eine tiefergehende Analyse der beteiligten Regionen, sowie der weiteren biochemischen Prozesse im Rahmen der Transkription und Translation, bevor eine klinische Anwendung möglich ist.

## 6 Literaturverzeichnis

1. Shumway DA, Sabolch A, Jagsi R. Breast Cancer Facts and Figures 2020-American Cancer Society. Am Cancer Soc. 2020;1–43.
2. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchio C, Reis-Filho JS. Breast cancer precursors revisited. Histopathology. 2010;57(2):171–92.
3. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015 Sep;26:v8–30.
4. Barnes B, Kraywinkel K, Nowossadeck E, Schönfeld I, Starker A, Wienecke A, et al. Bericht zum Krebsgeschehen in Deutschland 2016. Robert Koch-Institut; 2016.
5. Krebs in Deutschland für 2015/2016. 12. Ausgabe. Robert Koch-Institut (Hrsg) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg). Berlin, 2019.
6. Rossi S, Basso M, Strippoli A, Dadduzio V, Cerchiaro E, Barile R, et al. Hormone Receptor Status and HER2 Expression in Primary Breast Cancer Compared With Synchronous Axillary Metastases or Recurrent Metastatic Disease. Clin Breast Cancer. 2015;15(5):307–12.
7. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2016;34(10):1134–50.
8. van Poznak C, Somerfield MR, Bast RC, Cristofanilli M, Goetz MP, Gonzalez-Angulo AM, et al. Use of Biomarkers to Guide Decisions on Systemic Therapy for Women With Metastatic Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2015;33(24):2695–704.
9. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science. 1985;229(4717):974–6.
10. Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, et al. The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. Proc Natl Acad Sci U S A. 1999;96(9):4995–5000.
11. Karunagaran D, Tzahar E, Beerli RR, Chen X, Graus-Porta D, Ratzkin BJ, et al. ErbB-2 is a common auxiliary subunit of NDF and EGF receptors: implications for breast cancer. EMBO J. 1996;15(2):254–64.
12. Petit AM, Rak J, Hung MC, Rockwell P, Goldstein N, Fendly B, et al. Neutralizing antibodies against epidermal growth factor and ErbB-2/neu receptor tyrosine kinases down-regulate vascular endothelial growth factor production by tumor cells in vitro and in vivo: angiogenic implications for signal transduction therapy of so. Am J Pathol. 1997;151(6):1523–30.
13. Giatromanolaki A, Koukourakis MI, Simopoulos C, Polychronidis A, Gatter KC,

- Harris AL, et al. c-erbB-2 related aggressiveness in breast cancer is hypoxia inducible factor-1alpha dependent. *Clin Cancer Res.* 2004;10(23):7972–7.
14. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science.* 1987;235(4785):177–82.
  15. Yaziji H, Goldstein LC, Barry TS, Werling R, Hwang H, Ellis GK, et al. HER-2 testing in breast cancer using parallel tissue-based methods. *JAMA.* 2004;291(16):1972–7.
  16. Owens MA, Horten BC, Da Silva MM. HER2 amplification ratios by fluorescence in situ hybridization and correlation with immunohistochemistry in a cohort of 6556 breast cancer tissues. *Clin Breast Cancer.* 2004;5(1):63–9.
  17. Drebin JA, Link VC, Weinberg RA, Greene MI. Inhibition of tumor growth by a monoclonal antibody reactive with an oncogene-encoded tumor antigen. *Proc Natl Acad Sci U S A.* 1986;83(23):9129–33.
  18. National Cancer Institute. caMOD - Cancer Models Database. Model Characteristics - Model: BALB-neuT [Internet]. [zitiert am 04.09.2015]. Available from:  
[https://cancermodels.nci.nih.gov/camod/ViewModelAction.do?unprotected\\_method=populateModelCharacteristics&aModelID=448](https://cancermodels.nci.nih.gov/camod/ViewModelAction.do?unprotected_method=populateModelCharacteristics&aModelID=448)
  19. Waddington CH. The Epigenotype. *Int J Epidemiol.* 2012 Feb;41(1):10–3.
  20. Zhang G, Pradhan S. Mammalian epigenetic mechanisms. *IUBMB Life.* 2014 Apr 5;66(4):240–56.
  21. Baxter E, Windloch K, Gannon F, Lee JS. Epigenetic regulation in cancer progression. *Cell Biosci.* 2014;4(1):45.
  22. Singal R, Ginder GD. DNA methylation. *Blood.* 1999 Jun 15;93(12):4059–70.
  23. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002;16(1):6–21.
  24. Jovanovic J, Rønneberg JA, Tost J, Kristensen V. The epigenetics of breast cancer. *Mol Oncol.* 2010 Jun;4(3):242–54.
  25. Schubeler D. Function and information content of DNA methylation. *Nature.* 2015;517(7534):321–6.
  26. Okano M, Bell DW, Haber DA, Li E. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell.* 1999 Oct;99(3):247–57.
  27. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* 2009;324(5929):930–5.
  28. Guo JU, Su Y, Zhong C, Ming G, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell.* 2011;145(3):423–34.
  29. Valinluck V, Sowers LC. Endogenous Cytosine Damage Products Alter the Site

- Selectivity of Human DNA Maintenance Methyltransferase DNMT1. *Cancer Res.* 2007 Feb 1;67(3):946–50.
30. Ji D, Lin K, Song J, Wang Y. Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Mol Biosyst.* 2014;10(7):1749.
  31. Shi D-Q, Ali I, Tang J, Yang W-C. New Insights into 5hmC DNA Modification: Generation, Distribution and Function. *Front Genet.* 2017 Jul 19;8(JUL).
  32. Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* 2012 Jun;40(11):4841–9.
  33. Liu X, Gao Q, Li P, Zhao Q, Zhang J, Li J, et al. UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat Commun.* 2013 Jun 5;4(1):1563.
  34. Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I, et al. Recognition of 5-Hydroxymethylcytosine by the Uhrf1 SRA Domain. Xu S, editor. *PLoS One.* 2011 Jun 22;6(6):e21306.
  35. Hahn MA, Szabó PE, Pfeifer GP. 5-Hydroxymethylcytosine: A stable or transient DNA modification? *Genomics.* 2014 Nov;104(5):314–23.
  36. Jenuwein T, Allis CD. Translating the histone code. *Science.* 2001;293(5532):1074–80.
  37. Audia JE, Campbell RM. Histone Modifications and Cancer. *Cold Spring Harb Perspect Biol.* 2016 Apr 1;8(4):a019521.
  38. Højfeldt JW, Agger K, Helin K. Histone lysine demethylases as targets for anticancer therapy. *Nat Rev Drug Discov.* 2013 Dec 15;12(12):917–30.
  39. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res.* 2011 Apr 22;21(4):564–78.
  40. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. *Exp Mol Med.* 2017 Apr 28;49(4):e324–e324.
  41. Black JC, Van Rechem C, Whetstine JR. Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Mol Cell.* 2012 Nov;48(4):491–507.
  42. Shilatifard A. The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. *Annu Rev Biochem.* 2012 Jul 7;81(1):65–95.
  43. Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun Z-W, Schmid M, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature.* 2000 Aug;406(6796):593–9.
  44. Brower-Toland B, Riddle NC, Jiang H, Huisinga KL, Elgin SCR. Multiple SET Methyltransferases Are Required to Maintain Normal Heterochromatin Domains in the Genome of *Drosophila melanogaster*. *Genetics.* 2009 Apr 1;181(4):1303–19.

45. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity that Marks Chromosomal Polycomb Sites. *Cell*. 2002 Oct;111(2):185–96.
46. Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, et al. Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex. *Cell*. 2002 Oct;111(2):197–208.
47. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol*. 2012 Feb 23;13(2):115–26.
48. Edmunds JW, Mahadevan LC, Clayton AL. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J*. 2008 Jan 23;27(2):406–20.
49. Nguyen AT, Zhang Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev*. 2011 Jul 1;25(13):1345–58.
50. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 2004;119(7):941–53.
51. Huang Y, Greene E, Murray Stewart T, Goodwin AC, Baylin SB, Woster PM, et al. Inhibition of lysine-specific demethylase 1 by polyamine analogues results in reexpression of aberrantly silenced genes. *Proc Natl Acad Sci U S A*. 2007;104(19):8023–8.
52. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, et al. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature*. 2009 Sep 2;461(7262):415–8.
53. Ng SS, Kavanagh KL, McDonough MA, Butler D, Pilka ES, Lienard BMR, et al. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature*. 2007;448(7149):87–91.
54. Kouzarides T. Chromatin Modifications and Their Function. *Cell*. 2007 Feb;128(4):693–705.
55. Sterner DE, Berger SL. Acetylation of Histones and Transcription-Related Factors. *Microbiol Mol Biol Rev*. 2000 Jun 1;64(2):435–59.
56. Gray SG, Ekström TJ. The Human Histone Deacetylase Family. *Exp Cell Res*. 2001 Jan;262(2):75–83.
57. Voelter-Mahlknecht S, Ho AD, Mahlkecht U. Chromosomal organization and localization of the novel class IV human histone deacetylase 11 gene. *Int J Mol Med*. 2005 Oct;16(4):589–98.
58. Struhl K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev*. 1998 Mar 1;12(5):599–606.
59. Houlston RS, Peto J. The search for low-penetrance cancer susceptibility alleles. *Oncogene*. 2004;23(38):6471–6.
60. Antoniou A, Pharoah PDP, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2

- mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet.* 2003;72(5):1117–30.
61. Goldberg JI, Borgen PI. Breast cancer susceptibility testing: past, present and future. *Expert Rev Anticancer Ther.* 2006;6(8):1205–14.
  62. Loveday C, Turnbull C, Ramsay E, Hughes D, Ruark E, Frankum JR, et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet.* 2011;43(9):879–82.
  63. Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet.* 2010;42(5):410–4.
  64. Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, et al. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet.* 2011;44(1):40–6.
  65. Bae H, Kim B, Lee H, Lee S, Kang H-S, Kim SJ. Epigenetically regulated Fibronectin leucine rich transmembrane protein 2 (FLRT2) shows tumor suppressor activity in breast cancer cells. *Sci Rep.* 2017;7(1):272.
  66. Feinberg AP, Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun.* 1983;111(1):47–54.
  67. Feinberg AP, Tycko B. The history of cancer epigenetics. *Nat Rev Cancer.* 2004 Feb;4(2):143–53.
  68. Sproul D, Nestor C, Culley J, Dickson JH, Dixon JM, Harrison DJ, et al. Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer. *Proc Natl Acad Sci U S A.* 2011;108(11):4364–9.
  69. Feinberg AP, Gehrke CW, Kuo KC, Ehrlich M. Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer Res.* 1988;48(5):1159–61.
  70. Gama-Sosa MA, Slagel VA, Trewyn RW, Oxenhandler R, Kuo KC, Gehrke CW, et al. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res.* 1983;11(19):6883–94.
  71. Jones PA, Gonzalzo ML. Altered DNA methylation and genome instability. *Proc Natl Acad Sci U S A.* 1997;94(6):2103–5.
  72. Goelz SE, Vogelstein B, Hamilton SR, Feinberg AP. Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science.* 1985;228(4696):187–90.
  73. Dong C, Wu Y, Yao J, Wang Y, Yu Y, Rychahou PG, et al. G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *J Clin Invest.* 2012 Apr 2;122(4):1469–86.
  74. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc Natl Acad Sci U S A.* 2003;100(20):11606–11.

75. Yokoyama Y, Matsumoto A, Hieda M, Shinci Y, Ogihara E, Hamada M, et al. Loss of histone H4K20 trimethylation predicts poor prognosis in breast cancer and is associated with invasive activity. *Breast Cancer Res.* 2014;16(3):R66.
76. Ma L, Yuan L, An J, Barton MC, Zhang Q, Liu Z. Histone H3 lysine 23 acetylation is associated with oncogene TRIM24 expression and a poor prognosis in breast cancer. *Tumour Biol.* 2016;37(11):14803–12.
77. Chen X, Hu H, He L, Yu X, Liu X, Zhong R, et al. A novel subtype classification and risk of breast cancer by histone modification profiling. *Breast Cancer Res Treat.* 2016;157(2):267–79.
78. Messier TL, Gordon JAR, Boyd JR, Tye CE, Browne G, Stein JL, et al. Histone H3 lysine 4 acetylation and methylation dynamics define breast cancer subtypes. *Oncotarget.* 2016;7(5):5094–109.
79. Kapoor-Vazirani P, Kagey JD, Powell DR, Vertino PM. Role of hMOF-dependent histone H4 Lysine 16 acetylation in the maintenance of TMS1/ASC gene activity. *Cancer Res.* 2008;68(16):6810–21.
80. Pfister S, Rea S, Taipale M, Mendrzyk F, Straub B, Itrich C, et al. The histone acetyltransferase hMOF is frequently downregulated in primary breast carcinoma and medulloblastoma and constitutes a biomarker for clinical outcome in medulloblastoma. *Int J Cancer.* 2008 Mar 15;122(6):1207–13.
81. Yang X, Karuturi RKM, Sun F, Aau M, Yu K, Shao R, et al. CDKN1C (p57KIP2) Is a Direct Target of EZH2 and Suppressed by Multiple Epigenetic Mechanisms in Breast Cancer Cells. Blagosklonny M V., editor. *PLoS One.* 2009 Apr 2;4(4):e5011.
82. Wang Y, Zhang H, Chen Y, Sun Y, Yang F, Yu W, et al. LSD1 Is a Subunit of the NuRD Complex and Targets the Metastasis Programs in Breast Cancer. *Cell.* 2009 Aug;138(4):660–72.
83. Wagener C, Müller O. *Molekulare Onkologie. 3., komple. s.l.: Thieme; 2009.*
84. Hosseini H, Obradović MMS, Hoffmann M, Harper KL, Sosa MS, Werner-Klein M, et al. Early dissemination seeds metastasis in breast cancer. *Nature.* 2016 Dec 22;540(7634):552–8.
85. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. *Science (80-).* 2011;331(6024):1559–64.
86. Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis. *Nat Rev Cancer.* 2005;5(8):591–602.
87. Chan KK, Matchett KB, McEnhill PM, Dakir EH, McMullin MF, El-Tanani Y, et al. Protein deregulation associated with breast cancer metastasis. *Cytokine Growth Factor Rev.* 2015;26(4):415–23.
88. Cifone MA, Fidler IJ. Increasing metastatic potential is associated with increasing genetic instability of clones isolated from murine neoplasms. *Proc Natl Acad Sci U S A.* 1981;78(11):6949–52.
89. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, et al. Genes that mediate breast cancer metastasis to lung. *Nature.* 2005;436(7050):518–24.

90. Patel SA, Vanharanta S. Epigenetic determinants of metastasis. *Mol Oncol*. 2017 Jan;11(1):79–96.
91. Jacob LS, Vanharanta S, Obenauf AC, Pirun M, Viale A, Socci ND, et al. Metastatic Competence Can Emerge with Selection of Preexisting Oncogenic Alleles without a Need of New Mutations. *Cancer Res*. 2015;75(18):3713–9.
92. Massague J, Obenauf AC. Metastatic colonization by circulating tumour cells. *Nature*. 2016;529(7586):298–306.
93. Oskarsson T, Batlle E, Massague J. Metastatic stem cells. *Cell Stem Cell*. 2014;14(3):306–21.
94. Valastyan S, Weinberg RA. Tumor metastasis. *Cell*. 2011;147(2):275–92.
95. Fang F, Turcan S, Rimner A, Kaufman A, Giri D, Morris LGT, et al. Breast Cancer Methylomes Establish an Epigenomic Foundation for Metastasis // Breast cancer methylomes establish an epigenomic foundation for metastasis. *Sci Transl Med*. 2011;3(75):75ra25.
96. Aryee MJ, Liu W, Engelmann JC, Nuhn P, Gurel M, Haffner MC, et al. DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. *Sci Transl Med*. 2013;5(169):169ra10.
97. Reyngold M, Turcan S, Giri D, Kannan K, Walsh LA, Viale A, et al. Remodeling of the methylation landscape in breast cancer metastasis. *PLoS One*. 2014;9(8):e103896.
98. Gebhard C, Benner C, Ehrich M, Schwarzfischer L, Schilling E, Klug M, et al. General Transcription Factor Binding at CpG Islands in Normal Cells Correlates with Resistance to De novo DNA Methylation in Cancer Cells. *Cancer Res*. 2010;70(4):1398–407.
99. Sproul D, Meehan RR. Genomic insights into cancer-associated aberrant CpG island hypermethylation. *Brief Funct Genomics*. 2013;12(3):174–90.
100. Sproul D, Kitchen RR, Nestor CE, Dixon JM, Sims AH, Harrison DJ, et al. Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol*. 2012;13(10):R84.
101. Chen J, Wang F, Xu J, He Z, Lu Y, Wang Z. The role of PAQR3 gene promoter hypermethylation in breast cancer and prognosis. *Oncol Rep*. 2016;36(3):1612–8.
102. Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, et al. DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia. *Cancer Cell*. 2010 Jan;17(1):13–27.
103. Yang L, Rau R, Goodell MA. DNMT3A in haematological malignancies. *Nat Rev Cancer*. 2015 Mar 19;15(3):152–65.
104. Bacher U, Haferlach C, Schnittger S, Kohlmann A, Kern W, Haferlach T. Mutations of the TET2 and CBL genes: novel molecular markers in myeloid malignancies. *Ann Hematol*. 2010 Jul 2;89(7):643–52.



105. Yang X, Wong MPM, Ng RK. Aberrant DNA Methylation in Acute Myeloid Leukemia and Its Clinical Implications. *Int J Mol Sci.* 2019 Sep 16;20(18):4576.
106. Xu J, Zhang W, Yan X-J, Lin X-Q, Li W, Mi J-Q, et al. DNMT3A mutation leads to leukemic extramedullary infiltration mediated by TWIST1. *J Hematol Oncol.* 2016 Dec 10;9(1):106.
107. Malouf GG, Taube JH, Lu Y, Roysarkar T, Panjarian S, Estecio MR, et al. Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. *Genome Biol.* 2013;14(12):R144.
108. Tiwari N, Tiwari VK, Waldmeier L, Balwierz PJ, Arnold P, Pachkov M, et al. Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. *Cancer Cell.* 2013;23(6):768–83.
109. Zingg D, Debbache J, Schaefer SM, Tuncer E, Frommel SC, Cheng P, et al. The epigenetic modifier EZH2 controls melanoma growth and metastasis through silencing of distinct tumour suppressors. *Nat Commun.* 2015 May 22;6(1):6051.
110. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646–74.
111. Hanahan D, Weinberg RA. The Hallmarks of Cancer. *Cell.* 2000 Jan;100(1):57–70.
112. Urrutia G, Laurito S, Marzese DM, Gago F, Orozco J, Tello O, et al. Epigenetic variations in breast cancer progression to lymph node metastasis. *Clin Exp Metastasis.* 2015;32(2):99–110.
113. Klein CA, Blankenstein TJF, Schmidt-Kittler O, Petronio M, Polzer B, Stoecklein NH, et al. Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet (London, England).* 2002;360(9334):683–9.
114. Schmidt-Kittler O, Ragg T, Daskalakis A, Granzow M, Ahr A, Blankenstein TJF, et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A.* 2003;100(13):7737–42.
115. Klein CA. Parallel progression of primary tumours and metastases. *Nat Rev Cancer.* 2009;9(4):302–12.
116. Hüsemann Y, Geigl JB, Schubert F, Musiani P, Meyer M, Burghart E, et al. Systemic spread is an early step in breast cancer. *Cancer Cell.* 2008;13(1):58–68.
117. Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, et al. EMT and dissemination precede pancreatic tumor formation. *Cell.* 2012;148(1–2):349–61.
118. Eyles J, Puaux A-L, Wang X, Toh B, Prakash C, Hong M, et al. Tumor cells disseminate early, but immunosurveillance limits metastatic outgrowth, in a mouse model of melanoma. *J Clin Invest.* 2010;120(6):2030–9.
119. Banys M, Gruber I, Krawczyk N, Becker S, Kurth R, Wallwiener D, et al. Hematogenous and lymphatic tumor cell dissemination may be detected in patients diagnosed with ductal carcinoma in situ of the breast. *Breast Cancer Res Treat.* 2012;131(3):801–8.

120. Sanger N, Effenberger KE, Riethdorf S, van Haasteren V, Gauwerky J, Wiegratz I, et al. Disseminated tumor cells in the bone marrow of patients with ductal carcinoma in situ. *Int J cancer*. 2011;129(10):2522–6.
121. Rykova EY, Tsvetovskaya GA, Sergeeva GI, Vlassov V V, Laktionov PP. Methylation-Based Analysis of Circulating DNA for Breast Tumor Screening. *Ann N Y Acad Sci*. 2008 Aug;1137(1):232–5.
122. Szyf M. DNA methylation signatures for breast cancer classification and prognosis. *Genome Med*. 2012;4(3):26.
123. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012 Apr 4;9(4):357–9.
124. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
125. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell*. 2010 May;38(4):576–89.
126. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014 Dec 5;15(12):550.
127. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010 May 2;28(5):495–501.
128. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019 Dec 3;10(1):1523.
129. Freemont PS, Ollis DL, Steitz TA, Joyce CM. A domain of the klenow fragment of *Escherichia coli* DNA polymerase I has polymerase but no exonuclease activity. *Proteins Struct Funct Genet*. 1986 Jan;1(1):66–73.
130. Dewey FE, Pan S, Wheeler MT, Quake SR, Ashley EA. DNA Sequencing. *Circulation*. 2012 Feb 21;125(7):931–44.
131. Chaitankar V, Karakulah G, Ratnapriya R, Giuste FO, Brooks MJ, Swaroop A. Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research. *Prog Retin Eye Res*. 2016 Nov;55:1–31.
132. Lindqvist BM, Wingren S, Motlagh PB, Nilsson TK. Whole genome DNA methylation signature of HER2-positive breast cancer. *Epigenetics*. 2014 Aug 6;9(8):1149–62.
133. Lian ZQ, Wang Q, Li WP, Zhang AQ, Wu L. Screening of significantly hypermethylated genes in breast cancer using microarray-based methylated-CpG island recovery assay and identification of their expression levels. *Int J Oncol*. 2012 Aug;41(2):629–38.
134. Pearson K. LIII. On lines and planes of closest fit to systems of points in space. London, Edinburgh, Dublin *Philos Mag J Sci*. 1901 Nov 8;2(11):559–72.

135. Duntean G. Basic Concepts of Principal Components Analysis. In: *Principal Components Analysis* [Internet]. 2455 Teller Road, Newbury Park California 91320 United States of America: SAGE Publications, Inc.; 2011. p. 16–23. [zitiert am 30.04.2021]. Available from: <http://methods.sagepub.com/book/principal-components-analysis/n2.xml>
136. Van Der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-sne. *J Mach Learn Res* [Internet]. 2008;9:2579–605. [zitiert am 30.04.2021]. Available from: <http://www.jmlr.org/papers/v9/vandermaaten08a.html>
137. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci*. 2006 Jan 31;103(5):1412–7.
138. Kulis M, Esteller M. DNA Methylation and Cancer. In: *Advances in genetics*. 2010. p. 27–56.
139. ShahidSales S, Mehramiz M, Ghasemi F, Aledavood A, Shamsi M, Hassanian SM, et al. A genetic variant in CDKN2A/B gene is associated with the increased risk of breast cancer. *J Clin Lab Anal*. 2018 Jan;32(1):e22190.
140. Barekati Z, Radpour R, Lu Q, Bitzer J, Zheng H, Toniolo P, et al. Methylation signature of lymph node metastases in breast cancer patients. *BMC Cancer*. 2012 Dec 13;12(1):244.
141. Spitzwieser M, Entfellner E, Werner B, Pulverer W, Pfeiler G, Hacker S, et al. Hypermethylation of CDKN2A exon 2 in tumor, tumor-adjacent and tumor-distant tissues from breast cancer patients. *BMC Cancer*. 2017 Dec 12;17(1):260.
142. Zhang Y, Ding S, Yang J, Chen X, Huang W. Identification of miR-146a is Associated with the Aggressiveness and Suppresses Proliferation via Targeting CDKN2A in Breast Cancer. *Pathol Oncol Res*. 2020 Jan 18;26(1):245–51.
143. Wang L, Zhan X, Shen X, Li M, Yang J, Yu W, et al. P16 promotes the growth and mobility potential of breast cancer both in vitro and in vivo: the key role of the activation of IL-6/JAK2/STAT3 signaling. *Mol Cell Biochem*. 2018 Sep;446(1–2):137–48.
144. Chatterjee A, Rodger EJ, Eccles MR. Epigenetic drivers of tumourigenesis and cancer metastasis. *Semin Cancer Biol*. 2018 Aug;51:149–59.
145. Cha YJ, Jung WH, Koo JS. Expression of DNA methylation-related proteins in metastatic breast cancer. *Neoplasma*. 2017;64(3):412–20.
146. Munoz DP, Lee EL, Takayama S, Coppe J-P, Heo S-J, Boffelli D, et al. Activation-induced cytidine deaminase (AID) is necessary for the epithelial-mesenchymal transition in mammary epithelial cells. *Proc Natl Acad Sci*. 2013 Aug 6;110(32):E2977–86.
147. Chatterjee A, Stockwell PA, Ahn A, Rodger EJ, Leichter AL, Eccles MR. Genome-wide methylation sequencing of paired primary and metastatic cell lines identifies common DNA methylation changes and a role for EBF3 as a candidate epigenetic driver of melanoma metastasis. *Oncotarget*. 2017 Jan 24;8(4):6085–101.

148. Bockhorn J, Prat A, Chang YF, Liu X, Huang S, Shang M, et al. Differentiation and loss of malignant character of spontaneous pulmonary metastases in patient-derived breast cancer models. *Cancer Res.* 2014;74(24):7406–17.
149. Li Y, Melnikov AA, Levenson V, Guerra E, Simeone P, Alberti S, et al. A seven-gene CpG-island methylation panel predicts breast cancer progression. *BMC Cancer.* 2015 Dec 19;15(1):417.
150. Kong P-Z, Yang F, Li L, Li X-Q, Feng Y-M. Decreased FOXF2 mRNA Expression Indicates Early-Onset Metastasis and Poor Prognosis for Breast Cancer Patients with Histological Grade II Tumor. de Wever O, editor. *PLoS One.* 2013 Apr 19;8(4):e61591.
151. Lo P-K, Lee JS, Liang X, Sukumar S. The dual role of FOXF2 in regulation of DNA replication and the epithelial-mesenchymal transition in breast cancer progression. *Cell Signal.* 2016;28(10):1502–19.
152. Fujii H, Biel MA, Zhou W, Weitzman SA, Baylin SB, Gabrielson E. Methylation of the HIC-1 candidate tumor suppressor gene in human breast cancer. *Oncogene.* 1998 Apr 30;16(16):2159–64.
153. Parrella P, Poeta ML, Gallo AP, Prencipe M, Scintu M, Apicella A, et al. Nonrandom Distribution of Aberrant Promoter Methylation of Cancer-Related Genes in Sporadic Breast Tumors. *Clin Cancer Res.* 2004 Aug 15;10(16):5349–54.
154. Parrella P, Scintu M, Prencipe M, Poeta ML, Gallo AP, Rabitti C, et al. HIC1 promoter methylation and 17p13.3 allelic loss in invasive ductal carcinoma of the breast. *Cancer Lett.* 2005 May;222(1):75–81.
155. Müller HM, Widschwendter A, Fiegl H, Ivarsson L, Goebel G, Perkmann E, et al. DNA methylation in serum of breast cancer patients: an independent prognostic marker. *Cancer Res.* 2003 Nov 15;63(22):7641–5.
156. Cheng G, Sun X, Wang J, Xiao G, Wang X, Fan X, et al. HIC1 Silencing in Triple-Negative Breast Cancer Drives Progression through Misregulation of LCN2. *Cancer Res.* 2014 Feb 1;74(3):862–72.
157. Hur H, Lee JY, Yun HJ, Park BW, Kim MH. Analysis of HOX gene expression patterns in human breast cancer. *Mol Biotechnol.* 2014;56(1):64–71.
158. Abba MC, Gong T, Lu Y, Lee J, Zhong Y, Lacunza E, et al. A molecular portrait of high-grade ductal carcinoma in situ. *Cancer Res.* 2015 Sep 15;75(18):3980–90.
159. Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, et al. Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature.* 2000 Jun;405(6789):974–8.
160. Teo WW, Merino VF, Cho S, Korangath P, Liang X, Wu R, et al. HOXA5 determines cell fate transition and impedes tumor initiation and progression in breast cancer through regulation of E-cadherin and CD24. *Oncogene.* 2016 Oct 20;35(42):5539–51.
161. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, et al.

- Global Histone Modifications in Breast Cancer Correlate with Tumor Phenotypes, Prognostic Factors, and Patient Outcome. *Cancer Res.* 2009 May 1;69(9):3802–9.
162. Fujii S, Tokita K, Wada N, Ito K, Yamauchi C, Ito Y, et al. MEK-ERK pathway regulates EZH2 overexpression in association with aggressive breast cancer subtypes. *Oncogene.* 2011;30(39):4118–28.
163. Dworkin AM, Huang THM, Toland AE. Epigenetic alterations in the breast: Implications for breast cancer detection, prognosis and treatment. *Semin Cancer Biol.* 2009 Jun;19(3):165–71.
164. Müller HM, Fiegl H, Widschwendter A, Widschwendter M. Prognostic DNA methylation marker in serum of cancer patients. In: *Annals of the New York Academy of Sciences.* 2004. p. 44–9.
165. de Almeida BP, Apolónio JD, Binnie A, Castelo-Branco P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. *BMC Cancer.* 2019 Dec 12;19(1):219.
166. Fan Y, Mu J, Huang M, Imani S, Wang Y, Lin S, et al. Epigenetic identification of ADCY4 as a biomarker for breast cancer: an integrated analysis of adenylate cyclases. *Epigenomics.* 2019 Nov;11(14):1561–79.
167. Jones PA, Issa J-PJ, Baylin S. Targeting the cancer epigenome for therapy. *Nat Rev Genet.* 2016 Oct 15;17(10):630–41.
168. Christman JK. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene.* 2002 Aug 5;21(35):5483–95.
169. Kelly AD, Issa J-PJ. The promise of epigenetic therapy: reprogramming the cancer epigenome. *Curr Opin Genet Dev.* 2017 Feb;42:68–77.
170. Ronnekleiv-Kelly SM, Sharma A, Ahuja N. Epigenetic therapy and chemosensitization in solid malignancy. *Cancer Treat Rev.* 2017 Apr;55:200–8.
171. Pathiraja TN, Nayak SR, Xi Y, Jiang S, Garee JP, Edwards DP, et al. Epigenetic reprogramming of HOXC10 in endocrine-resistant breast cancer. *Sci Transl Med.* 2014;6(229):229ra41.

## 7 Danksagung

Zuerst möchte ich mich ganz herzlich bei meinem Doktorvater Prof. Dr. Michael Rehli bedanken. Von dem ersten Moment an, als er mir den Anspruch einer Arbeit in seiner Arbeitsgruppe erklärt und dabei mein Interesse geweckt hat, war es eine große Freude, Teil seiner Arbeitsgruppe zu sein. Er hat mich all die Zeit mit großer Geduld und viel Verständnis betreut und hatte immer ein offenes Ohr.

Auch Prof. Dr. Herr, in dessen Klinik für Innere Medizin III diese Arbeit entstanden ist und in der ich auch als Assistenzärztin in einem tollen Team arbeiten durfte, gilt ein großes Dankeschön.

Diese Arbeit entstand im Rahmen des FOR2127. Hier gilt mein Dank Prof. Dr. Klein und seinem Team für die Organisation vieler Vorträge und Retreats, die immer sehr bereichernd für mich waren. Außerdem möchte ich mich bei Christian Mulas bedanken, der mir bei der Gewinnung der Kontrollproben von BALB/c-Mäusen behilflich war.

Der Fakultät für Medizin der Universität Regensburg habe ich die Unterstützung durch ein Promotionsstipendium zu verdanken, das mir ermöglichte, mich für mehrere Monate nur auf die Arbeit im Labor zu konzentrieren.

Ich möchte mich außerdem beim ganzen Team im Labor bedanken, die den Kern meiner Erfahrungen während der Doktorarbeit ausgemacht haben. Insbesondere bei Dagmar, die mir am Anfang alles geduldig und genau gezeigt und erklärt hat und die stets zur Stelle war, wenn Fragen aufkamen. Und auch Johanna, Lucia, Julia W, Julia M, Sandra und Claudia standen mir stets mit Rat und Tat zur Seite und haben maßgeblich zu der warmherzigen Atmosphäre im H1 beigetragen und ihn somit zu einem tollen Arbeitsplatz gemacht, an dem viel gelacht wurde und den ich immer wieder gerne besucht habe. Und danke auch allen aus der AG Kreutz, die immer für einen Spaß zu haben waren.

Meiner Familie habe ich so viel auf diesem Weg zu verdanken. Vielen, vielen Dank Mama, Papa, Felix und Paul für die kontinuierliche Unterstützung, euren Glauben an mich und diese Arbeit, für unzählige motivierende Gespräche und dafür, dass ihr in allen Lebenslagen für mich da seid.

Und last but definitely not least: Ich möchte Dir von Herzen danken, Vincent! Ohne Dich wäre diese Arbeit nicht möglich gewesen. Du bist immer an meiner Seite und gehst mit mir durch dick und dünn. Auf den letzten Metern hast Du mir den Rücken freigehalten und mich unterstützt und angefeuert, wo es nur ging. Und bei Dir, allerliebste Lilian, möchte ich mich dafür bedanken, dass Du mir Tage stets mir Deinem strahlenden Lächeln versüßt hast, mein kleiner Sonnenschein!