



**Differenzielle Methylierungsanalysen mittels verschiedener Next-
Generation Sequencing-basierter Techniken:
Die Bedeutung von differenziell methylierten Regionen in der
menschlichen Hirnevolution und bei der Krebsentstehung**

Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades
an der Julius-Maximilians-Universität Würzburg

vorgelegt von

Julia Böck

geboren in

Augsburg

Würzburg, 2018

Eingereicht am:

Mitglieder der Promotionskommission:

Vorsitzender:

Gutachter: Univ.-Prof. Dr. med. Thomas Haaf

Gutachter: Univ.-Prof. Dr. Arthur Korte

Tag des Promotionskolloquiums:

Doktorurkunde ausgehändigt am:

ZUSAMMENFASSUNG

1. DNA-Methylierungsmuster corticaler Neuronen und Gliazellen im Menschen und Schimpansen

Die Evolution der Primaten zeigt eine Verbindung zwischen der zunehmenden Komplexität des sozialen Verhaltens und der Vergrößerung des humanen Gehirns, insbesondere des präfrontalen Cortex. Deshalb stellt der präfrontale Cortex bezüglich der Evolution des Menschen eine der interessantesten Strukturen im humanen Gehirn dar. Es wird angenommen, dass nicht allein die Größe, sondern auch die Funktion, vor allem das Zusammenspiel von Neuronen und nicht-neuronalen Zellen, wie z.B. Gliazellen, zur Differenzierung des menschlichen Gehirns von dem rezenter Primaten geführt hat. Daraus lässt sich schließen, dass die Gehirnfunktionen über eine ausgeglichene und gut aufeinander abgestimmte transkriptionelle Landschaft kontrolliert werden, die durch ein zugrundeliegendes genetisches und epigenetisches Rückgrat organisiert ist. In dieser Studie wurden das Methylierungsprofil neuronaler und nicht-neuronaler Zellen des präfrontalen Cortex (Brodmann-Areal 10) von drei Menschen und drei Schimpansen miteinander verglichen. Die intra- und interspezifischen differenziell methylierten Regionen (DMRs) waren in bestimmten genomischen Regionen angereichert. Intraspezifische Methylierungsunterschiede zwischen neuronalen und nicht-neuronalen Zellen konnten dreimal häufiger beobachtet werden als interspezifische Unterschiede in den einzelnen Zelltypen. Rund 90% der humanen intraspezifischen DMRs wiesen eine Hypomethylierung in den neuronalen Zellen im Vergleich zu den nicht-neuronalen Zellen auf. In den intraspezifischen DMRs (Mensch und Schimpanse) waren Gene angereichert, die mit verschiedenen neuropsychiatrischen Erkrankungen assoziiert sind. Der Vergleich zwischen Menschen und Schimpanse in den neuronalen und nicht-neuronalen Zelltypen zeigte eine Anreicherung von Genen mit human-spezifischer Histonsignatur. In den nicht-neuronalen Zellen konnten mehr interspezifische DMRs (n=666) detektiert werden als in den neuronalen Zellen (n=96). Ungefähr 95% der nicht-neuronalen interspezifischen DMRs waren im Menschen, im Vergleich zum Schimpansen, hypermethyliert. Daraus ergibt sich der Eindruck, dass mehrere hundert der nicht-neuronalen Gene während der humanen Gehirnevolution einer Methylierungswelle unterlagen. Dies führt zu der Annahme, dass der Einfluss dieser Veränderungen in den nicht-neuronalen Zellen auf die Vergrößerung des menschlichen Gehirns bisher stark unterschätzt wurde.

2. Epimutationscreening multipler Tumorsuppressorgene in mutationsnegativen early-onset und high-risk Brustkrebspatienten

Die bekannteste genetische Ursache für erblichen Brust- und Eierstockkrebs sind Mutationen in den Tumorsuppressorgenen (TSG) *BRCA1* und *BRCA2*. Dennoch können nur rund 20-25% der familiären Brustkrebserkrankungen über Keimbahnmutationen in *BRCA1/BRCA2* erklärt werden, besonders bei Frauen, deren Erkrankung vor dem vierzigsten Lebensjahr auftritt. Epigenetische Veränderungen, die zu einer aberranten Genexpression führen, spielen ebenfalls eine wichtige Rolle bei der Karzinogenese und der Entwicklung einer Brustkrebserkrankung. Es ist bekannt, dass TSG nicht nur durch den Verlust der Heterozygotie (engl. *loss of heterozygosity*, LOH) oder homozygote Deletionen, sondern auch durch transkriptionelle Stilllegung via DNA-Methylierung inaktiviert werden können. Im Rahmen dieser Arbeit wurde überprüft, welchen Einfluss aberrante Methylierungsmuster im Promotorbereich von TSG auf die Brustkrebskarzinogenese und die Expression der Gene haben. Für die Quantifizierung der Epimutationen wurden die Promotorbereiche von acht TSG (*BRCA1*, *BRCA2*, *RAD51C*, *ATM*, *PTEN*, *TP53*, *MLH1*, *RB1*) und des estrogenen Rezeptors (*ESR1*) Gens, welches eine Rolle in der Tumorphyse spielt, mittels Deep Bisulfite Amplicon Sequencing (DBAS) analysiert. Es wurden Blutproben von zwei unabhängigen *BRCA1/BRCA2*-mutationsnegativen Brustkrebs (BC)-Patientenkohorten, sowie von zwei unabhängigen alters-gematchten, gesunden Kontrollkohorten untersucht. BC-Kohorte 1 beinhaltet early-onset (EO) BC-Patientinnen. Kohorte 2 enthält BC-Patientinnen mit einem Risiko von >95% eine heterozygote Mutation in *BRCA1/BRCA2* (high-risk, HR) zu tragen. Allele mit >50% methylierten CpGs werden als funktionell relevante Epimutationen erachtet, da bekannt ist, dass TSG über eine Methylierung im Promotorbereich transkriptionell stillgelegt werden. Im Vergleich zu *ESR1* (\emptyset Methylierung, 3%), welches die Methylierungslevel eines durchschnittlichen Promotors widerspiegelt, zeigten die TSG sehr geringe durchschnittliche Methylierungswerte von weniger als 1%. Zudem waren die durchschnittlichen Epimutationsraten (EMR; <0,0001-0,1%) der TSG sehr gering. Mit der Ausnahme von *BRCA1*, welches eine erhöhte EMR in der BC-Kohorte verglichen zu den Kontrollen (0,31% gegen 0,06%) zeigte, gab es keine signifikanten Gruppenunterschiede zwischen BC-Patientinnen und Kontrollen. Eine von 36 HR BC-Patientinnen zeigte im Vergleich zu den restlichen Proben eine stark erhöhte EMR von 14,7% in *BRCA1*. Rund ein Drittel (15/44) der EO BC-Patientinnen wiesen eine erhöhte Rate an Einzel-CpG Fehlern in

mehreren TSG auf. Die nachfolgenden Expressionsanalysen ergaben eine erniedrigte Expression vieler TSG je analysierter Patientin. Diese Ergebnisse führen zu der Annahme, dass epigenetische Veränderungen in normalen Körperzellen als ein möglicher Indikator für einen gestörten Mechanismus, der für die Aufrechterhaltung des unmethylierten Status und der daraus resultierenden normalen Genexpression zuständig ist, angesehen werden können. Dies kann mit einem erhöhten BC-Risiko assoziiert werden.

SUMMARY

1st DNA methylation patterns of cortical neurons and glia in human and chimpanzee

The increasing complexity of social behavior along the ascending scale of primates, peaking in human spoken language, is accompanied by an impressive expansion of the human brain, particularly of the prefrontal cortex. Hence, prefrontal cortex appears to be one of the most interesting structures of the human brain, at least from an evolutionary perspective. But not only size but also function, in particular the interplay of neurons and glia cells, are suspected to distinguish the human brain from great apes and other primates. It is plausible to assume that proper brain function is controlled by a coordinated and well balanced transcriptional landscape, orchestrated by the underlying genetic and epigenetic backbone. Using reduced representation bisulfite sequencing (RRBS), we have compared the methylation profiles of neuronal and non-neuronal cells from three human and three chimpanzee prefrontal cortices (Brodmann area 10). Bioinformatic analyses revealed a genome-wide significant enrichment of differentially methylated regions (DMRs) in specific genomic areas. Intraspecific methylation differences between neuronal and non-neuronal cells are about three times more abundant than the interspecific methylation differences. More than 90% of human intraspecific DMRs were hypomethylated in neuronal cells, compared to non-neuronal cells. Intraspecific DMRs showed enrichment of genes associated with different neuropsychiatric disorders. Comparison between humans and chimpanzees yielded enrichments of genes showing human-specific brain histon modification. Interspecific DMRs were much more frequent in non-neuronal cells (n=666) than in neurons (n=96). Approximately 95% of interspecific DMRs in non-neuronal cells were hypermethylated in humans, compared to chimpanzees. It can be assumed that several hundreds of non-neuronal genes underwent a wave of methylation during human brain evolution. The impact of these changes in non-neuronal cells on the extension of the human brain may have been largely underestimated so far.

2nd Aberrant methylation patterns in the promoter region of tumor suppressor genes in breast cancer patients

The most prominent genetic cause for inherited breast and ovarian cancer are mutations in the *BRCA1* and *BRCA2* tumor suppressor genes (TSG). However, *BRCA1/BRCA2* germline mutations explain less than 50% of all familial breast cancers, even for women diagnosed before the age of 40 years. It has also been reported that epigenetic abnormalities, which contribute to changes in gene expression, play an important role in carcinogenesis and breast cancer development. To rapidly quantify the number of epimutations in different TSG, in both a qualitative and quantitative manner, we have developed a deep bisulfite sequencing assay targeting the promoter regions of eight TSG (*BRCA1*, *BRCA2*, *RAD51C*, *ATM*, *PTEN*, *TP53*, *MLH1* and *RB1*) and the estrogen receptor (*ESR1*) gene, which plays a role in tumor progression. We have analyzed blood samples of two independent *BRCA1/BRCA2*-mutation negative breast cancer (BC) cohorts and two independent age-matched healthy control cohorts. BC cohort 1 represents patients with early-onset BC and BC cohort 2 patients with a high risk to carry a heterozygous mutation. Since it is well known that tumor suppressor genes are transcriptionally silenced by promoter methylation, alleles with >50% methylated CpGs are considered as functionally relevant epimutations. Compared to *ESR1*, which is representative for an average promoter, TSG exhibited very low (<1%) average methylation levels and also very low mean epimutation rates (EMR; <0.0001% to 0.1%). With exception of *BRCA1*, which showed an increased EMR in BC (0.31% vs. 0.06%), there was no significant difference between patients and controls detectable. One of 36 HR BC patients showed a dramatically increased EMR (14.7%) in *BRCA1*. We identified in approximately one third (15 of 44) of EO BC patients increased rates of single CpG methylation errors in multiple TSG. Both EO and HR BC patients exhibited global underexpression of blood TSG. We propose that epigenetic abnormalities in normal body cells are indicative of disturbed mechanisms for maintaining low methylation and appropriate expression levels and may be associated with an increased BC risk.

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG.....	I
1. DNA-Methylierungsmuster corticaler Neuronen und Gliazellen im Menschen und Schimpansen.....	I
2. Epimutationscreening multipler Tumorsuppressorgene in mutationsnegativen early-onset und high-risk Brustkrebspatienten	II
SUMMARY	IV
1st DNA methylation patterns of cortical neurons and glia in human and chimpanzee	IV
2nd Aberrant methylation patterns in the promoter region of tumor suppressor genes in breast cancer patients	V
I Einleitung	6
1. Die Evolution des Menschen und sein Gehirn.....	6
1.1 Abstammungsgeschichte und Evolution	6
1.1.1 Die Erschaffung der Welt.....	6
1.1.2 Die Evolution der Primaten	7
1.1.3 Hominine Evolution	8
1.2 Faszination menschliches Gehirn	6
1.2.1 Aufbau des menschlichen Gehirns	6
1.2.2 Zelltypen des menschlichen Gehirns.....	8
1.2.2.1 Neuronale Zellen	8
1.2.2.2 Nicht-neuronale Zellen	9
1.2.3 Gehirngröße und Intelligenz.....	11
2. Epigenetik.....	13
2.1 DNA-Methylierung.....	14
2.1.1 Methylom und Genom	16
2.1.2 DNA-Methylierung im Gehirn und Plastizität.....	18
2.1.2 Methylierung und Krebs.....	19
3. Krebserkrankungen	21
3.1 Tumorgenese.....	22
3.2 Brustkrebs.....	23
3.2.1 Einteilung in Subtypen.....	25
3.2.2 Molekulare Ursachen	26
3.2.2.1 Genetische Faktoren	27
3.2.2.2 Methylierungsveränderungen und Brustkrebs	29
4. Zielsetzung der Arbeit	33

4.1	DNA-Methylierungsanalysen corticaler Neuronen und Gliazellen im Menschen und Schimpansen	33
4.2	Methylierungsanalysen zur Detektion von Epimutationen in Brustkrebspatienten.....	34
II	MATERIAL	35
1.	Probensammlung	35
1.1	Methylierungs- und Expressionsanalysen bei Brustkrebspatientinnen	35
1.1.1	Vergleichende Methylierungsanalysen	35
1.1.2	Expressionsanalysen.....	36
1.2	Vergleichende Methylierungsanalysen zwischen Mensch und Schimpanse	37
1.2.1	Präfrontales Cortexgewebe.....	37
2.	Primer und Adapter	38
3.	Kommerzielle Kit-Systeme.....	43
4.	Reagenzien und Enzyme.....	43
5.	Chemikalien und Puffer/Lösungen	44
6.	Geräte.....	45
7.	Software und Analysetools.....	46
III	METHODEN	48
1.	Isolation genomischer DNA und RNA-Extraktion	48
1.1	DNA-Extraktion	48
1.1.1	Blutproben.....	48
1.1.2	Neuronale und nicht-neuronale Zellen	48
1.1.3	Präfontales Cortexgewebe	48
1.2	RNA-Extraktion	49
2.	Natriumbisulfitbehandlung genomischer DNA	49
2.1	Brustkrebspatientenproben	50
2.2	gDNA der neuronalen und nicht-neuronalen Zellen für die Datenvalidierung.....	51
2.3	gDNA aus dem präfrontalen Cortex-Gesamtgewebe für die Datenvalidierung.....	51
3.	Next-Generation Sequencing	51
3.1	Deep Bisulfite Amplicon Sequencing.....	51
3.1.1	Genspezifische PCR.....	52
3.1.2	Aufreinigung mittels magnetischen Beads.....	54
3.1.2.1	Manuelle Aufreinigung.....	55
3.1.2.2	Automatisierte Aufreinigung.....	56
3.1.3	Quantifizierung der PCR-Produkte	57
3.1.4	Library-Präparation	58

3.1.4.1	Herstellung der Pools	58
3.1.4.1.1	Verdünnen der PCR-Produkte	58
3.1.4.1.2	Pooling der verdünnten PCR-Produkte	59
3.1.4.2	Aufreinigung mittels magnetischen Beads.....	59
3.1.4.3	Modifikation der Fragmentenden	59
3.1.4.4	Ligation der Adapter.....	60
3.1.4.5	Aufreinigung der Ligation via magnetischen Beads	61
3.1.4.6	PCR zur Überprüfung der Ligations-Effizienz.....	61
3.1.4.7	Finale Barcoding-PCR.....	62
3.1.4.8	Aufreinigung der finalen PCR mittels magnetischen Beads	63
3.1.4.9	Quantifizierung der finalen PCR	63
3.1.4.10	Verdünnung und finales Pooling	64
3.1.5	Sequenzierung via MiSeq	64
3.2	Reduced Representation Bisulfite Sequencing	66
3.2.1	MspI Restriktionsverdau.....	67
3.2.2	Modifikation der Fragmentenden	68
3.2.3	DNA-Aufreinigung mittels magnetischen Beads	68
3.2.4	Multiplexed Adapter-Ligation	68
3.2.5	DNA-Aufreinigung und Determination der Fragmentlängen	69
3.2.6	Ligations-Effizienz PCR.....	69
3.2.7	Natriumbisulfitbehandlung	70
3.2.8	Aufreinigung mittels Säulchen	70
3.2.9	Bestimmung der optimalen Zyklenzahl.....	71
3.2.10	Finale Amplifikation der DNA-Fragmente	71
3.2.11	DNA-Aufreinigung mittels magnetischen Beads	72
3.2.12	Quantifizierung der RRBS Library	73
3.2.13	Sequenzierung der RRBS Libraries.....	73
4.	Datenanalyse und Statistik.....	74
4.1	Auswertung der NGS-Daten	74
4.1.1	DBAS	74
4.1.2	RRBS.....	75
4.1.2.1	Datenanalyse	75
4.1.2.2	Enrichmentanalysen	76
4.2	Statistische Analyse	77
5.	Datenvalidierung via Pyrosequenzierung.....	78

5.1	Genspezifische PCR.....	78
5.2	Pyrosequenzierung.....	78
6.	Genexpression.....	80
IV	ERGEBNISSE	83
1.	Genomweite Methylierungsanalysen im präfrontalen Cortex via RRBS.....	83
1.1	Analyse der RRBS-Daten.....	83
1.1.1	Intra- und interspezifische differenziell methylierte Regionen.....	86
1.1.2	Gene Ontology Term Analysen.....	91
1.1.3	DMRs und neuropsychiatrische Erkrankungen	94
1.1.4	DMRs und human-spezifische Expressions- und Histonmethylierungs-Signaturen.....	97
1.1.5	Repetitive Elemente assoziiert mit DMRs	100
1.2	Validierung der DMRs via Bisulfit-Pyrosequenzierung.....	106
2.	Methylierungsanalysen bei Brustkrebs via DBAS.....	109
2.1	Analyse der DBAS-Daten	109
2.1.1	DNA-Methylierungsprofile von TSG	110
2.1.2	Epimutationsraten in TSG.....	112
2.1.3	Hypermethylierung auf Einzel-CpG-Ebene.....	118
2.1.4	Kummulative Promotor-Hypermethylierung	128
2.2	Expressionsanalysen für ausgewählte Proben und Gene	131
V	DISKUSSION	134
1.	Identifizierung differenziell methylierter Regionen im präfrontalen Cortex	134
1.1	RRBS zur Detektion aberranter Methylierungsmuster	135
1.2	Intraspezifisch differenziell methylierte Regionen.....	136
1.2.1	Kandidatengenanalysen in intraspezifischen DMRs.....	137
1.2.1.1	Neuropsychiatrische Erkrankungen	137
1.2.2	Repetitive Elemente in intraspezifischen DMRs	138
1.3	Interspezifisch differenziell methylierte Regionen.....	140
1.3.1	Kandidatengenanalysen in interspezifischen DMRs.....	141
1.3.1.1	Neuropsychiatrische Erkrankungen	141
1.3.1.2	Human-spezifisch regulierte Gene	142
1.4	Schlussfolgerung und Ausblick	144
2.	Methylierungsanalysen von TSG bei Brustkrebspatientinnen	146
2.1	DBAS zur Detektion aberranter Methylierungsmuster	146
2.2	Methylierungslevel der Promotorbereiche ausgewählter TSG	147

2.3	Analyse der Epimutationen in TSG	149
2.3.1	Biologische Aspekte für die Entstehung der Epimutationen.....	150
2.3.2	Störfaktoren — Datenanalyse und technische Artefakte	151
2.4	Einzel-CpG-Fehler	153
2.5	Kummulative Promotor-Hypermethylierung	155
2.6	Expressionsanalysen.....	157
2.7	Schlussfolgerung und Ausblick	159
	REFERENZEN	161
1.	Literaturangaben.....	161
2.	Internetquellen.....	179
	VERZEICHNISSE	182
1.	Abkürzungsverzeichnis	182
2.	Abbildungsverzeichnis.....	185
3.	Tabellenverzeichnis	186
	PUBLIKATIONEN & KONGRESSBEITRÄGE	187
1.	Publikationen.....	187
2.	Kongressbeiträge.....	189
2.1	Erstauthor.....	189
2.2	Ko-Author	189
	DANKSAGUNG	191
	EIDESSTÄTTLICHE VERSICHERUNG	192
	ANHANG	193

I EINLEITUNG

1. Die Evolution des Menschen und sein Gehirn

1.1 Abstammungsgeschichte und Evolution

1.1.1 Die Erschaffung der Welt

„¹ Im Anfang erschuf Gott Himmel und Erde. [...] ²⁰ Dann sprach Gott: Das Wasser wimmele von Schwärmen lebendiger Wesen und Vögel sollen über der Erde am Himmelsgewölbe fliegen. ²¹ [...] Gott erschuf die großen Wassertiere und alle Lebewesen [...] ²⁴ Dann sprach Gott: Die Erde bringe Lebewesen aller Art hervor, von Vieh, von Kriechtieren und von Wildtieren der Erde nach ihrer Art. Und so geschah es. ²⁷ Dann sprach Gott: Lasst uns Menschen machen als unser Bild, uns ähnlich! Sie sollen walten über die Fische des Meeres, über die Vögel des Himmels, über das Vieh, über die ganze Erde und über alle Kriechtiere, die auf der Erde kriechen. ³¹ Gott sah alles an, war er gemacht hatte: Und siehe, es war sehr gut. Es wurde Abend und es wurde Morgen: sechster Tag.“

Auszug aus *Genesis 1,1* [1]

Mit diesen Worten erklärt die Bibel in der sogenannten Schöpfungsgeschichte die Erschaffung der Welt, sowie die Entstehung der Lebewesen und des Menschen – lange Zeit die einzige Erklärung für das menschliche Dasein auf der Erde. Die Etablierung der ersten Evolutionstheorien begann im 19. Jahrhundert. Zu den auch heute noch bekanntesten Evolutionstheoretikern zählen Jean-Baptiste de Lamarck, Charles Darwin und Alfred R. Wallace. Die bekannteste Theorie stammt wohl von Darwin, der das Prinzip „survival of the fittest“ prägte, welches das Überleben derer beschreibt, die am besten an die Umweltbedingungen angepasst sind. Diese Individuen geben dann wiederum ihre Gene an die nächste Generation weiter, wodurch es, über einen Zeitraum mehrerer Generation, zu einer natürlichen Selektion und zur Entstehung neuer Arten kommt [2, 3]. Bis heute ist die Evolutionstheorie einem stetigen Wandel unterworfen und es gibt nicht die eine Theorie, welche die Evolution in einer gerechten Weise erklären kann. Darwins Theorie von Variation und natürlicher Selektion wurden bis heute nicht verworfen, sondern eher ergänzt bzw. erweitert [2].

1.1.2 Die Evolution der Primaten

Die Primaten werden der Klasse der Säugetiere zugeordnet und teilen sich in die Unterklassen der Feuchtnasenprimaten (Strepsirrhini) und Trockennasenprimaten (Haplorrhini) auf. Letztere beinhaltet die Überfamilie der Menschartigen (Hominoidea), woraus die Familie der Menschenaffen (Hominidae) und die Unterfamilie der Homininae hervorgehen. Die hominine Evolution hin zum modernen Menschen begann vor circa fünf bis sieben Millionen Jahren (MYA) durch die Abspaltung eines gemeinsamen Vorfahren zwischen den Hominini und den Panini (Abbildung 1). Somit zählen der Schimpanse und die Bonobos zu unseren nächsten Verwandten innerhalb der Primaten [4].

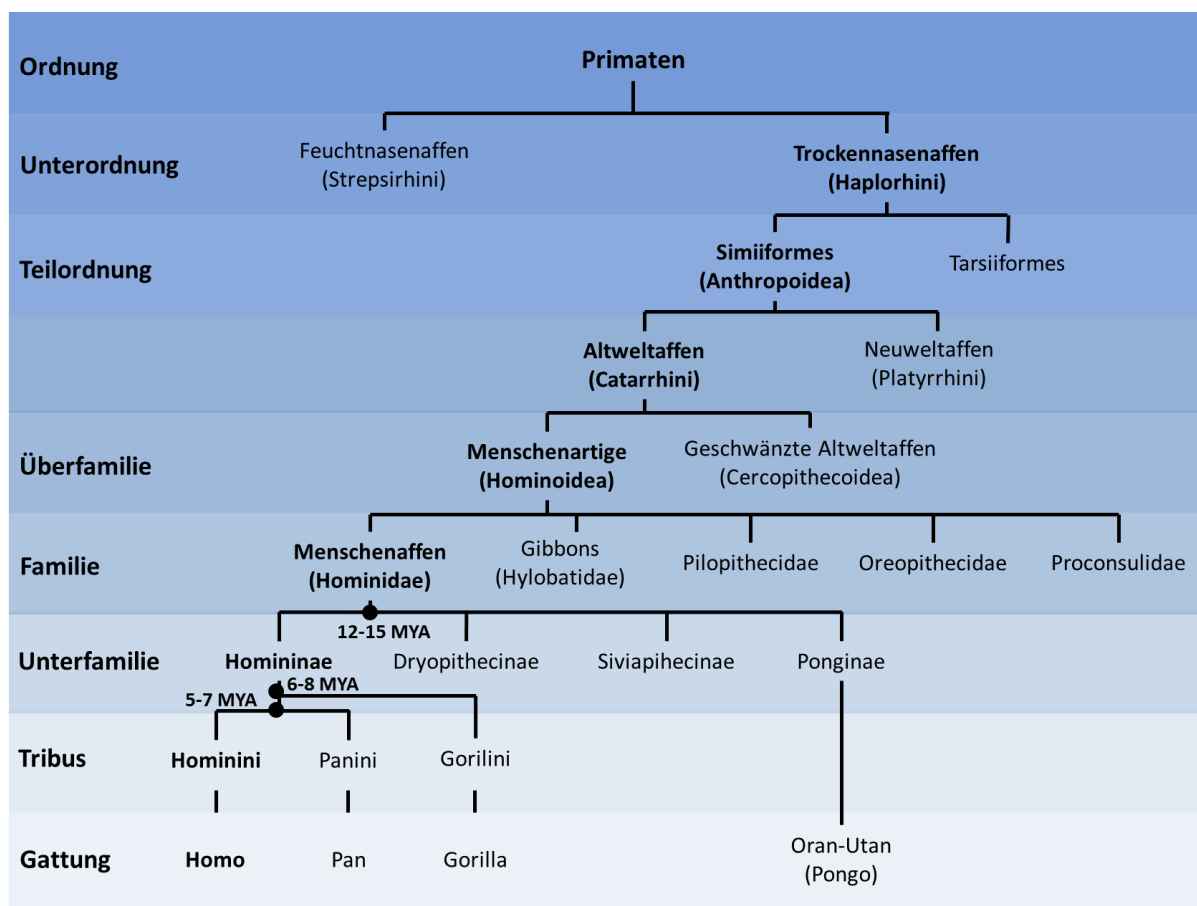


Abbildung 1: Taxonomische Klassifizierung der Primaten und phylogenetische Beziehung rezenter großer Primaten. Die Abbildung zeigt einen stark verkürzten Stammbaum der Primaten. Es sind ausschließlich die Aufspaltungen der Primaten bis hin zur Unterfamilie der Homininae aufgeführt (Abb. erstellt nach [5] und ergänzt durch Jahreszahlen, entnommen aus [4]).

1.1.3 Hominine Evolution

Nach der Spaltung der *Homininae* durchlief der Tribus der *Hominini* (Menschenartigen) eine bemerkenswerte Entwicklung bis hin zum modernen Menschen. Einige tausend Fossilienfunde aus verschiedenen Zeitperioden zeigen an, dass diese Entwicklung nicht geradlinig stattfand (Abbildung 2). Es gab verschiedene Seitenzweige, die alle im Laufe der Evolution ausgestorben sind. Auch unter Berücksichtigung neuester Funde in Asien, geht eine Vielzahl von Forschern nach wie vor davon aus, dass der Ursprung des modernen Menschen in Afrika liegt [6].

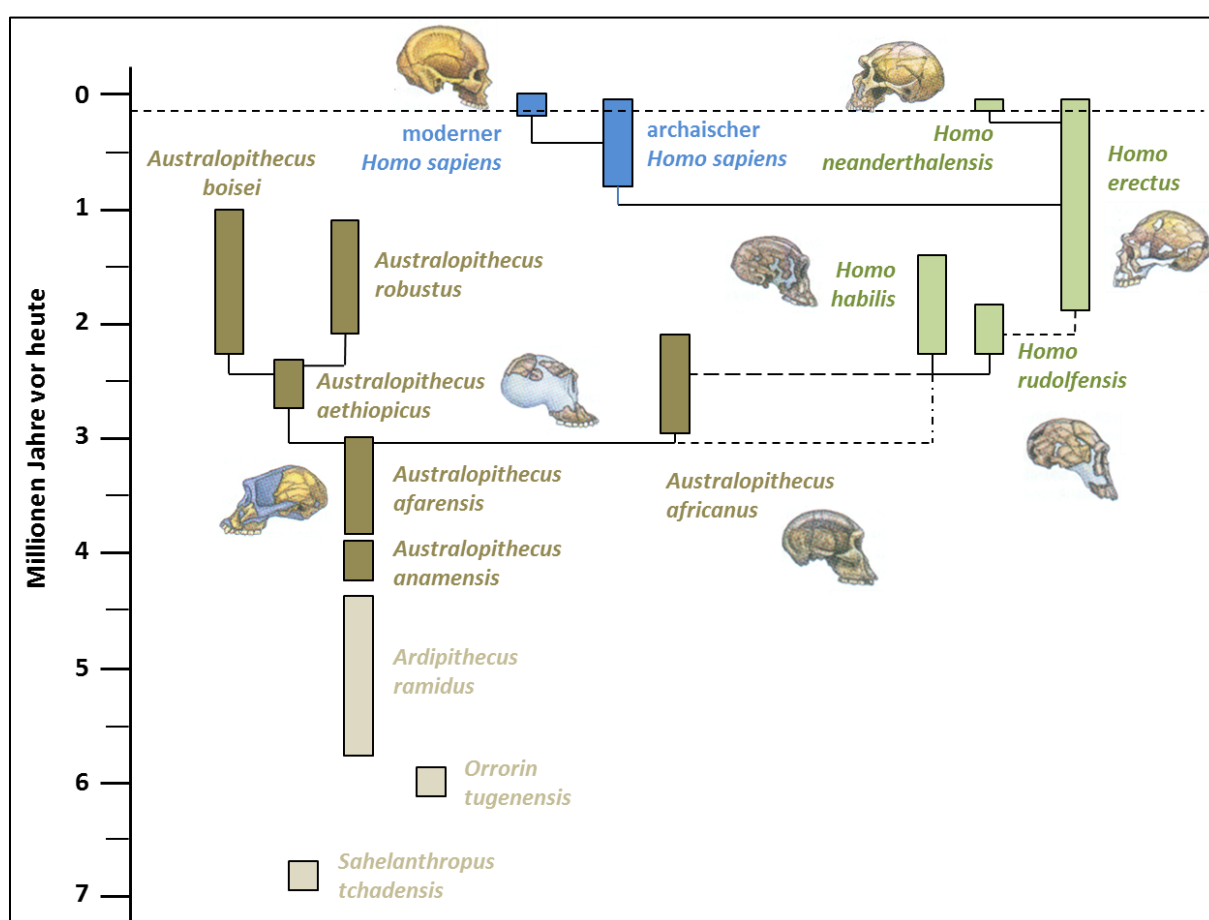


Abbildung 2: Stammbaum der Hominini. Dargestellt ist eine vereinfachte und gekürzte Version des homininen Stammbaums (Abb. erstellt nach dem Vorbild von [7, 8], ergänzt mit Schädelbildern aus [9]).

Zu den ersten Menschenartigen zählen *Sahelanthropus tchadensis* (Fundort: Toros-Menalla, Djurab Wüste, nördlicher Tschad) [10, 11], *Orrorin tugenensis* (Fundort: Lukeino Formation, Tugen Berge, Kenia) [11, 12] und *Ardipithecus ramidus* (Fundort: Aramis, Middle Awash, nördliches Äthiopien) [11, 13]. In der Folge entwickelte sich die Gattung der

Australopithecinen vor rund 4-5 MYA. Die Schädelproportionierung dieser „Südaffen“ war noch ursprünglich, jedoch weisen Skelettfunde auf einen aufrechten Gang hin. Der Ursprung der Gattung *Homo* ist nach wie vor unter Paläoanthropologen umstritten. Jedoch wird davon ausgegangen, dass *Homo rudolfensis* (Fundort: Koobi Fora, Lake Turkana, nördliches Kenia) [14] und *Homo habilis* (Fundort: Olduvai Gorge, Tanzania) [15] die ältesten Menschenarten repräsentieren. Diese beiden Arten lebten zeitgleich mit einigen Australopithecinen-Arten (Fundorte: Ostafrika; Äthiopien, Kenia, Tanzania) [16] vor rund 2,5-1,6 MYA. In diesem Zeitraum wird auch die Entstehung des *Homo erectus* (Fundort: Java, Indonesien) [17] angesiedelt, der als Vorfahr des *Homo neanderthalensis* (Fundorte: Neandertal, Deutschland) [18] und des archaischen *Homo sapiens* gilt. Der exakte Verlauf der homininen Evolution im Zeitraum der letzten 2 MYA ist unter den Wissenschaftlern stark diskutiert. Jedoch sind die Veränderungen der körperlichen Merkmale der Hominidenevolution durch Fossilienfunde gut belegt. So weiß man, dass der aufrechte Gang bereits früh (vor ca. 7 MYA, *Sahelanthropus tchadensis*) in der homininen Evolution entstand und konsequenterweise zu einer Verlängerung der Beine und längerfristig zu weiteren Veränderungen am Skelett führte, besonders im Bereich des Beckens, der Füße und der Wirbelsäule. Des Weiteren durchlief der Schädel eine starke Veränderung. Der Gesichtsschädel wurde durch die Rückbildung des Kauapparates verkleinert, wohingegen der Gehirnschädel an Masse stark zunahm. Dadurch kam es zu einer Steigerung des Hirnvolumens von circa 500 cm³ bei *Homo habilis* und *Homo rudolfensis*, auf 1500 cm³ beim archaischen *Homo sapiens*. Nachdem die Hände nicht weiterhin zur Fortbewegung benutzt werden mussten, konnten sie andere Tätigkeiten, wie den Gebrauch von Werkzeugen und das Tragen dergleichen über weite Entfernungen hin übernehmen. Dadurch konnten neue und energiereiche Nahrungsquellen erschlossen werden. Die bessere Versorgung des Gehirns mit Energie trug entscheidend dazu bei, dass sich das Gehirn vergrößern konnte und sich immer komplexere Fähigkeiten herauskristallisiert haben. Das Gehirn eines adulten modernen Menschen macht ca. 2% des gesamten Körpergewichts aus und verbraucht rund 25% an Energie im Ruhezustand (RMR, *resting metabolic rate*) [7]. Der zeitliche Verlauf dieser körperlichen Veränderungen kann der nachfolgenden Tabelle 1 entnommen werden.

Tabelle 1: Körperliche Merkmale und Besonderheiten der Hominiden. Die Tabelle ist dem Lehrbuch für Ökologie und Evolution entnommen und in einer gekürzten Version dargestellt [7]. Der Übersicht halber werden nur Angaben zu den homininen Arten gemacht, welche in Abbildung 2 dargestellt sind.

Art	Zeit (MYA)	Hirnvolumen (cm ³)	Schädelmerkmale	Besonderheiten
<i>Sahelanthropus tchadensis</i>	ca. 7	ca. 380	kleine Eckzähne, fehlende Zahnlucke	evtl. schon aufrechter Gang
<i>Orrorin tugenensis</i>	ca. 6	unbekannt	Zähne relative ursprünglich	„
<i>Ardipithecus ramidus</i>	4,4	„	„	fähig zum aufrechten Gang
<i>Australopithecus afarensis</i>	3,6-2,9	375-550	großer Gesichts-, kleiner Hirnschädel; fast geschlossene Zahnreihe	aufrechter Gang
<i>Australopithecus africanus</i>	3,0-2,3	bis 500	großer Gesichts-, kleiner Hirnschädel; geschlossene Zahnreihe	„
<i>Australopithecus aethiopicus</i>	2,6-2,3	420	großer, breiter Gesichtsschädel, massiver Unterkiefer, flacher Hirnschädel mit Scheitelkamm	„
<i>Australopithecus boisei</i>	2,3-1,4	450-545	„	„
<i>Australopithecus robustus</i>	2,0-1,5	475-530	„	„, Gebrauch von Knochen als Werkzeug
<i>Homo habilis</i>	2,4-1,6	500-650	Gesichtsschädel unter dem Hirnschädel, Hirnschädel mit Stirn	erste Steinwerkzeuge, einfache sprachliche Kommunikation?
<i>Homo rudolfensis</i>	2,5-1,6	600-800	ähnlich H. habilis, große Molaren	„
<i>Homo erectus</i>	1,8-0,5	800-1200	dicke Schädelknochen, flache Stirn, flacher Schädel, Überaugenwülste, kein Kinn	Nutzung von Feuer, Ausbreitung nach Europa und Asien, einfache Sprache
<i>Homo neanderthalensis</i>	0,12-0,03	ca. 1500	Schädel lang, flach und massig; reduzierte Überaugenwülste	Sozialstruktur, Leben in Gruppen
<i>Homo sapiens archaisch</i>	0,5-0,16	1100-1400	reduzierte Überaugenwülste, fliehendes Kinn, flacher Schädel	
<i>Homo sapiens modern</i>	0,16-heute	1200-1700	keine Überaugenwülste, hohe Stirn, höherer runder Schädel	Beeinflussung der eigenen Entwicklung

1.2 Faszination menschliches Gehirn

1.2.1 Aufbau des menschlichen Gehirns

Das menschliche Gehirn setzt sich aus dem Großhirn, dem Kleinhirn und dem Hirnstamm zusammen. Das Großhirn besteht aus einer rechten und einer linken Hemisphäre und wird in vier Lappen untergliedert: Frontal-, Parietal-, Temporal- und Okzipitallappen. Im Großhirn werden die höheren Funktionen, wie das Interpretieren von Berührung, Sehen und Hören, sowie Sprache, logisches Denken, Emotionen, Lernen und Kontrolle von Bewegungen ausgeübt bzw. gesteuert. Die äußerste Schicht des Großhirns wird als Cortex bezeichnet. Dieser enthält rund 70% aller Nervenzellen und wird aufgrund der gräulichen Färbung auch als graue Substanz bezeichnet. Der Cortex weist eine gefaltete Struktur auf, jede Faltung wird als Gyrus bezeichnet und jede Furche als Sulcus. Unterhalb des Cortex liegen die Axone, welche mit einer Myelinschicht überzogen sind und deshalb als weiße Substanz bezeichnet werden [19].

Der Neocortex gilt als Zentrale der höheren kognitiven Funktionen und stellt den stammesgeschichtlich jüngsten Teil der Großhirnrinde bei Säugern dar [20]. Seit der Entstehung des Neocortex vor rund 220 MYA, durchlief dieser eine extreme Vergrößerung, was eine positive Selektion durch die Erhöhung der Anzahl an Neuronen widerspiegelt. Im Menschen macht dieser zwei Drittel der gesamten Gehirnmasse aus und beinhaltet rund 16 Milliarden neuronale Zellen [21]. Der zugrundeliegende Mechanismus, welcher zu dieser Expansion beitrug, könnte eine veränderte Neurogenese sein. Unter Neurogenese versteht man den Bildungsprozess neuronaler Zellen aus unterschiedlichen neuronalen Vorläuferzellen während der Entwicklung eines Individuums. Diese Unterschiede betreffen primär die Wachstumskapazität, die Menge und die Abstammungsbeziehungen bestimmter Typen von neuronalen Vorläuferzellen [20].

Der präfrontale Cortex (PFC) ist ein Teil des Neocortex und ist der am weitesten vorne liegende Teil des Frontallappens. Dieser setzt sich aus dem stirnwärts gelegenen Pol des Gehirns (Brodmann-Areale (BA) 9, 10, 46), dem Orbitallappen (BA 11, 12) und den medianen Feldern (BA 32) zusammen. Der PFC ist reziprok mit dem Thalamus und dem limbischen System verbunden und beinhaltet konvergente Projektionen der sensorischen Bereiche. Der PFC ist zuständig für die höheren kognitiven Prozesse unter Einbezug von Emotionen und hat somit eine große Bedeutung für die Persönlichkeit [22]. Läsionen im dorsolateralen Teil, das BA 10 eingeschlossen, sind mit Störungen höherer kognitiver Funktionen verbunden, wie

dem Extrahieren von Bedeutungen aus permanenten Erfahrungen, der Organisation mentaler Inhalte, die das kreative Denken und die Sprache kontrollieren, sowie der Umsetzung von kreativem Denken und der Planung zukünftiger Aktionen [23].

Das im PFC gelegene BA 10 nimmt rund 25-30 cm³ ein und ist beim Menschen im Vergleich zu den Menschenaffen und anderen Lebewesen erheblich größer [24-26] (Abbildung 3). Neben diesen Fakten weist auch die Struktur dieses Areals und die niedrigere Zelldichte im Menschen, verglichen zu den Affen und Menschenaffen, darauf hin, dass das BA10 wichtig für die menschlichen kognitiven Fähigkeiten ist. Es wird angenommen, dass durch die niedrige Zelldichte mehr Raum für Verbindungen zu den übergeordneten Assoziationsregionen zur Verfügung steht [24]. Dieser Gedanke wird durch die Anzahl an dendritischen Dornen pro Zelle und der Dornendichte, welche höher im Vergleich zu anderen corticalen Regionen ist, unterstützt [24, 27]. Ein weiterer Grund, der für die Rolle des BA10 bei kognitiven Funktionen spricht, ist, dass der rostrale PFC eine sehr späte Maturation während der ontogenetischen Entwicklung erfährt. Diese Region wird zuletzt myelinisiert und zeigt beim Menschen die höchste Wachstumsrate erst im Alter von 5 bis 11 Jahren [28, 29].

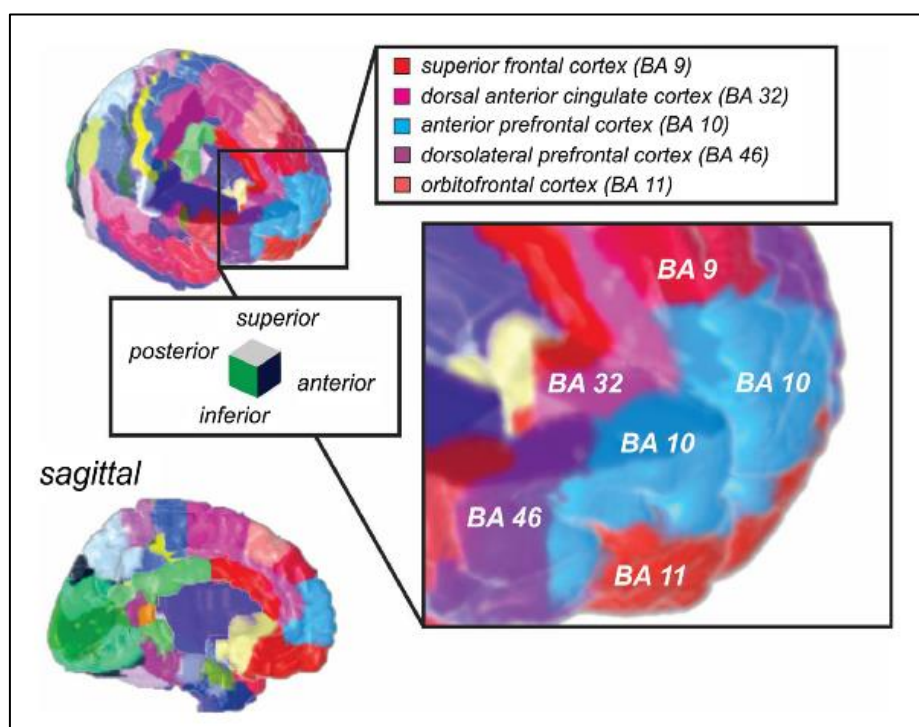


Abbildung 3: Lage des Brodmann-Areals 10 im präfrontalen Cortex. (Abbildung entnommen aus [30])

1.2.2 Zelltypen des menschlichen Gehirns

1.2.2.1 Neuronale Zellen

Die neuronalen Zellen besitzen einzigartige strukturelle Eigenschaften und können über elektro-chemische Signale Informationen prozessieren und weitergeben. Obwohl sich die verschiedenen neuronalen Zelltypen in ihrer Morphologie unterscheiden, weisen sie alle einen Zellkörper, Dendriten, Axone und axonale Enden auf. Die Axone dienen der Informationsweiterleitung zwischen den Zellen und die Dendriten dem Empfangen von Informationen. Die Kommunikation mit anderen Neuronen findet über das Aussenden von Neurotransmitterstoffen an den Synapsen, zwischen den Axonen und Dendriten benachbarter Neurone, statt. Die Neurone sind in einem Zentralnervensystem (ZNS), dem Gehirn, und einem Längsnervenstrang organisiert. Neurone, die die Information vom ZNS hin- bzw. weggleiten, bilden das periphere Nervensystem. Grundsätzlich werden die Neurone in Interneurone, Motorneurone und sensorische Neurone untergliedert (Abbildung 4). Jede dieser Klassen besteht wiederum aus hunderten von verschiedenen Zelltypen mit spezifischen Fähigkeiten. [31, 32] [Int. 1].

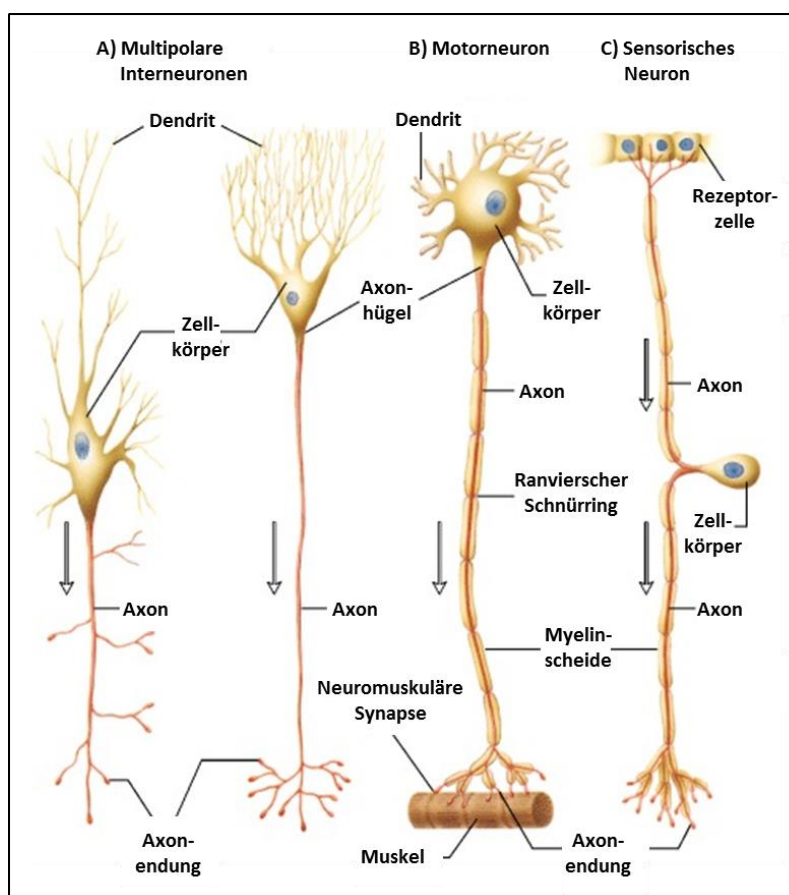


Abbildung 4: Schematische Darstellung der typischen Neuronen bei Säugern. **A)** Multipolare Interneuronen besitzen viele verzweigte Dendriten, die Signale von mehreren hundert anderen Neuronen über ihre Synapsen erhalten, und ein einzelnes langes Axon, das lateral verzweigt ist. **B)** Motorneurone sind für die Innervierung von Muskelzellen zuständig. Sie haben typischerweise ein einzelnes langes Axon, welches sich vom Zellkörper bis hin zur Effektorzelle erstreckt. **C)** Bei sensorischen Neuronen verzweigen sich die Axone erst nach Verlassen des Zellkörpers. Die periphere Verzweigung leitet die Nervenimpulse von der Rezeptorzelle zum Zellkörper, der in den Hinterwurzelganglien nahe dem Rückenmark lokalisiert ist. Die zentrale Verzweigung leitet die Impulse vom Zellkörper zum Rückenmark oder Gehirn weiter. Die Pfeile geben die Leitrichtung des Aktionspotentials im Axon an. (Abb. entnommen aus [31] und modifiziert)

Obwohl die Mehrheit der Neuronen bereits nach der Geburt vorhanden ist, wird angenommen, dass es auch im postnatalen Gehirn zur Bildung neuer Neurone kommt. Die sogenannte Neurogenese spielt sich in Regionen mit einer hohen Konzentration an neuronalen Vorläuferzellen ab. Die dort gebildeten Neurone werden dann entweder über radiale Gliazellen zu den Zielbereichen transportiert oder finden ihren Weg über eigens produzierte chemische Signale. Fehler bei der Migration der Neurone sind mit verschiedenen Krankheiten, wie kindlicher Epilepsie, Schizophrenie oder der Lernschwäche Dyslexie assoziiert [Int. 1]. Während der Entwicklung des Nervensystems steht das Wachstum der neuronalen Elemente im Vordergrund, jedoch kommt es auch natürlicherweise zum Zelltod während der Entwicklung des Nervensystems. Diese nicht-pathologischen Ereignisse spielen eine wichtige Rolle bei der Etablierung der komplexen Netzwerke bei der Gehirnentwicklung [33]. Ein unkontrollierter Zelltod im adulten Gehirn kann jedoch zur Ausbildung von Krankheiten, wie Morbus Parkinson, Alzheimer-Krankheit oder Chorea Huntington führen [Int. 1]. Für das effiziente Funktionieren der Neurone benötigen diese die Hilfe der Gliazellen, die, je nach Typ, u.a. der Versorgung der Neurone dienen, zur Isolation der Axone beitragen und die Zusammensetzung der extrazellulären Flüssigkeit kontrollieren [32].

1.2.2.2 Nicht-neuronale Zellen

Die nicht-neuronalen Zellen setzen sich aus Glia-, Ependym- und Endothelzellen sowie Perizyten und Vorläuferzellen zusammen. Gliazellen werden grundsätzlich in Mikroglia und Makroglia unterteilt. Die Makroglia umfassen zudem die Astrozyten, Tanyzyten, Oligodendrozyten und Ependymzellen (Abbildung 5) [34].

Die Mikroglia sind die primären Immunzellen des Zentralnervensystems und werden im Gegensatz zu den anderen Gliazellen im Mesoderm gebildet und nicht im Neuroektoderm.

Sie unterscheiden sich aber auch von immunologischen Zellen, da Mikroglia sich selbstständig erneuern können und nicht auf myeloide Vorläufer angewiesen sind [35]. Sie übernehmen die Phagozytose zellulärer Ablagerungen, beschädigter Zellen, Plaques und fremdartiger Substanzen [36]. Einige Subtypen besitzen die Fähigkeit, neuronale Wachstumsfaktoren, Neurotrophine und neurotrophine Faktoren freizusetzen [37]. Es wird zudem angenommen, dass Mikroglia in die Regulation der Neuronenzahl und deren Entwicklung involviert sind [38-42], sowie die Angiogenese während der Entwicklung fördern [38, 43-45]. Mikroglia können die neuronale Aktivität durch das Fördern der synaptischen Plastizität und die Freisetzung neurotrophischer Faktoren und anti-inflammatorischer Cytokine modulieren [46-48].

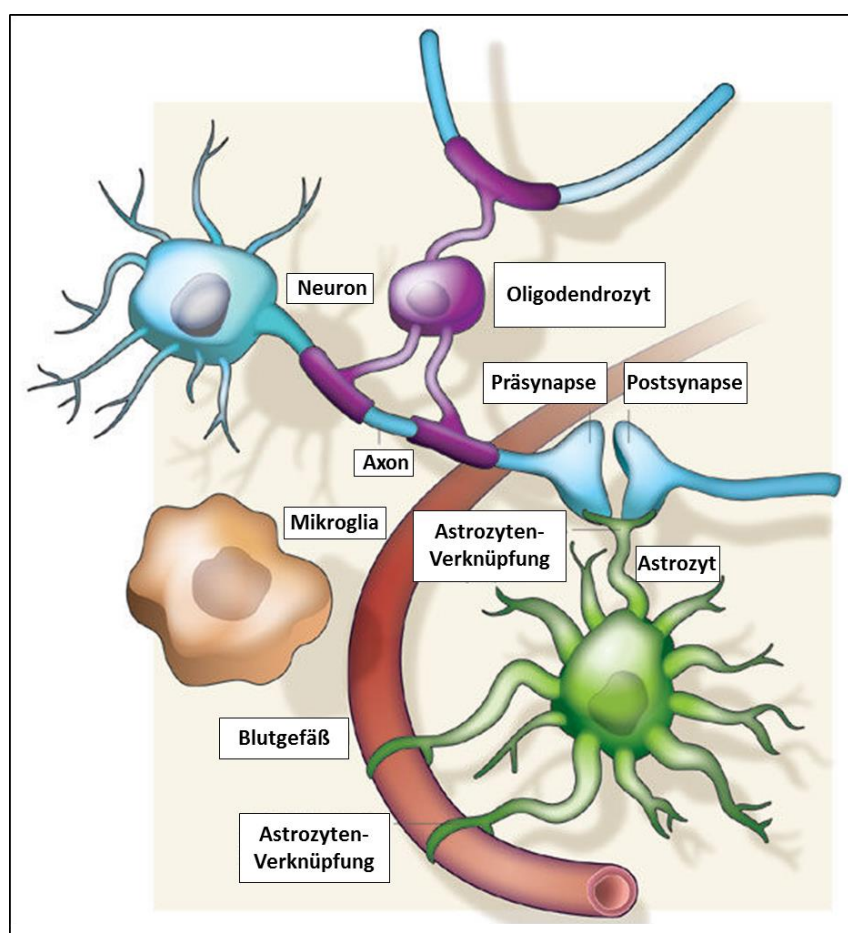


Abbildung 5: Schematische Darstellung verschiedener Gliazelltypen. In der Abbildung sind ausgewählte nicht-neuronale Zelltypen und deren Interaktion mit einem Neuron und einem Blutgefäß dargestellt. Die Oligodendrozyten ummanteln die Axone mit Myelin. Die Astrozyten umhüllen die Blutgefäße und Synapsen. Die Mikroglia dienen u.a. dem Schutz des Gehirns vor Infektionen. (Abb. entnommen aus [49] und modifiziert)

Frühere Studien, die dazu dienten, ein besseres Verständnis für die Funktionsweise des Gehirns zu erlangen, legten den Fokus auf die neuronalen Zellen, da angenommen wurde, dass nicht-neuronale Zellen nur eine unterstützende Funktion gegenüber den Neuronen besitzen. Es konnten jedoch im Laufe der Zeit neue Erkenntnisse gewonnen werden, die klar zeigen, dass die nicht-neuronalen Zellen eine wichtige Rolle bei der Entwicklung des Gehirns einnehmen und an der Neurogenese, Synaptogenese und der Aufrechterhaltung der synaptischen Plastizität bei metabolischen Kreisläufen beteiligt sind. Einige Gliazellen, wie beispielsweise Tanyzyten und Astrozyten, transportieren zirkulierende Nährstoffe und metabolische Faktoren, die entscheidend für das neuronale Überleben und ihre Aktivität im Hypothalamus sind. Die nicht-neuronalen Zellen exprimieren Rezeptoren für verschiedene metabolische Faktoren und Hormone, was zu der Annahme führt, dass diese Zellen eine metabolische Funktion einnehmen [34].

1.2.3 Gehirngröße und Intelligenz

Das Gehirn des Menschen stellt eine Besonderheit unter den Gehirnen der Säuger dar. Es ist zu den höchsten kognitiven Leistungen fähig, größer als die Körpergröße vermuten lässt, ausgestattet mit einem überentwickelten cerebralen Cortex, der mehr als 80% der Gehirnmasse ausmacht, und beinhaltet 100 Millionen Neuronen und zehnmal mehr Gliazellen [49-54]. Da das menschliche Gehirn offensichtlich nicht das Größte unter den Lebewesen auf der Erde ist, können die überlegenen kognitiven Fähigkeiten nicht über die Gehirngröße erklärt werden. Im Vergleich zu anderen Primaten und Menschenaffen besitzt die menschliche Spezies einen kleineren Körper und ein größeres Gehirn, welches um ein fünf- bis sieben-faches zu groß ist, gemessen an der Körpergröße [55, 56]. Die Anzahl an Neuronen und deren relative Menge in verschiedenen Gehirnregionen führt zu dem Schluss, dass diese ein bestimmender Faktor für die neurale Funktion und das Verhalten ist [57]. Bei Säugern zeigen die Spezies mit den größten Gehirnen auch eine größere Bandbreite bezüglich des Verhaltens – verglichen zu denen mit einem kleineren Gehirn [58, 59]. Es findet sich jedoch keine Korrelation, wenn man die absolute Gehirngröße und die kognitiven Fähigkeiten bei Spezies innerhalb der gleichen Ordnung mit einer ähnlichen Gehirngröße miteinander vergleicht. Beispielsweise besitzen Affen sehr viel kleinere Gehirne als die Huftiere (vier- bis fünf-fach größer), weisen aber höhere kognitive Fähigkeiten und eine höhere Flexibilität im Verhalten auf [60].

Der Enzephalisierungsquotient (EQ) gibt die Abweichung der beobachteten Gehirngröße einer Spezies zur erwarteten Körpergröße an, wobei sich der Erwartungswert an den Durchschnitten eines Taxons orientiert. Ein EQ von 1 gibt an, dass die beobachtete mit der erwarteten Gehirngröße übereinstimmt und ein EQ von >1 bedeutet, dass die Gehirnmasse größer ist als aufgrund der Körpermasse erwartet. Im Vergleich zu den Säugetieren besitzt der Mensch den größten EQ mit einem Wert zwischen 7 und 8 [55]. Selbst im Vergleich zu den anthropoiden Primaten hat der Mensch noch einen EQ von 3 [56]. Der hohe EQ als Erklärung für die höheren kognitiven Fähigkeiten der menschlichen Spezies birgt einige Schwierigkeiten und wurde in der Vergangenheit des Öfteren kritisiert. Dies führte zur Suche nach einem allgemein gültigen Algorithmus, der aus physiologischen und anatomischen Parametern vorhersagbare Intelligenzeigenschaften einer Spezies erlaubt [61-64]. Roth und Dicke (2005) entnahmen dem Vergleich verschiedener Parameter, wie Gehirngröße, relative Gehirngröße, Enzephalisierung, Nervenleitungsgeschwindigkeit und der geschätzten Anzahl an Neuronen, dass die Anzahl corticaler Neuronen und die Nervenleitungsgeschwindigkeit besser mit der Intelligenz bei verschiedenen Spezies korreliert [65]. In einer späteren Arbeit zeigten Roth und Dicke (2015), dass die intracorticale Informationsverarbeitung in Primaten, Affen und Menschen wesentlich schneller ist als bei Tieren mit einem größeren Gehirn, wie beispielsweise bei Elefanten oder Walen. Die Geschwindigkeit der Informationsverarbeitung ist wahrscheinlich schneller in sehr viel kleineren Gehirnen mit noch höherer neuronaler Packungsdichte, obwohl diese weniger Neurone besitzen. Deshalb wird davon ausgegangen, dass die Kombination aus einer hohen Anzahl an corticalen Neuronen und einer relativ hohen Informationsverarbeitungs-Kapazität verantwortlich für die Intelligenzeigenschaften unseres Gehirns ist [63].

Bisher konnten weder anatomische noch physiologische Eigenschaften identifiziert werden, die das menschliche Gehirn von anderen Säuger- oder Primatengehirnen in einer qualitativen Weise unterscheiden, mit wohl einer Ausnahme: dem Broca-Sprachzentrum. Die Evolution der syntaktisch-grammatikalischen Sprache war ein kompliziertes Ereignis, welches verbunden ist mit wesentlichen Modifikationen des Stimmapparats und der Evolution des Broca-Sprachzentrums, welches die Verbindung zwischen dorsalen präfrontalen Regionen und motorischer Kontrolle des Stimmapparats darstellt. Die Herstellung der Verbindung zwischen dem Wernicke-Sprachzentrum und dem Broca-Sprachzentrum war ebenfalls entscheidend für die Evolution der Sprache [66, 67]. Dies führt zu der Annahme, dass die

Entwicklung der syntaktisch-grammatikalischen Sprache eine wichtige Rolle bei der Zunahme der Intelligenz des modernen Menschen eingenommen hat [68].

2. Epigenetik

Conrad Waddington gilt als Schöpfer des Begriffs der „Epigenetik“. Er definierte diesen bereits 1942 als „die kausale Interaktion zwischen Genen und deren Produkten, welche eine phänotypische Expression ermöglicht“ [69]. Im Laufe der Zeit wurde diese Definition Waddingtons mehrfach umgewandelt und beschreibt heutzutage alle hereditären Veränderungen des Phänotyps, welche nicht durch Veränderungen in der primären DNA-Sequenz hervorgerufen werden [70, 71]. Dieses Teilgebiet der Genetik beinhaltet im Wesentlichen die DNA-Methylierung, Histon-Modifikationen und nicht-kodierende RNAs (Abbildung 6) [72].

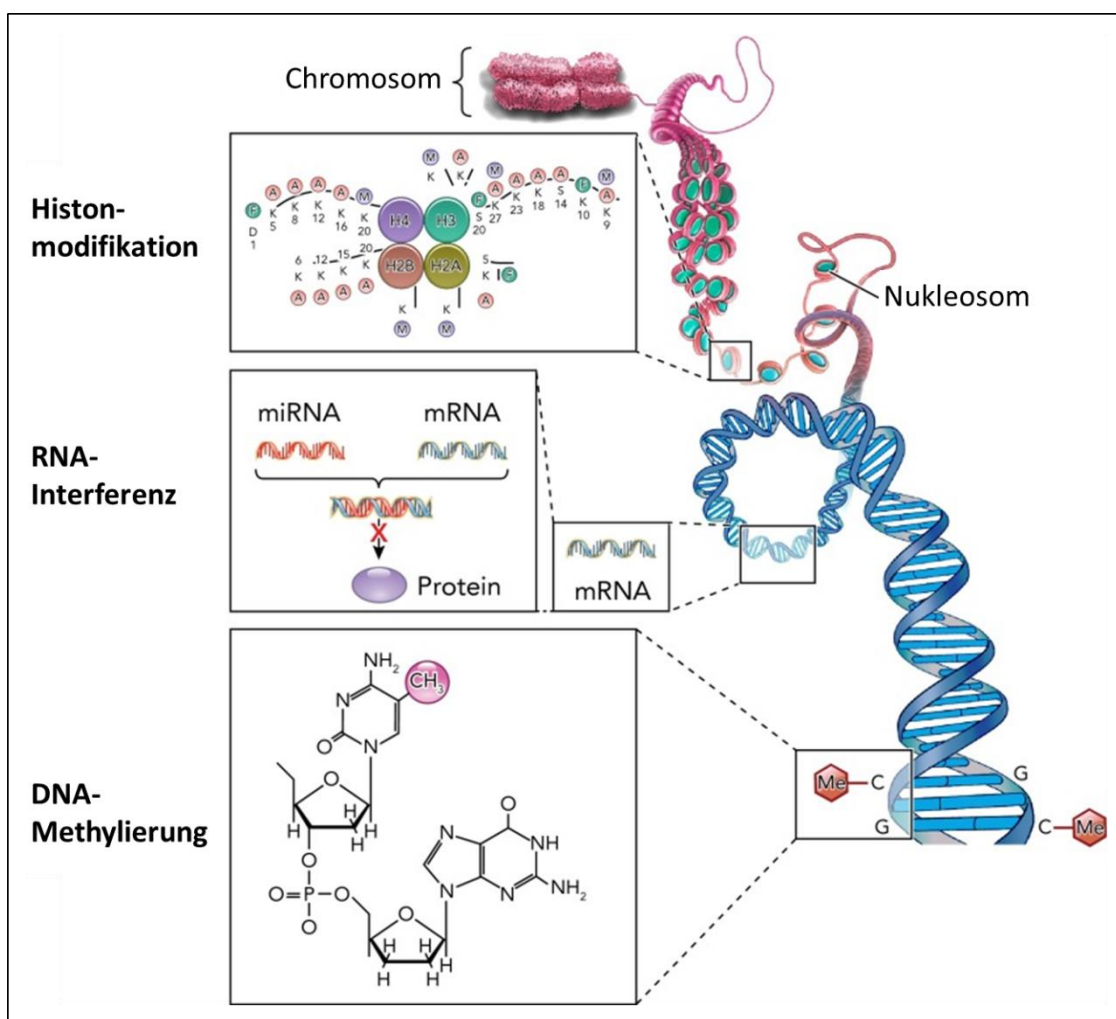


Abbildung 6: Schematische Darstellung der drei wesentlichen epigenetischen Mechanismen zur Kontrolle der Genexpression. (Abb. entnommen aus [73] und modifiziert)

2.1 DNA-Methylierung

Unter den epigenetischen Prozessen gilt die DNA-Methylierung als mit am besten erforscht [74]. Hierbei handelt es sich meist um eine biochemische Modifikation der Base Cytosin (C). Neben der kovalenten Verknüpfung einer Methylgruppe (CH₃-Gruppe) über S-Adenosylmethionin (SAM) an das 5'-Kohlenstoff-Atom (C-5) des Cytosins [75], konnten noch zwei weitere Formen der DNA-Methylierung beschrieben werden, N6-Methyladenin und N4-Methylcytosin, welche in Prokaryoten und einzelnen Eukaryoten vorkommen [76-78]. Die nicht-CpG Methylierung wurde erstmalig im Genom von Pflanzen entdeckt und beschreibt die Anheftung einer CH₃-Gruppe an Cytosine gefolgt von Adenin, Thymin oder einem weiteren Cytosin [79]. Es wird angenommen, dass die Methylierung an nicht-CpG-Stellen in humanen embryonalen Stammzellen und Gehirngewebe häufig vorkommt und 0,02% aller methylierten Cytosine in differenzierten somatischen Zellen ausmacht [80, 81]. 5-Methylcytosin (5-mC) stellt bei Eukaryoten die vorherrschende Form dar und tritt bei Säugern am Häufigsten als 5-mC innerhalb eines Cytosin-Phosphat-Guanin (CpG)-Dinukleotid-Kontexts auf [82]. Die Methylierung wird in Säugern hauptsächlich von drei DNA-Methyltransferasen (DNMTs) kontrolliert. DNA Methyltransferase 1 (DNMT1) sorgt für die Aufrechterhaltung der bestehenden Methylierungsmuster bei der DNA-Replikation und -Reparatur, indem dieses Enzym die Anheftung von CH₃-Gruppen an die hemimethylierten CpG-Stellen vermittelt [83]. Das Substrat S-Adenosylmethion (SAM) fungiert als Methylgruppendonor [84]. DNMT3A und 3B werden besonders während der Embryogenese exprimiert und dienen der Methylierung von CpG-Stellen, die bisher unmethyliert vorlagen [85]. Eine passive Demethylierung kann während der DNA Replikation in Abwesenheit von DNMT1 erfolgen. Die DNA-Methylierung eines spezifischen Locus kann dadurch nach zwei Zellteilungen komplett ausgelöscht sein. Die weitaus seltenere aktive Demethylierung wird über die Enzyme ten-eleven translocation methylcytosine dioxygenase 1-3 (TET1-3) katalysiert. TET-Enzyme besitzen die Fähigkeit, 5-mC zu 5-Hydroxymethylcytosin (5-hmC) und weiter zu 5-Formylcytosin und 5-Carboxylcytosin umzuwandeln (Abbildung 7) [86, 87].

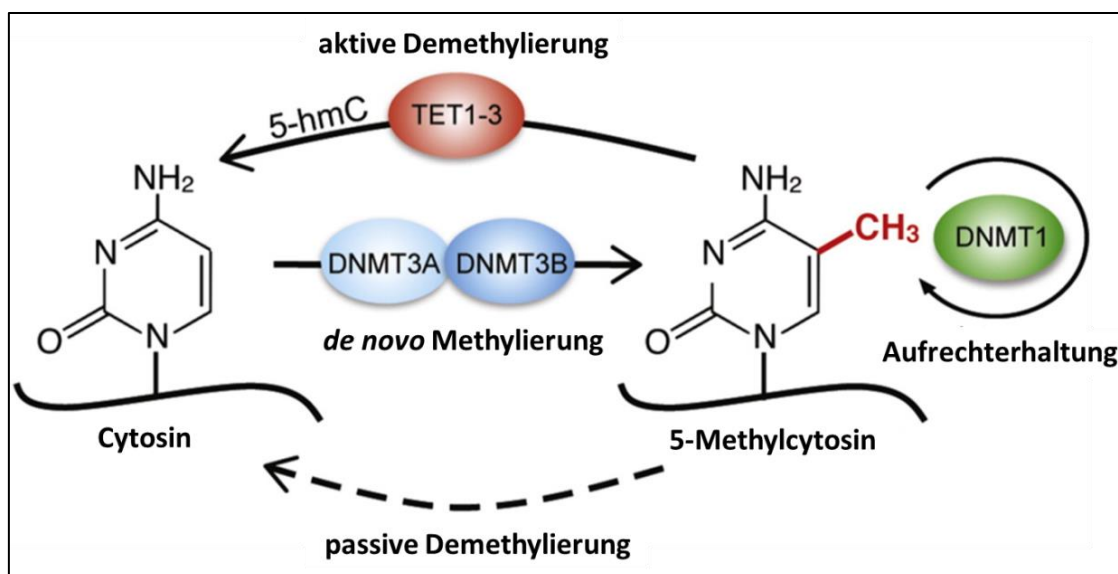


Abbildung 7: Vereinfachte schematische Darstellung der DNA-Methylierung und Demethylierung. In Eukaryoten erfolgt die Methylierung überwiegend am 5'-Kohlenstoffatom des Cytosins. Die *de novo* Anheftung der Methylgruppe (CH₃-Gruppe) wird über die DNA Methyltransferasen DNMT3A und DNMT3B katalysiert. Für die Aufrechterhaltung bereits etablierter Methylierungsmuster ist die Methyltransferase DNMT1 zuständig. Liegt während der Zellteilung kein aktives DNMT1 vor, kommt es zu einem Verlust der Methylierungsmuster am replizierten Strang. Hier spricht man von einer passiven Demethylierung. Es gibt jedoch auch eine aktive Form der Demethylierung welche über das Enzym tet methylcytosine dioxygenase 1-3 (TET1-3) katalysiert wird. Hierbei kommt es zur Oxidation von 5-mC zu 5-Hydroxymethylcytosin (5-hmC) und weiter zu 5-Formylcytosin und 5-Carboxylcytosin (nicht dargestellt) (Abb. entnommen aus [88] und modifiziert).

Das aktuell entschlüsselte menschliche Genom beinhaltet rund 3×10^7 CpG-Dinukleotidstellen, von denen in etwa 60% eine Methylierung am C-5 des Cytosinrings tragen [89]. Die DNA-Methylierung ist assoziiert mit einer langfristigen reprimierenden Wirkung auf die Transkription und wird auch bei der Mitose an die Tochterzellen weitergegeben [90, 91]. Diese kann zum einen direkt, über die Inhibierung der Bindung von Transkriptionsfaktoren, und zum anderen indirekt, über die Rekrutierung von Methyl-CpG-Bindeproteinen, welche einen Einfluss auf die Chromatinremodellierung haben, erfolgen [92]. Die Hypermethylierung von CpG-Stellen im Genkörper ist dagegen mit einer erhöhten Genexpression assoziiert [93]. Bei Säugern ist die Methylierung unter anderem beteiligt an einer gameten-spezifischen Reprimierung bestimmter Gene während der Entwicklungsphase aber auch ganzer Chromosomen, wie beispielsweise bei der X-Inaktivierung [94, 95]. Sie spielt ebenso eine essentielle Rolle für die allel-spezifische Genexpression beim genomischen Imprinting [96, 97]. Des Weiteren ist sie involviert in die Stilllegung repetitiver Elemente und trägt somit zur Aufrechterhaltung der genomischen Stabilität bei [98-100].

Transposable Elemente, wie beispielsweise short interspersed repetitive elements (SINEs), long interspersed nuclear elements (LINEs) und long terminal repeats (LTRs), weisen eine sehr hohe und dichte Methylierung auf. Das restliche Genom hingegen ist unterschiedlich stark methyliert. Promotor-assoziierte CpG-Inseln und das erste Exon liegen dahingegen meist unmethyliert vor. 75% aller Promotoren liegen in CpG-Inseln und sind unmethyliert [81, 89, 101, 102], die restlichen 25% weisen eine niedrige CpG-Dichte auf und werden wahrscheinlich nicht über die Methylierung reguliert (Abbildung 8) [89, 101].

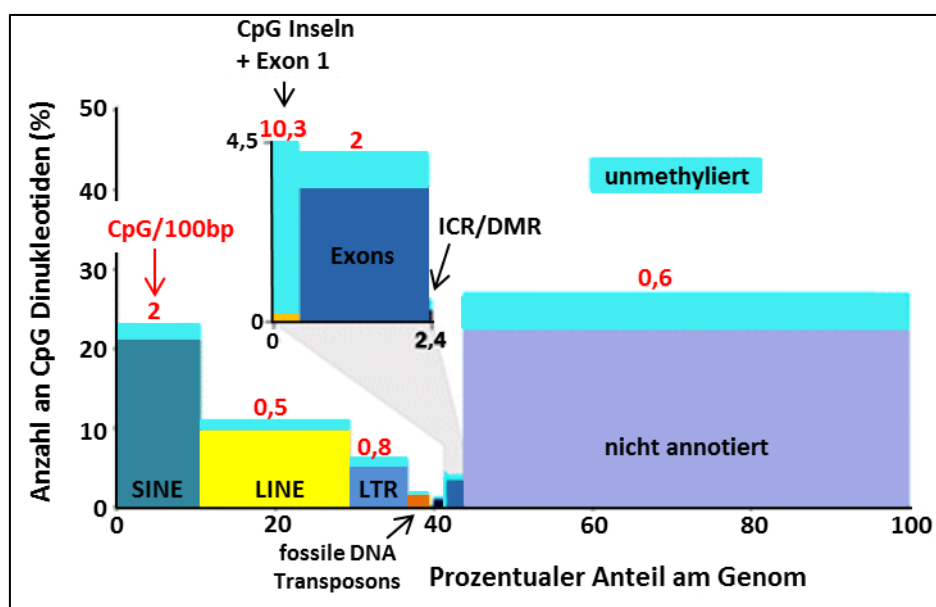


Abbildung 8: Verteilung der DNA-Methylierung im menschlichen Genom. Auf der Y-Achse ist die Anzahl der CpG-Dinukleotide und auf der X-Achse der Anteil im Genom dargestellt. Die hellblauen Balken zeigen den unmethylierten Bereich innerhalb jedes Kompartiments. Die roten Zahlen geben die Anzahl an CpG-Dinukleotide pro 100 bp an. CpG-Inseln und Exon 1 ist überwiegend unmethyliert und entspricht <0,5% des Genoms. Die ICR/DMR (differenziell methylierte Bereiche der Imprinting Kontrollregionen) nehmen ~0,001% des Genoms ein. Introns sowie putative Enhancer sind im nicht annotierten Bereich enthalten. Diese Abbildung bzw. die dazugehörigen Berechnungen wurden anhand von Bisulfidsequenzierungsdaten des Hippocampus erhoben. Andere adulte Gewebe zeigen ebenfalls ein ähnliches Verteilungsmuster (Abb. entnommen aus [103] und modifiziert).

2.1.1 Methylohm und Genom

Es kann mit Sicherheit gesagt werden, dass die DNA-Methylierung einen maßgeblichen Einfluss auf die Evolution der genomischen Sequenz genommen hat. In den frühen Achtzigerjahren des 20. Jahrhunderts (Jhds.) wurde herausgefunden, dass das Genom vieler Tierarten, meist von Wirbeltieren, arm an CpG-Dinukleotiden ist. Wenn man die Basenzusammensetzung des humanen Genoms genauer untersuchte, würde man eine

Häufigkeit von 0,04 für das Vorkommen einer CpG-Stelle erwarten, jedoch treten diese nur mit einer Häufigkeit von 0,008 auf [104, 105]. Ursächlich dafür ist eine spontane Desaminierungsreaktion, welche zu einer Umwandlung methylierter CpG-Stellen in TpG Dinukleotide führt [106]. Die Mutationsrate von CpG-Dinukleotiden ist zehnmal so hoch wie für andere Dinukleotide [107] und zeigt sich auch in Nukleotidabweichungen zwischen den Primaten. Genomweit gesehen, beläuft sich die durchschnittliche Abweichung zwischen Mensch und Schimpanse auf ~1% und steigt bis zu 15% in CpG-Stellen an [108]. In unmethylierten CpG-Inseln dahingegen bleibt der hohe CpG-Gehalt erhalten, da diese nicht anfällig für eine erhöhte Mutabilität sind. Deshalb wird angenommen, dass CpG-Inseln als eine Art Schalter in der Entwicklung dienen, über welchen die Zelle ein epigenetisches Gedächtnis für die Generierung von Zelltyp-spezifischen hyper- und hypomethylierten Mustern erlangt [109-111]. In stark methylierten, CpG-reichen Subtelomerregionen wird die hohe Rate der Desaminierung durch einen hohen Austausch mit Cytosin- und Guanin-Basen balanciert. Dies erfolgt meist über einen Genkonvertierungsmechanismus. In rund 15% solcher genomischer Regionen wird der durch die Desaminierung verursachte CpG-Verlust dennoch nicht kompensiert [112].

Evolutionäre Studien im Menschen konnten genetische Varianten identifizieren, die mit den Methylierungslevels in der Nähe gelegener CpG-Stellen assoziiert sind [113, 114]. Diese sogenannten „methylation quantitative trait loci“ (mQTL) stehen häufig mit Methylierungsveränderungen vieler CpG-Stellen innerhalb einer bis zu 3 kb großen Region in Verbindung. Des Weiteren können die mQTL mit Veränderungen auf Ebene der Genregulation, wie Histonmodifikation, DNaseI-Zugänglichkeit, Chromatinzugänglichkeit und der Expression benachbarter Gene assoziiert werden. Dies führt zu der Annahme, dass genetische Varianten zu gerichteten molekularen Veränderungen führen. Ein möglicher Faktor dieser koordinierten Veränderungen regulatorischer Mechanismen ist eine veränderte Bindung von Transkriptionsfaktoren (TF). Einzelnukleotid-Polymorphismen (engl. *single nucleotide polymorphism*, SNP), welche die TF-Bindeaffinität beeinflussen, konnten häufig mit der DNA-Methylierung nahe gelegener CpG-Stellen in einen Zusammenhang gebracht werden [113].

Weiterhin wurden Retrotranspositionereignisse mit graduellen Nukleotidveränderungen, die zur Akkumulation neuer CpG-Stellen führen, assoziiert [115]. Diese Regionen stellen die notwendige Basis für die Entstehung von neuen regulatorischen Bereichen dar. Einige

Studien kommen zu der Annahme, dass es sich bei der Prägung bestimmter Gene um ein Nebenprodukt bei der Stilllegung von Retrotransposons mittels DNA-Methylierung handeln könnte. Als Beispiel kann hier das retinoblastoma (*RB1*) Gen genannt werden, welches eine geprägte Region im Intron 2 beherbergt; diese Prägung resultiert aus der differenziellen Methylierung eines prozessierten Pseudogens [116].

2.1.2 DNA-Methylierung im Gehirn und Plastizität

Der DNA-Methylierung wird eine wichtige Rolle bei der epigenetischen Regulation während der Gehirnentwicklung und Neurogenese zugesprochen. Sie ist möglicherweise auch in die frühe Entwicklung des Gehirns und in der Spezifizierung von Regionen über die Genexpression involviert [117]. Der auffälligste Unterschied zwischen Menschen und nicht-menschlichen Primaten liegt in den hoch entwickelten kognitiven Fähigkeiten des Menschen. Die genetischen Ursachen, welche diese Unterschiede hervorrufen, sind nach wie vor noch recht unklar. Durch die große Sequenzähnlichkeit zwischen dem Genom des Menschen und dem der nicht-menschlichen Primaten rückte die Erforschung von Unterschieden auf Ebene der Genregulation im Gehirn immer stärker in den Vordergrund. Es wird angenommen, dass diese entscheidend für die menschliche Gehirnevolution waren [118-121]. Da epigenetische Modifikationen eine wichtige Rolle bei Lernprozessen und Gedächtnisbildung einnehmen, wird davon ausgegangen, dass diese auch zur menschlichen Gehirnevolution maßgeblich beigetragen haben [122-124].

Das menschliche Gehirn stellt ein ideales System dar bei dem gut beobachtet werden kann, welchen Einfluss die Veränderung externer Stimuli auf die Struktur und Funktion eines so komplexen Organs nehmen können. Über neuronale Schaltkreise läuft eine dynamische Filterung der enormen Menge an Signalen ab, denen der Mensch von Geburt an ausgesetzt ist und resultiert in der Komplexität unseres Verhaltens. Erfahrungsabhängige Plastizität beschreibt das Leistungsvermögen neuronaler Schaltkreise, sich selbst zu remodellieren und sich den Modifikationen zu unterziehen, welche auf Veränderungen der Aktivität und sensorischen Inputs basieren. Der Mechanismus der Plastizität ist wohl die aufregendste Fähigkeit des menschlichen Gehirns; die plastischen Phänomene können zum einen auf struktureller Ebene beobachtet werden, wie die Veränderungen der Schaltkreise und synaptischen Verbindungen, und zum anderen auf molekularer Ebene, wie Modifikationen in der Chromatinstruktur und schlussendlich in der Genexpression. Die Zellen des Gehirns

stehen in Abhängigkeit komplexer und stark regulierter Mechanismen, um über die Aktivierung oder Stilllegung bestimmter Gene in geeigneter Weise auf die Umgebung reagieren zu können. Dieses komplex regulierte System führt zu dem Schluss, dass epigenetische Mechanismen eine Schlüsselrolle bei der Aufrechterhaltung des transkriptionellen Programms innehaben, welches verantwortlich für die Etablierung und die Feinabstimmung der synaptischen Verbindungen und neuronalen Schaltkreise ist. Veränderungen in der Chromatinstruktur und Genexpression spielen eine Rolle bei der Vermittlung verschiedener Aspekte der erfahrungsabhängigen Plastizität, wie Lernen und Gedächtnisbildung, der fehlerhaft angepassten Plastizität (z.B. bei Drogenmissbrauch) und der Entwicklungsplastizität (z.B. visuelle corticale Plastizität) führen [125, 126].

Epigenetische Veränderungen können zu Störungen bei der Entwicklung des Nervensystems und zur Entstehung von neurodegenerativen Erkrankungen führen [127-130]. Mutationen in Genen, die für die Erkennung der DNA-Methylierung zuständig oder für eine normale neurale Entwicklung essentiell sind, führen zu epigenetischen Erkrankungen, wie dem Rett-Syndrom, Fragilen-X-Syndrom, Rubinstein-Taybi-Syndrom, Coffin-Lowry-Syndrom und Alpha-Thalassämie/Mentale Retardierung-Syndrom. Veränderungen der Histonmodifikationen oder DNA-Methylierung können ursächlich für neurodegenerative Erkrankungen sein, wie Morbus Parkinson, Alzheimer-Krankheit oder Chorea Huntington [131, 132]. Das Fragile-X-Syndrom (FXS) ist eine der häufigsten Ursachen erblich kognitiver Behinderung und wird zu den X-chromosomalen Mentalen Retardierungen gezählt. Diese Erkrankung ist mit dem *FMR1* (fragile X mental retardation-1) Gen assoziiert, welches stark in den Neuronen exprimiert wird und eine wichtige Rolle bei der Regulation der Dendriten- und Synapsenbildung einnimmt. Die wesentliche Ursache für das FXS ist die außergewöhnliche Expansion des CGG-Repeats in der 5'-UTR des *FMR1* Gens, welche zu Histon-Deacetylierung und DNA-Hypermethylierung führt und in einer transkriptionellen Stilllegung des *FMR1* Gens resultiert [133, 134].

2.1.2 Methylierung und Krebs

Im Laufe der Entwicklung eines Menschen können sich die epigenetischen Muster durch bestimmte Umwelteinflüsse oder während des Alterungsprozesses erheblich verändern; diese können unter anderem bestimmte Krebserkrankungen oder Krankheiten auslösen, die nicht nur, wie lange angenommen, rein durch Mutationen auf DNA-Sequenz Ebene

hervorgerufen werden können [135]. Wie wichtig die Etablierung und Aufrechterhaltung der Methylierungsmuster für die Entwicklung bei Säugern und die Funktionalität eines erwachsenen Organismus ist, zeigt die stetig wachsende Zahl an Erkrankungen, die mit einer veränderten Methylierung in Zusammenhang gebracht werden können [136].

Eine erste Verbindung zwischen der DNA-Methylierung und Krebs konnte 1983 von Feinberg und Vogelstein hergestellt werden. Es wurde gezeigt, dass das Genom einer Krebszelle in hypomethylierter Form im Vergleich zu einer nicht entarteten Zelle vorliegt [137, 138]. Diese Hypomethylierung in Tumorzellen entsteht primär durch die Demethylierung repetitiver Regionen und resultiert in genomischer Instabilität, was als Kennzeichen von Tumorzellen gilt (Abbildung 9, A) [139]. Die dadurch entstandene Reaktivierung der Promotoren von Transposons führt möglicherweise ebenfalls zu einer aberranten Genregulation über transkriptionelle Interferenz oder aber zu der Generierung von antisense Transkripten [131]. Der Verlust der genomweiten Methylierung ist eine häufige Beobachtung im frühen Stadium der Krebsentwicklung. Diese korreliert in den meisten Tumortypen mit der Schwere der Erkrankung und dem Metastasierungspotential der Krebszellen [140]. Dahingegen scheint eine gen-spezifische Demethylierung ein eher spätes Ereignis zu sein. Dies ermöglicht der Tumorzelle eine bessere Anpassung an ihre Umgebung und wirkt begünstigend auf die Metastasierung (Abbildung 9, B) [136].

Neben der genomweiten Hypomethylierung kann in Krebszellen auch eine gen-spezifische Hypermethylierung beobachtet werden. Diese tritt normalerweise innerhalb von CpG-Inseln auf, welche in somatischen Zellen meist unmethyliert vorliegen und führt zu Veränderungen der Chromatinstruktur, wie beispielsweise einer Hypoacetylierung der Histone, was schlussendlich zur Repression der Transkription führt (Abbildung 9, C) [92, 141, 142]. Meist sind davon Gene betroffen, welche an der Regulation des Zellzyklus, bei der Invasion von Tumorzellen, der DNA-Reparatur, der Chromatinremodellierung, zellulärer Signaltransduktion, Transkription und Apoptose maßgeblich beteiligt sind. Durch den Eingriff in diese wichtigen Mechanismen wird zum einen das Tumorzellwachstum gefördert, und zum anderen die genetische Instabilität erhöht, was zu weiteren genetischen Veränderungen in der Krebszelle führt und die Metastasierung zusätzlich unterstützt. In Tumoren, deren Fortschritt gut untersucht und verstanden ist, konnte gezeigt werden, dass eine aberrante Hypermethylierung bereits in den frühen Tumorstadiumen auftritt. Deshalb wird davon

ausgegangen, dass eine Hypermethylierung zur Transformation der Zellen beiträgt und kein spätes Ereignis darstellt, welches auf genetischen Veränderungen basiert [136, 143].

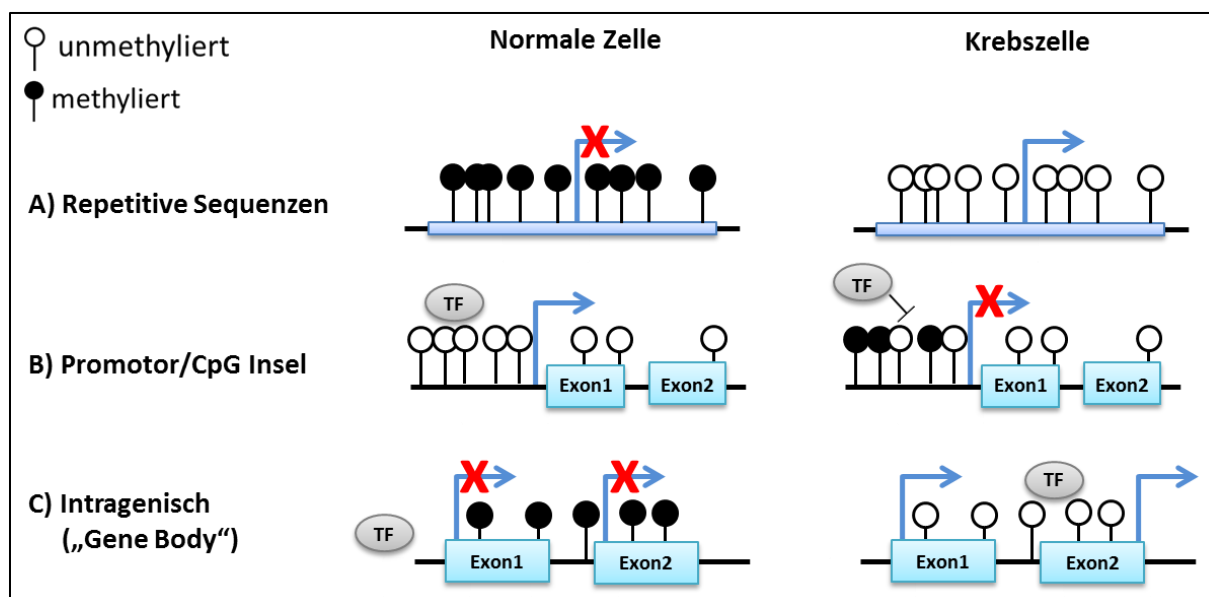


Abbildung 9: Schematische Darstellung der DNA-Methylierungsmuster in einer normalen Zelle und einer Krebszelle. A) Repetitive Sequenzen weisen generell eine Methylierung auf. Die globale Demethylierung in Krebszellen führt zu chromosomaler Instabilität und zur Aktivierung dieser Sequenzen. **B)** CpG-Inseln in Promotorbereichen liegen typischerweise unmethyliert vor. In Krebszellen dagegen sind diese oftmals hypermethyliert, was in einer transkriptioneller Repression resultiert. **C)** Intragenisch gelegene CpG-Dinukleotide sind häufig methyliert in nicht entarteten Zellen. Bei Krebszellen sind diese unmethyliert und führen zur Aktivierung der Transkription an falschen Positionen. (Abb. angefertigt nach Vorbild von [144])

3. Krebserkrankungen

Krebserkrankungen können sowohl durch Veränderungen der Basensequenz als auch der epigenetischen Muster hervorgerufen werden und beinhalten mehr als 300 Entitäten. Diese zeigen ähnliche genetische, epigenetische und auch pathologische Merkmale. Ab den frühen Siebzigerjahren des 20. Jhds. beschäftigten sich hauptsächlich drei Theorien mit der Entstehung von Krebs [145, 146]. Eine dieser Theorien postulierte, dass es sich bei Krebs um eine Krankheit mit einer abnormalen Zelldifferenzierung handelt. Als Ursache für die Entstehung maligner Zellen sind epigenetische Veränderungen verantwortlich, welche zu einer Störung des normalen Differenzierungsprozesses führen. Diese Theorie basiert hauptsächlich auf pathologischen Beobachtungen. Eine weitere Theorie wurde von Virologen formuliert und besagt, dass Krebs durch Viren hervorgerufen wird. Die dritte

Theorie verfolgte den Ansatz, dass es eine starke Korrelation zwischen mutagenen und tumorigenen Eigenschaften bei chemischen und physikalischen Agentien gibt, und somit die Bildung von Krebszellen das Resultat von Mutationen ist [147]. Im Laufe der Zeit konnten alle drei Theorien bestätigt werden. Zum einen half hierbei die Entdeckung des Onkogens *v-src* im Rous-Sarkom-Virus, sowie die Entdeckung des Protoonkogens *c-src* in humanen Zellen [148]. Zum anderen wurde die Mutationstheorie durch den Fund einer einzelnen Punktmutation im Onkogen *H-ras* bestätigt, welche zu einem Aminosäureaustausch von Glycin nach Valin führt und in einer Aktivierung des Onkogens resultiert [149].

3.1 Tumorgenese

Bei der Tumorgenese handelt es sich um einen mehrstufigen Prozess, dessen Verlauf in Abhängigkeit zu einer fortlaufenden Akkumulation von Mutationen innerhalb einer einzelnen Zelle steht [150]. Beim Menschen sind mindestens vier bis sechs Mutationen in einer neoplastischen Zelle erforderlich, um die homöostatischen Mechanismen abzuschalten, die bestimmend für eine normale Zellproliferation sind [151]. Diese genetischen Veränderungen führen zu einem erhöhten Zellwachstum, einer herabgesetzten Zellsterblichkeit und der Schaffung einer Zellumgebung, die das Wachstum unterstützt. Diese Komponenten resultieren schlussendlich in der Störung der gesamten Gewebshomöostase [152, 153].

Grundsätzlich lässt sich der Prozess der Tumorentstehung in vier Phasen unterteilen; die Initiationsphase stellt hierbei den Ausgangspunkt dar. Es kommt zu einem Mutationsereignis innerhalb einer Zelle - dafür ursächlich können zum einen krebsauslösende Substanzen, oxidativer Stress, Strahlung und Fehler während der Zellteilung sein. Es wird angenommen, dass sich jeden Tag tausende von Mutationsereignissen in einer Zelle abspielen. Jedoch werden diese Mutationen umgehend durch bestimmte Reparaturmechanismen behoben, oder aber direkt ein programmierter Zelltod eingeleitet. Wenn diese Mechanismen nicht greifen, kommt es zu einer dauerhaften Veränderung der DNA und in seltenen Fällen zu einer möglichen Entartung der betroffenen Zelle. Diese Zelle entwickelt sich dann zu einem sogenannten Mikrotumor, der aus bis zu zehn Millionen entarteter Zellen bestehen kann, bei einem Durchmesser von ca. 0,5 mm². Dieser Abschnitt wird als Promotion bezeichnet [154, 155]. In der Progressionsphase kommt es zur Ausbildung neuer Blutgefäße – auch als Angiogenese bezeichnet – um eine ausreichende Versorgung des Tumors mit Nährstoffen

und Sauerstoff zu gewährleisten. Jedoch setzt die Stimulation der Angiogenese die Entwicklung neuer Mutationen innerhalb des Mikrotumors voraus, der dem Tumor die Eigenschaft verleiht, die Bildung neuer Blutgefäße zu induzieren. Hierauf folgt die vaskuläre Phase, bei der es sich um das exponentielle Wachstum des Tumors handelt [154]. Durch die Gewährleistung der Blutversorgung ist es möglich, dass der Mikrotumor innerhalb von zwei Wochen seine Größe um das 16.000-fache steigert [155].

Die Vielfalt genetischer Veränderungen innerhalb eines Tumors führt zu dem Schluss, dass genetische Instabilität ein charakteristisches Merkmal vieler Krebserkrankungen ist und eine entscheidende Rolle bei der Krebsentstehung einnimmt [152, 153]. Bei 2-5% aller Erkrankungen kann eine familiäre Häufung beobachtet werden; hier ist von einer erblichen Prädisposition der Erkrankung auszugehen. Es konnte gezeigt werden, dass sowohl bei der Entwicklung von erblichen als auch von nicht erblichen Tumorerkrankungen die gleichen Gene beteiligt sind. Es handelt sich hier häufig um Tumorsuppressorgene oder Protoonkogene bzw. Onkogene. Beide Gruppen sind an der Regulation der Zellteilung beteiligt [156]. Die Tumorsuppressorgene gehören zu einer großen Molekül-Gruppe, zu deren Hauptaufgaben die Inhibierung bzw. Kontrolle der Zellteilung, Induktion der Apoptose, DNA-Reparatur und Inhibierung der Metastasierung zählen. Der funktionelle Verlust solcher Moleküle kann durch unkontrollierte Zellteilungen zur Entstehung von Krebszellen führen [157].

3.2 Brustkrebs

Brustkrebs ist die am häufigsten auftretende Krebserkrankung bei Frauen mit einer jährlichen Neuerkrankungsrate von ca. 1,5 Millionen Fällen weltweit. Laut einer Statistik der World Health Organization (WHO) starben 2015 rund 570.000 Frauen an Brustkrebs, dies entspricht etwa 15% der Todesfälle aller Krebserkrankungen [Int. 2]. Betrachtet man beispielsweise die Anzahl der Brustkrebsfälle in den USA von 1975-2012, ist ein deutlicher Anstieg sowohl der *in situ* als auch der invasiven Form dieser Erkrankung zu erkennen (Abbildung 10) [Int. 3]. Zu den *in situ* (nicht-invasiven) Formen werden das DCIS (Ductales Carcinoma *in situ*) und das LCIS (Lobular Carcinoma *in situ*), sowie weitere nicht-invasive Brustkrebsformen gezählt, welche duktale und lobuläre Eigenschaften aufweisen oder einen unbekanntem Ursprung haben. Beim DCIS handelt es sich um eine Brustkrebs-Frühform, bei

der die normalen Epithelzellen der Milchgänge durch abnormale Zellen ersetzt werden. Dies führt zu einer Vergrößerung der Milchgänge und der Milchdrüsenläppchen (Abbildung 11).

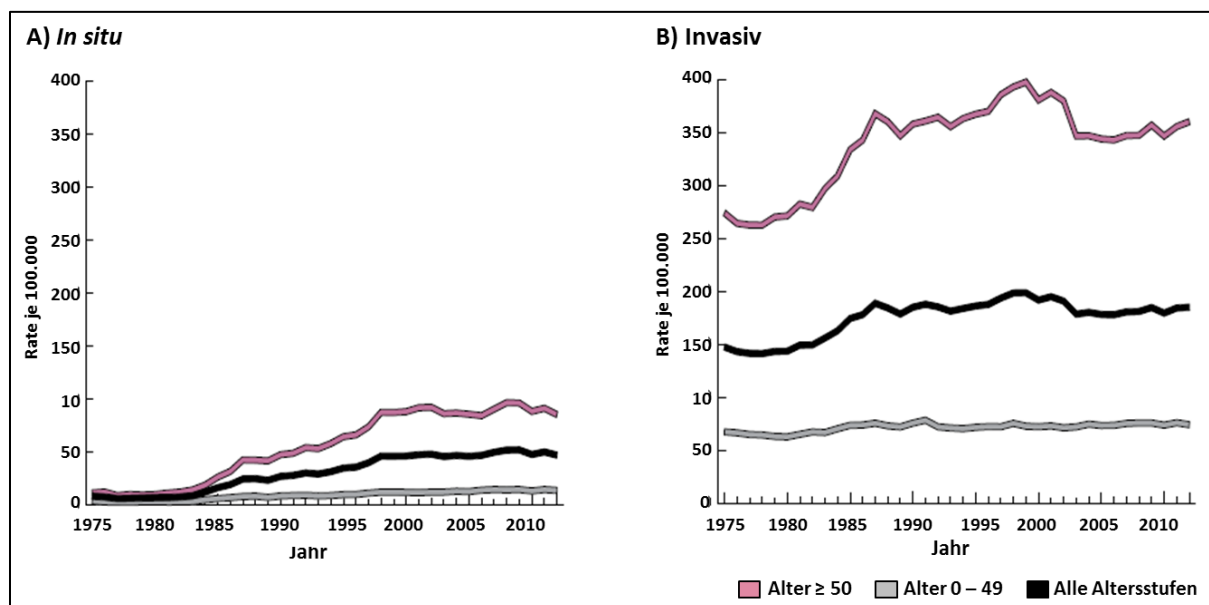


Abbildung 10: Anzahl an aufgetretenen *in situ* und invasiven Brustkrebskrankungen in den USA (1975-2012). Die Daten wurden durch das Surveillance, Epidemiology and End Results (SEER) Programm des US National Cancer Institutes erhoben (Abb. entnommen aus [Int. 3] und modifiziert).

Das DCIS ist der häufigste Typ der nicht-invasiven Fälle und machte rund 83% der *in-situ* Fälle zwischen 2008 und 2012 in den USA aus. Bei dieser Form besteht unter der Brustkrebs-Vorstufen das höchste Risiko, dass es zu einer Weiterentwicklung hin zu einer invasiven Form kommt. Deshalb werden die meisten Frauen mit DCIS sicherheitshalber behandelt, auch wenn dies für einige der Betroffenen nicht notwendig wäre. Das LCIS, bei dem die Zellen, welche eine krebszellen-ähnliche Optik besitzen, innerhalb der Milchdrüsenläppchen wachsen (Abbildung 11), gilt nicht als Vorläufer der invasiven Form; stattdessen wird das LCIS als Indikator für ein erhöhtes Risiko für die Entwicklung der invasiven Form erachtet. Das LCIS ist mit einer Häufigkeit von rund 13% (2008-2012, USA) wesentlich seltener als das DCIS. Am häufigsten handelt es sich bei Brustkrebskrankungen um die invasive oder infiltrierende Form. Hierbei durchbrechen die Krebszellen die Wände der Drüsen oder Gänge und wachsen in das umliegende Brustgewebe ein. Diese Karzinome werden untergliedert in duktal (Häufigkeit: 70-80%), lobulär (Häufigkeit: 10-15%) oder einige seltenere Varianten [Int. 3, Int. 4].

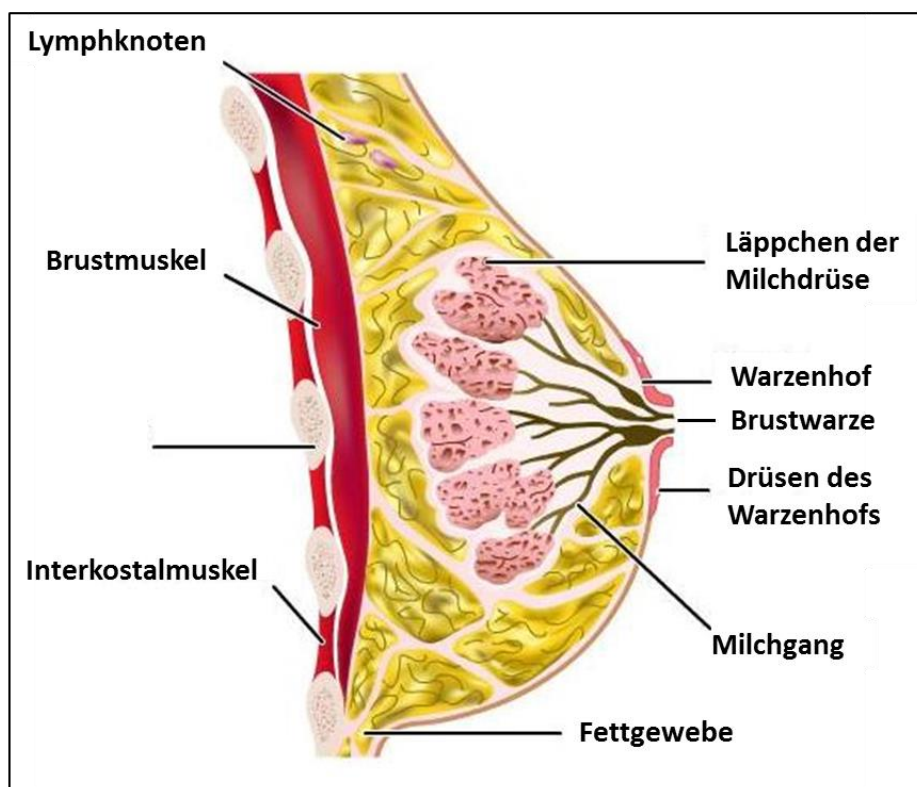


Abbildung 11: Aufbau der weiblichen Brust, schematisch dargestellt. (Abb. entnommen aus [Int. 5] und modifiziert)

Die Zunahme an Neuerkrankungen lässt sich durch das Ansteigen zahlreicher Risikofaktoren erklären, dazu zählen beispielsweise ein jüngeres Alter beim ersten Eintreten der Regelblutung, ein höheres Alter bei der Geburt des ersten Kindes, weniger Schwangerschaften, der Verzicht auf das Stillen oder eine Verkürzung der Stillzeit sowie ein späterer Beginn der Menopause. Zudem kommen weitere Risikofaktoren wie Übergewicht, der Konsum von Alkohol, Bewegungsmangel und Hormonersatztherapien hinzu [158].

3.2.1 Einteilung in Subtypen

Brustkrebserkrankungen können zwei Hauptgruppen zugeordnet werden: Estrogenrezeptor (ER)-positiv und ER-negativ. Diese beiden Gruppen werden weiterhin unterteilt in biologisch und klinisch signifikante Untergruppen. Bei ER-positiven Tumoren findet eine Expression des ER statt, sowie von ER-interagierenden Genen und Genen, die für charakteristische Proteine der luminalen Epithelzellen codieren. Diese Gruppe wird als luminal Gruppe bezeichnet und wird weiter in luminal A und B untergliedert. Diese Unterteilung steht in Abhängigkeit des

Expressionslevels charakteristischer Gene oder von Genen, die im Bezug zu dem Wachstumscluster und/oder HER-2 stehen [159-161]. Die ER-positiven Tumore repräsentieren die größte Patientengruppe [162]. Die Gruppe der ER-negativen Tumoren wird in drei Unterklassen eingeteilt: HER-2-positiv, „basal-like“ und sogenannte „normal breast-like“ Tumore. HER-2-positive Tumore verfügen über eine hohe Expression von Genen, die im epidermal growth factor receptor (HER-2) Amplikon auf Chromosom 17 lokalisiert sind und zeigen eine gleichzeitig erniedrigte Expression von ER und dem GATA binding protein 3 (GATA3) [163, 164]. Die „basal-like“ Tumore, auch als triple-negativ bezeichnet, weisen einen Verlust des ER, Progesteronrezeptors und HER-2, sowie die Expression von charakteristischen Genen basaler Epithelzellen auf. Diese Form tritt häufig bei Patienten mit einer Keimbahnmutation im *BRCA1* Gen auf [165]. Die Gruppe der „normal breast-like“ Tumore ähnelt gesundem Brustgewebe und zeigt ein hohes Expressionslevel vieler Gene, die typisch für Fettzellen und andere nicht-epitheliale Zelltypen sind [159, 162, 166]. Die richtige Zuordnung der Brusttumore in die verschiedenen Formen ist entscheidend für die Wahl des jeweiligen Therapieansatzes. So werden beispielsweise die luminal A Tumore mit einer hormonellen Therapie behandelt, wohingegen bei Tumoren, die eine Überexpression von HER-2 aufweisen, eine Behandlung mit monoklonalen Antikörpern angewandt wird, welche gegen HER-2 gerichtet sind, oder HER-2 Tyrosin Kinase Inhibitoren [167].

3.2.2 Molekulare Ursachen

Jahrzehntelange Forschung förderte das substantielle Verständnis für die Beteiligung verschiedener Faktoren an der Brustkrebsentwicklung; dabei handelt es sich oftmals um verschiedene genetische Veränderungen wie Genamplifikation, Gendeletion, Punktmutationen, Verlust der Heterozygotie (engl. *loss of heterozygosity*, LOH), chromosomale Umlagerungen und Aneuploidien, aber auch epigenetische Veränderungen, wie beispielsweise aberrante Methylierungsmuster [168]. Den größten Anteil an genetischen Veränderungen, die bei Brustkrebspatienten detektiert wurden, machen gain-of-function Mutationen in Proto-Onkogenen und loss-of-function Mutationen in Tumorsuppressorgenen aus [169].

3.2.2.1 Genetische Faktoren

Es konnte wiederholt beobachtet werden, dass eine Krebserkrankung gehäuft innerhalb einer Familie auftritt. Dies kann zum einen durch eine ähnliche Lebensführung und – umgebung und zum anderen durch die Genetik hervorgerufen werden. Die Berechnung des familiären relativen Risikos (FRR) hilft bei der Risikoeinschätzung für Angehörige von Patienten und wird über das Verhältnis von generellem Krankheitsrisiko in der Bevölkerung zum Krankheitsrisiko des Angehörigen eines Patienten ermittelt. Das FRR für Brustkrebs variiert je nach Alter bei der Krebsdiagnose des Indexpatienten und dem Alter des Verwandten [170, 171]. Keiner der für Brustkrebs bekannten Umweltrisikofaktoren hat einen Einfluss auf das FRR. Studien haben ergeben, dass bei 12% aller Frauen mit einer Brustkrebserkrankung ein weiterer Verwandter und bei 1% der Frauen zwei oder mehr Verwandte betroffen waren [170].

Mit Brustkrebs assoziierte genetische Varianten können eingeteilt werden in hoch penetrante Mutationen, welche selten in der Population auftreten, aber ein hohes Risiko darstellen, in moderat penetrante Varianten, welche mit einem mäßig erhöhten Risiko verbunden sind, und in niedrig penetrante Polymorphismen, welche ein häufiges Vorkommen in der Population aufweisen und mit einem leicht erhöhten Risiko assoziiert sind [172].

In einigen familiären Brustkrebsfällen handelt es sich um spezifische familiäre Brustkrebs syndrome, bei denen die Erkrankung auf einzelne Allele zurückzuführen ist, welche ein hohes Risiko hervorrufen [173]. In den Neunzigerjahren des 20. Jhds. wurde erstmals gezeigt, dass Mutationen in den Tumorsuppressorgenen *BRCA1* und *BRCA2* zu einem erhöhten Erkrankungsrisiko führen [174, 175]. Mutationen in diesen Genen prädisponieren ebenfalls für die Entwicklung von Eierstockkrebs [176]. Den Großteil der pathogenen Mutationen machen kleine Deletionen oder Insertionen aus, welche zur Translation eines trunkeierten Proteins führen [177]. Als weitere hoch penetrante Mutationen konnten Keimbahnmutationen in *TP53* identifiziert werden, für das Li-Fraumeni-Syndrom [178], in *PTEN* für das Cowden Syndrom [179] und Mutationen in *STK11/LKB1* für das Peutz-Jegher Syndrom [180]. Obwohl diese hoch penetranten Mutationen zu einem zehnfach erhöhten Erkrankungsrisiko beitragen, kommen sie in der Gesamtbevölkerung mit einer geringen Allelfrequenz (engl. *minor allele frequency*, MAF) von <0,005 vor und machen schätzungsweise nur rund 20-25% aller familiären Brustkrebsfälle aus [181, 182].

Einige der Sequenzveränderungen zählen zu den moderat penetranten Varianten. Diese treten sehr selten mit einer MAF von 0,005-0,01 auf und sind mit einem schwachen Risiko (zwei- bis vierfach erhöht) assoziiert [172, 182]. Sie bedingen weniger als 3% aller familiären Fälle [183]. Diese Gruppe beinhaltet protein-verkürzende Varianten in *CHEK2* [184] und Varianten in *PALB2* [183], *ATM* [185] und *BRIP1* [186]. Diese vier Gene sind maßgeblich an DNA-Reparaturmechanismen beteiligt.

Die sogenannten niedrig-penetranten Polymorphismen entsprechen einem polygenetischen Model, welches eine Kombination aus vielen individuellen Varianten darstellt. Diese Polymorphismen treten vergleichsweise häufig in der Bevölkerung mit einer MAF <0,05 auf und sind mit einem schwachen Risiko (<1,5-fach erhöht) assoziiert [182]. Sie machen den größten ungeklärten Teil des FRR aus [187, 188]. Mit der Hilfe von Assoziationsstudien, bei denen die genotypische Häufigkeit einer bestimmten Variante zwischen Betroffenen und nicht-verwandten Kontrollen miteinander verglichen wurde, konnten zwölf vielversprechende Varianten identifiziert werden (bis 2016) [189-195], wie beispielsweise *CASP8* D302H – eine missense Variante im kodierenden Bereich des caspase-8 (*CASP8*) Gens – und die missense Variante *TGFB1* L10P im transforming growth factor beta (*TGFB1*) Gen [194].

Trotz der vielen durchgeführten Studien zur Identifikation weiterer Suszeptibilitätsgene für Brustkrebs, gibt es für rund 70% aller Erkrankungen mit familiärem Hintergrund keine genetische Erklärung (Abbildung 10) [196, 197]. Dies führt zu dem Schluss, dass die restlichen genetischen Varianzen möglicherweise durch eine Kombination aus verschiedenen Varianten bedingt werden, die jeweils zu einem moderaten Brustkrebsrisiko beitragen [182].

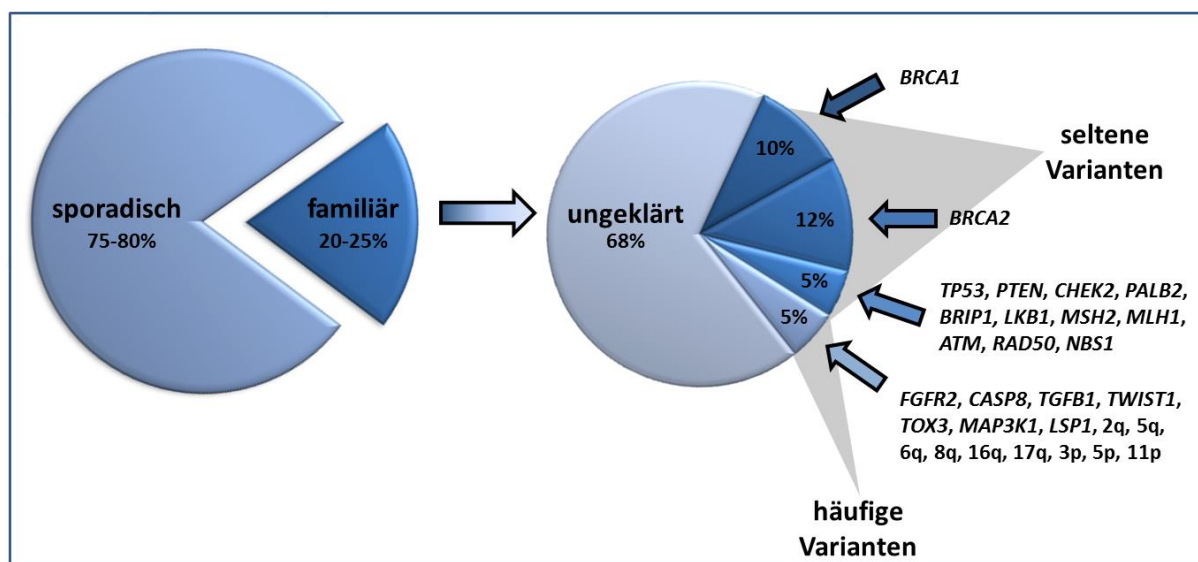


Abbildung 12: Verteilung der aktuell bekannten Brustkrebs-suszeptibilitätsgene bzw. -loci für familiären Brustkrebs. (Abb. basierend auf den Daten von [182, 198] angefertigt)

3.2.2.2 Methylierungsveränderungen und Brustkrebs

Epigenetische Veränderungen unterscheiden sich wesentlich von genetischen Veränderungen, da sie wesentlich häufiger auftreten, in definierten Genregionen erfolgen und reversibel sind, was durch die Behandlung mit pharmakologischen Agentien herbeigeführt werden kann [168]. In den letzten Jahren konnte eine Synergie zwischen genetischen und epigenetischen Veränderungen aufgezeigt werden. Deshalb wurde die „Two-Hit“ Hypothese von Knudson überarbeitet, welche besagt, dass eine Zelle erst dann kanzerogen wird, wenn beide Allele eines Gens in mutierter Form vorliegen (Abbildung 13) [199]. Tumorsuppressorgene können nicht, wie lange angenommen, nur durch LOH oder homozygote Deletionen, sondern auch durch transkriptionelle Stilllegung via DNA-Methylierung der Promotorregion inaktiviert werden [200, 201].

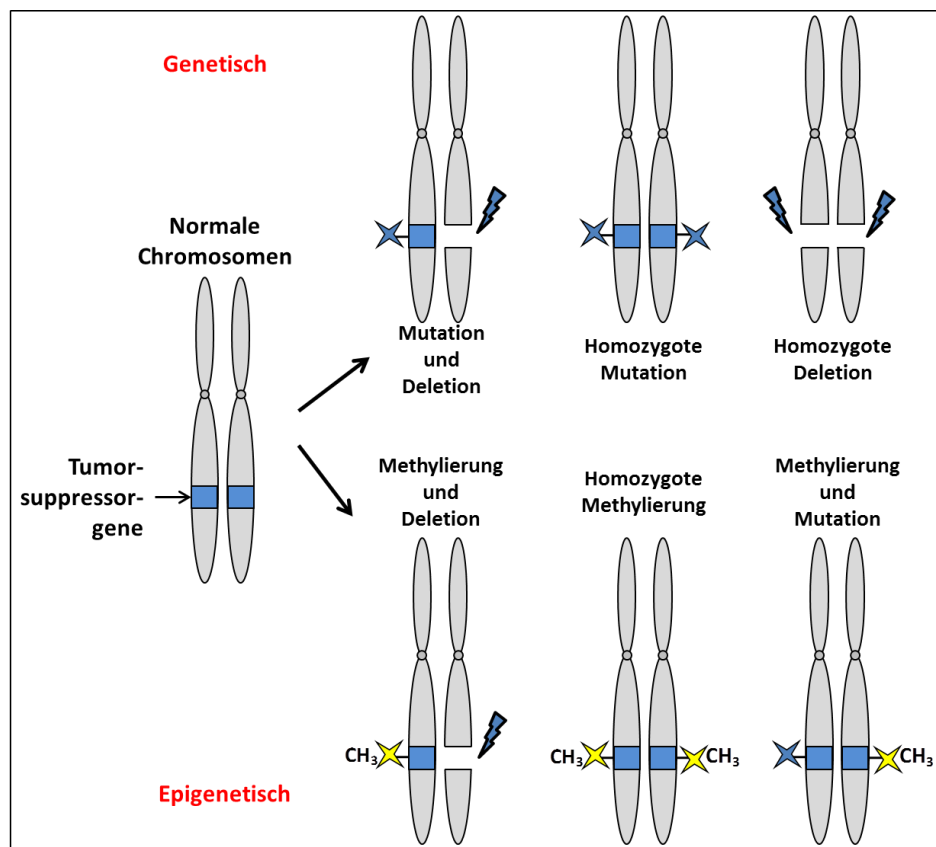


Abbildung 13: Die „Two-Hit“ Hypothese von Knudson. Bezogen auf die „Two-Hit“ Hypothese von Knudson sind genetische (oben) und epigenetische (unten) Veränderungen, die zur Inaktivierung von Tumorsuppressorgenen (TSG) führen können, dargestellt. Generell besagt diese Hypothese, dass eine Zelle erst dann kanzerogen wird, wenn beide Allele eines Tumorsuppressorgens in mutierter Form vorliegen. Der „First Hit“ kann eine Mutation in der codierenden Region eines TSG darstellen. Bei erblichem Krebs handelt es sich um Keimbahnmutationen und bei sporadischem Krebs um somatische Mutationen. Alternativ kann auch eine Hypermethylierung im Promoterbereich und die damit assoziierte Stilllegung der Transkription als „First Hit“ bei sporadischem Krebs, aber nicht bei erblichem Krebs fungieren. Bei sporadischen Krebsfällen stellen häufig Deletionen in der zweiten Genkopie den „Second Hit“ dar. In seltenen Fällen werden beide Genkopien über Punktmutationen inaktiviert. Eine Hypermethylierung beider Genkopien liegt dagegen häufiger vor. Bei erblichem Krebs mit einer Keimbahnmutation in einer Genkopie, handelt es sich bei dem „Second Hit“ entweder um eine Deletion oder eine Hypermethylierung im Promoterbereich. Die blauen Sterne stellen Genmutationen und die gelben Sterne Methylierungsveränderungen dar. Die Blitze deuten Deletionen an. (Abbildung erstellt nach dem Vorbild von [200] und [201])

Bei Brustkrebs sind zwei Arten von Methylierungsveränderungen bekannt; zum einen eine regionale Hypermethylierung spezifischer Gene und zum anderen eine globale Hypomethylierung [202, 203]. Diese paradoxe Koexistenz einer globalen Demethylierung mit einer lokusspezifischen Hypermethylierung weist auf zwei voneinander unabhängige und unterschiedliche Prozesse hin. Deshalb kann angenommen werden, dass die globale

Hypomethylierung aus einer globalen Steigerung der Aktivität von Demethylasen resultiert. Die spezifische Hypermethylierung könnte dahingegen durch lokale Veränderungen in der Chromatinstruktur, welche den Zugang der Demethylasen an die DNA verhindert, zustande kommen (Abbildung 14) [204, 205].

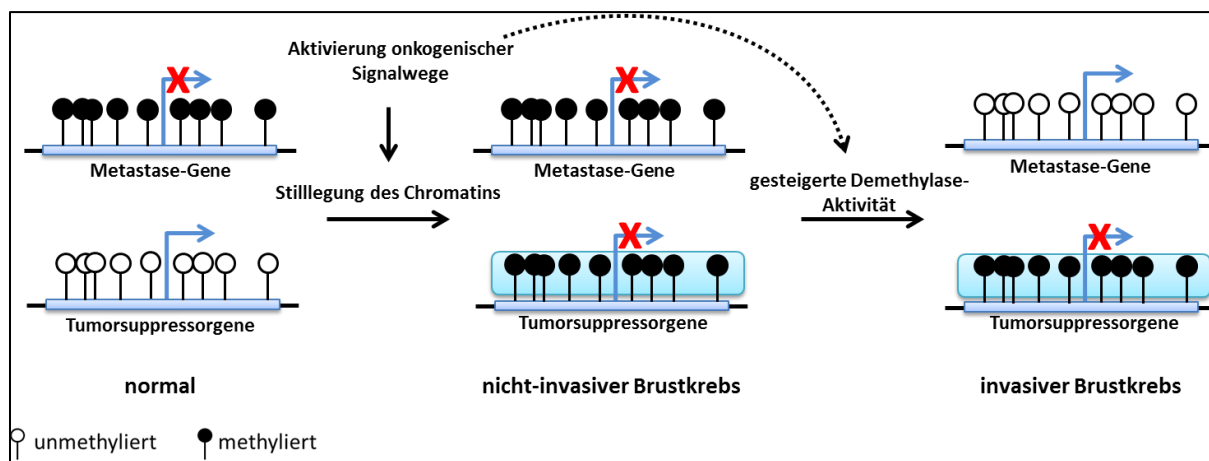


Abbildung 14: Methylierungsveränderungen bei Brustkrebs. In normalen Epithelzellen des Brustgewebes liegen die Tumorsuppressorgene in aktiver und unmethylierter Form vor. Metastase-Gene hingegen sind methyliert und somit inaktiviert. Die Aktivierung onkogener Signalwege führt zu einer lokalen Stilllegung des Chromatins und der DNA-Methylierung bestimmter Tumorsuppressorgene. In einem fortgeschrittenen Stadium führt derselbe onkogene Signalweg oder die Aktivierung weiterer onkogener Signalwege zu einer global gesteigerten Demethylase-Aktivität; dies resultiert in einer globalen Demethylierung und der Aktivierung von Genen, welche für die Metastasierung notwendig sind. Die Tumorsuppressorgene dagegen sind vor der Demethylierung durch lokale Repressoren geschützt, welche die inaktive Chromatinstruktur aufrechterhalten. (Abb. erstellt nach dem Vorbild von [206])

Die Forschung der letzten zehn Jahre offenbarte einige Gene, welche in hypermethylierter Form bei einer Krebserkrankung vorliegen, und daneben auch spezifische Methylierungssignaturen in Brustkrebszellen [207, 208]. Neben dem Wert für die Brustkrebsdiagnostik, zeigt sich auch anhand der Bandbreite an methylierten Genen, die mittels diverser Studien entdeckt worden sind, dass die Stilllegung dieser Gene eine wichtige Rolle in der Transformation einer gesunden Zelle hin zu einer Krebszelle spielt. Bei einem Teil dieser Gene handelt es sich um Tumorsuppressorgene, wie beispielsweise cyclin dependent kinase inhibitor 2A (*CDKN2A*; auch bekannt als *p16*), welches durch die Methylierung inaktiviert wird, was wiederum zur Abschaltung von Signalen führt, die zur Regulation des Zellwachstums nötig sind [209, 210]. Des Weiteren wurden Gene gefunden, die bei der

Schadensantwort beteiligt sind, wie *BRCA1* [211] oder Mismatch-Reparaturgene wie mutL homolog 1 (*MLH1*) und mutS homolog 2 (*MSH2*) [212]. Die Stilllegung von Reparaturgenen kann zu einer Anhäufung von sporadischen Mutationen führen [206]. Auch Gene der Steroidrezeptor Familie, wie estrogen receptor 1 (*ESR1*) [213] und retinoic acid receptor beta 2 (*RARβ2*), wurden in methylierter und dadurch inaktiver Form bei einigen Brustkrebsfällen detektiert [214, 215]. Es wird vermutet, dass die Interaktion zwischen *RARβ2* und Retinolsäure einen antiproliferativen Effekt hat und die Stilllegung einen selektiven Vorteil für die Brustkrebszellen bietet. Die Inaktivierung von Zelladhäsions- und Zelloberflächenmolekülen, wie cadherin 1 (*CDH1*) [216] und Proteaseinhibitoren, wie TIMP metalloproteinase inhibitor 3 (*TIMP3*) [217], kann eine unterstützende Wirkung auf die Metastase bei Brustkrebs haben.

Obwohl der Fokus der epigenetischen Brustkrebsstudien überwiegend auf der Hypermethylierung von Tumorsuppressorgenen liegt, ist die globale Hypomethylierung ebenfalls gut erforscht und dokumentiert. Überwiegend findet man eine Hypomethylierung bei repetitiven und Satelliten-DNA Sequenzen in Brusttumorgewebe [202, 218]. Aber auch Mitglieder der melanoma associated cancer/testis antigen (MAGE) Genfamilie, die normalerweise methyliert und inaktiviert vorliegen, wurden in vielen Tumoren in hypomethylierter Form entdeckt und werden in Brustkrebszellen exprimiert [219, 220]. Die Expression dieser Gene ist bei invasivem duktalem Brustkrebs mit einem wenig ausdifferenzierten Stadium assoziiert, und wird auch als anaplastisch bezeichnet [221]. Einzelne Gene, wie etwa urokinase plasminogen activator (*uPA*) [222] oder breast cancer specific gene 1 (*BCSG1*, auch bekannt als Synuclein gamma) [223, 224], zeigten ebenfalls eine Hypomethylierung bei Brustkrebspatienten. Diese Daten weisen darauf hin, dass in Krebszellen möglicherweise ein genereller Defekt der DNA-Methylierungsmaschinerie vorliegt. Es handelt sich wohl nicht um eine Reduktion der DNA Methyltransferase Aktivität in den Krebszellen, sondern, wie bei ER-Brustkrebszellen beobachtet, um eine erhöhte Expression der DNA Methyltransferasen während der G1 Phase des Zellzyklus [225, 226]. Darauf aufbauend kann angenommen werden, dass für die globale Demethylierung eine übermäßige Demethylierungsaktivität ursächlich ist (Abbildung 14) [227]. Abschließend kann gesagt werden, dass das Verständnis epigenetischer Veränderungen, wie die DNA-Methylierung und deren Beitrag zur Brusttumorgenese, neben den genetischen Mutationen

einen wichtigen Faktor für den weiteren Fortschritt im Bereich der Diagnostik, Prognose und Therapie darstellt [168].

4. Zielsetzung der Arbeit

4.1 DNA-Methylierungsanalysen corticaler Neuronen und Gliazellen im Menschen und Schimpansen

Wie konnte im Laufe der Evolution ein solch komplexes Organ wie das menschliche Gehirn entstehen? Welche Faktoren führten zur unterschiedlichen corticalen Expansion und Organisation bei den Säugern? Warum wurde das menschliche Gehirn vergleichsweise so groß und komplex und unterscheidet sich wesentlich vom Gehirn anderer Säugetiere? Basierend auf diesen und vielen ähnlichen Fragen wurden zahlreiche Theorien über die humane Gehirnevolution aufgestellt. Dank der sich stetig verbessernden wissenschaftlichen Erkenntnisse vergrößert sich das Wissen über unsere eigene Entstehungsgeschichte permanent. Dennoch bleiben viele Fragen weiterhin offen bzw. sind bisher nur unzureichend geklärt. Deshalb soll mit dieser Arbeit ein weiteres Puzzleteil für die Lösung dieser komplexen Fragestellung geschaffen werden. Da Gehirngewebe nicht fossilisiert, muss zur Erforschung der Gehirnevolution ein Umweg über den Vergleich mit unseren direkten Verwandten gemacht werden, nämlich dem Schimpansen. Aufgrund einer Ähnlichkeit von 98,8% zwischen der humanen DNA und jener des Schimpansen, wird angenommen, dass die unterschiedliche evolutive Entwicklung mehr auf der Regulation der Gene beruht. Transkriptionelle und epigenetische Modifikationen können zur Erhöhung der Variabilität der in der DNA-Sequenz gelegenen Informationen führen und so zu phänotypischer Vielfalt beitragen. Deshalb wurde ein genomweiter Next-Generation Sequencing (NGS)-basierter technischer Ansatz gewählt, um differenzielle Methylierungsmuster zwischen dem Menschen und dem Schimpansen identifizieren zu können. Der Sitz der höheren kognitiven Fähigkeiten des Menschen befindet sich im präfrontalen Cortex (Brodmann Areal 10, BA 10) und soll deshalb untersucht werden. Aufgrund der unterschiedlichen Anzahl und Funktion von neuronalen und nicht-neuronalen Zellen soll eine Zellseparation des Gehirngewebes in diese beiden Zellfraktionen unternommen werden, um einer Verfälschung der Ergebnisse entgegen zu wirken. Die Anzahl der neuronalen und nicht-neuronalen Zellen unterscheidet sich nicht nur innerhalb einer Art, sondern auch zwischen dem Gehirn des Menschen und des Schimpansen. Die aus den neuronalen und nicht-neuronalen Zellen isolierte DNA der

beiden Spezies soll dann mittels Reduced Representation Bisulfite Sequencing (RRBS) analysiert und einige der differenziell methylierten Bereiche via Bisulfit-Pyrosequenzierung validiert werden.

4.2 Methylierungsanalysen zur Detektion von Epimutationen in Brustkrebspatienten

Brustkrebs ist die am häufigsten auftretende Krebserkrankung bei Frauen (~25%) mit einer jährlichen Neuerkrankungsrate von ca. 1,5 Millionen Fällen weltweit [Int. 2]. Die Überlebensrate variiert sehr stark je nach Herkunftsland; so beläuft sich diese auf 80% in Nordamerika, Schweden und Japan, liegt bei 60% in Ländern mittleren Einkommens und fällt unter 40% in Ländern niedrigen Einkommens [228]. Bei rund 25% aller Frauen mit Brustkrebs kann eine familiäre Häufung beobachtet werden. Jedoch können nur bei 5-10% aller Patientinnen pathogene genetische Mutationen identifiziert werden [Int. 6]. Da neben den genetischen Ursachen auch die Auswirkungen epigenetischer Veränderungen auf die Karzinogenese bekannt sind, soll in dieser Arbeit das Methylierungsmuster ausgewählter Tumorsuppressorgene (TSG) bei *BRCA1/BRCA2*-mutationsnegativen Patientinnen analysiert werden. Mittels Deep Bisulfite Amplicon Sequencing (DBAS) sollen die Promoterbereiche der TSG (*BRCA1*, *BRCA2*, *RAD51C*, *ATM*, *PTEN*, *TP53*, *MLH1*, *RB1*) und des estrogenen Rezeptors (*ESR1*) untersucht werden, welches eine Rolle in der Tumorprogression spielt, und das als Kontrollgen fungiert. Es sollen Blutproben von zwei unabhängigen *BRCA1/BRCA2*-mutationsnegativen Brustkrebspatienten (BC)-Kohorten, sowie zwei unabhängige altersgematchte gesunde Kontrollkohorten untersucht werden. Die BC-Kohorte 1 beinhaltet early-onset (EO) BC-Patientinnen. Die BC-Kohorte 2 setzt sich aus BC-Patientinnen zusammen, die ein Risiko von >95% besitzen, um eine heterozygote Mutation in *BRCA1* und/oder *BRCA2* (HR) zu tragen. Inwieweit sich die Methylierungsmuster auf die Expression der TSG auswirken, soll mithilfe von anschließenden Expressionsstudien untersucht werden.

II MATERIAL

1. Probensammlung

1.1 Methylierungs- und Expressionsanalysen bei Brustkrebspatientinnen

1.1.1 — Vergleichende Methylierungsanalysen

In dieser Studie wurde insgesamt 79 aus EDTA-Blut extrahierte DNA-Proben von Brustkrebspatientinnen sowie 87 altersgemachte Kontrollproben untersucht. Über das Deutsche Konsortium für familiären Brust- und Eierstockkrebs konnte zum einen eine Kohorte (Kohorte 1) aus dem Uniklinikum Heidelberg mit 43 *early-onset* (EO) Patientinnen (Alter: 21–36 Jahre), zum anderen eine weitere Kohorte (Kohorte 2) aus dem Humangenetischen Institut der Universität Würzburg mit 36 Patientenproben (Alter: 27–85 Jahre) rekrutiert werden, die ein berechnetes Heterozygoten-Risiko von >95% (HR) aufweisen. Das Heterozygoten-Risiko wurde mittels des Berechnungsprogramms Cyrillic ermittelt. Im Rahmen einer genetischen Diagnostik konnten keine pathogenen Sequenz-Mutationen in *BRCA1* und *BRCA2* ermittelt werden. Bei den Kontrollproben handelt es sich um Personen, die bis zur Blutabnahme ohne Krebserkrankung waren und bei denen kein Auftreten von Brust- und Ovarialkrebs in deren familiären Umfeld bekannt war. Im Zuge der diagnostischen Beratung wurde das Einverständnis für eine Verwendung der DNA-Proben nach Beendigung der Diagnostik erteilt.

Tabelle 2: Übersicht über die analysierten Brustkrebspatienten- und Kontrollproben mit Angaben zum Alter bei der Blutabnahme (BA) und/oder Krebsdiagnose (KD). Proben, die bei der Blutabnahme nicht tumorfrei waren, sind mit einem Sternchen gekennzeichnet.

	Proben ID	Alter KD	Proben ID	Alter BA		Proben ID	Alter KD	Proben ID	Alter BA
Kohorte 1: Early-onset Brustkrebspatientinnen	EO BC 1	21	K 1	49	Kohorte 2: High-Risk Brustkrebspatientinnen (Heterozygoten-Risiko > 95%)	HR BC 1	28	K 51	27
	EO BC 2	27	K 2	29		HR BC 2	36	K 52	38
	EO BC 3	28	K 3	41		HR BC 3	38	K 53	38
	EO BC 4	30	K 4	47		HR BC 4	39	K 54	38
	EO BC 5	35	K 5	32		HR BC 5	39	K 55	39
	EO BC 6*	34	K 6	57		HR BC 6	40	K 56	40
	EO BC 7*	34	K 7	51		HR BC 7	40	K 57	40
	EO BC 8	31	K 8	36		HR BC 8	41	K 58	42
	EO BC 9	31	K 9	34		HR BC 9	42	K 59	43
	EO BC 10	31	K 10	30		HR BC 10*	42	K 60	44
	EO BC 11	34	K 11	58		HR BC 11*	43	K 61	45
	EO BC 12	31	K 12	48		HR BC 13*	47	K 62	46
	EO BC 13	32	K 13	58		HR BC 14*	47	K 63	47
	EO BC 14	35	K 14	33		HR BC 15	48	K 64	47
	EO BC 15*	36	K 15	51		HR BC 16	49	K 65	48
	EO BC 16	35	K 16	30		HR BC 17*	50	K 66	49
	EO BC 17	29	K 17	43		HR BC 18	50	K 67	50

EO BC 18	36	K 18	41	HR BC 19	51	K 68	51
EO BC 19	35	K 19	27	HR BC 37	85	K 69	51
EO BC 20	35	K 20	29	HR BC 20	51	K 70	52
EO BC 21	35	K 21	35	HR BC 21	51	K 71	52
EO BC 22	27	K 22	29	HR BC 22	52	K 72	53
EO BC 23	35	K 23	52	HR BC 23	53	K 73	54
EO BC 24	32	K 24	52	HR BC 24*	55	K 74	55
EO BC 25	35	K 25	41	HR BC 25	56	K 75	55
EO BC 26*	31	K 26	33	HR BC 26	57	K 76	57
EO BC 27*	32	K 27	27	HR BC 27	58	K 77	60
EO BC 28*	23	K 28	35	HR BC 28	58	K 78	58
EO BC 29	30	K 29	36	HR BC 29*	59	K 79	60
EO BC 30*	35	K 30	38	HR BC 30	62	K 80	62
EO BC 31	32	K 31	31	HR BC 31	63	K 81	62
EO BC 32	35	K 32	30	HR BC 32	68	K 82	70
EO BC 33*	29	K 33	37	HR BC 33	71	K 83	71
EO BC 34	28	K 34	34	HR BC 34	72	K 84	70
EO BC 35	31	K 35	51	HR BC 35	72	K 85	73
EO BC 36	27	K 36	31	HR BC 36	79	K 86	73
EO BC 37	32	K 37	55			K 87	75
EO BC 38*	32	K 38	47				
EO BC 39	26	K 39	29				
EO BC 40*	34	K 40	45				
EO BC 41*	33	K 41	33				
EO BC 42	31	K 42	30				
EO BC 43	33	K 43	31				
		K 44	27				
		K 45	33				
		K 46	27				
		K 47	61				
		K 48	41				
		K 49	35				
		K 50	47				

1.1.2 Expressionsanalysen

Zusätzlich zu den DNA-Analysen wurden in diesem Projekt Expressionsstudien angeschlossen. Da das Blut für eine DNA-Analyse im Rahmen einer genetischen Diagnostik untersucht wurde, musste für die RNA-Studie erneut Blut angefordert werden. Mit einer Ausnahme wurden deshalb Patientinnen, die ein aberrantes Methylierungsmuster gezeigt haben, für eine erneute Blutabnahme rekrutiert. Es konnte Blut von vier EO Patientinnen und drei HR Patientinnen gesammelt werden. Des Weiteren wurde Blut von drei gesunden Institutsmitarbeiterinnen als Kontrolle (K 90, K 91) und als Referenzwert (K 92) für die RNA-Extraktion und Expressionsanalyse verwendet.

Tabelle 3: Auflistung der Brustkrebspatienten- und Kontrollproben, die für die Expressionsanalyse verwendet wurden, versehen mit der Angabe zum Alter bei der Blutabnahme.

	Proben ID	Alter		Proben ID	Alter		Proben ID	Alter
Kohorte 1	EO BC 30	42	Kohorte 2	HR BC 8	44	gesunde Kontrollen	K 90	28
	EO BC 34	35		HR BC 11	47		K 91	24
	EO BC 36	34		HR BC 17	53		K 92	26
	EO BC 40	36						

1.2 Vergleichende Methylierungsanalysen zwischen Mensch und Schimpanse

1.2.1 Präfrontales Cortexgewebe

Zur Untersuchung der Methylierungsmuster im präfrontalen Cortex bei Menschen und Schimpansen stand Gehirngewebe von drei Individuen je Spezies zur Verfügung. Das Gewebe der Schimpansen stammte aus dem Biomedical Primate Research Centre (BPRC) in Rijswijk, das menschlichen Gehirngewebe aus dem Zentrum für Neuropathologie und Prionforschung (ZNP) in München. Bei keinem der untersuchten Individuen waren neurologische, neurodegenerative oder psychiatrische Erkrankungen bekannt. Zudem zeigte das Gewebe keine pathologisch oder sonstig auffällige strukturelle Veränderungen.

Tabelle 4: Übersicht über die Individuen für die vergleichende Methylierungsanalyse zwischen Mensch und Schimpanse. In der Tabelle finden sich Angaben zu Geschlecht, Alter, Herkunftsort des Gewebes und Todesursache.

Spezies	Proben ID	Geschlecht	Alter	Herkunft	Todesursache
Schimpanse	4	F	40	BPRC, Rijswijk, The Netherlands	ertrunken
	5	F	12	BPRC, Rijswijk, The Netherlands	Tod während Anästhesie
	6	F	43	BPRC, Rijswijk, The Netherlands	eingeschläfert
Mensch	1	F	77	ZNP, Munich, Germany	Herz-Kreislauf-Erkrankung
	2	F	77	ZNP, Munich, Germany	Herz-Kreislauf-Erkrankung
	3	F	67	ZNP, Munich, Germany	Herz-Kreislauf-Erkrankung

2. Primer und Adapter

Tabelle 5: Genspezifische PCR-Primer für das DBAS.

Gen	Primer	Sequenz (5' to 3')	CpG	Chromosomale Lage (bp) ^a	AL ^b	Typ	AT (°C) ^c
<i>hBRCA1</i>	F	GTATTTGAGAGGTTGTTGTTTAG	22	17:43125060-43125184	410 bp	Promoter	58
	R	ACTCAATACCCCTTCCTAATC		17:43125448-43125470			
<i>hBRCA2</i>	F	TTTGGGATAGTTGGGAGGGAA	52	13:32315354-32315375	597 bp	Promoter	56
	R	AATACCTAACAAAAATTCCTTT		13:32315926-32315950			
<i>hESR1</i>	F	GGGTTTGTTGTTGTTTAGT	31	6:151807611-151807531	449 bp	Promoter Flanking Region	56
	R	CCTCCAACACCTTTATAATACATA		6:151808035-151808059			
<i>hATM</i> Region 1	F	AGTTGGGAGGTAAAATTTAAAGTT	32	11:108222386-108222401	446 bp	Promoter	56
	R	AACTCTACCCACCTCTT		11:108222812-108222831			
<i>hATM</i> Region 2	F	AATGTTTTGGGGTAGTGT	32	11:108222920-108222938	385 bp	Promoter	56
	R	AATCAAAAACCACTCTAAAAAATAACAT		11:108223276-108223304			
<i>hTP53</i> Region 1	F	TTTTTTGGGAGTAGGTAGAAGAT	36	17:7688125-7688149	424 bp	Promoter	56
	R	AACAAATCTAAAACCTAAAAATAAAAT		17:7688520-7688548			
<i>hTP53</i> Region 2	F	GGGGTTTTTTTTTTTATGTGTTAAGAT	15	17:7687163-7687192	379 bp	Promoter	58
	R	ATAATCCTAACATCCCATCATCTA		17:7687516-7687541			
<i>hMLH1</i>	F	TTTGTTTTATTGTTGGATATTT	28	3:36993214-36993238	450 bp	Promoter	58
	R	CCAATTCTCAATCATCTCTTAATAACAT		3:36993635-36993664			
<i>hRB1</i>	F	TGAGGAGGATTTAGAGTAGGATAG	44	13:48303999-48304023	540 bp	Promoter	56
	R	ACCTCTCCACAAATAAACTCAAATA		13:48304511-48304538			
<i>hRAD51C</i>	F	GGTGGAGGTAAGGAAATGTG	32	17:58692335-58692355	429 bp	Promoter	58
	R	CCTCTAAAAATTCCTCAACAATCTAAA		17:58692737-58692764:1			
<i>hPTEN</i>	F	TGGGGAGTGGGAATTTGGAA	36	10:87862250-87862270:1	456 bp	Promoter	58
	R	TCAAAAACCCAAAAACACCTATCTA		10:87862679-87862705:1			

^a Die genomischen Koordinaten basieren auf Ensembl release 87.

^b AL = Amplikon-Länge.

^c AT = Annealing-Temperatur.

Tabelle 6: Indexprimer für das DBAS, enthalten in NEBNext Multiplex Oligos for Illumina (Dual Index Primer Set 1) Kit.

Indexprimer ^a	Sequenz (5' to 3') ^b	Index-Read Sequenz
NEBNext i501 Primer	5'-AATGATACGGCGACCACCGAGATC-TACACT <u>TATAGCCT</u> ACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	TATAGCCT
NEBNext i502 Primer	5'-AATGATACGGCGACCACCGAGATC-TACAC <u>ATAGAGG</u> CACACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	ATAGAGGC
NEBNext i503 Primer	5'-AATGATACGGCGACCACCGAGATC-TACAC <u>CCCTATCCT</u> ACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	CCTATCCT
NEBNext i504 Primer	5'-AATGATACGGCGACCACCGAGATC-TACAC <u>GGCTCTGA</u> ACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	GGCTCTGA
NEBNext i505 Primer	5'-AATGATACGGCGACCACCGAGATC-TACAC <u>AGGCGA</u> AGACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	AGGCGAAG
NEBNext i506 Primer	5'-AATGATACGGCGACCACCGAGATC-TACACT <u>AACTCTTA</u> ACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	TAATCTTA
NEBNext i507 Primer	5'-AATGATACGGCGACCACCGAGATC-TACAC <u>CAGGACG</u> TACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	CAGGACGT
NEBNext i508 Primer	5'-AATGATACGGCGACCACCGAGATC-TACAC <u>GTA</u> CTGACACTCTTTCCCTA-CACGACGCTCTCCGATC*T-3'	GTACTGAC
NEBNext i701 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>CGAGTAAT</u> GTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	ATTACTCG
NEBNext i702 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>TCTCCGG</u> AGTGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	TCCGGAGA
NEBNext i703 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>AATGAGCG</u> GTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	CGCTCATA
NEBNext i704 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>GGAATCT</u> CGTGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	GAGATTCC
NEBNext i705 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>TTCTGA</u> ATGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	ATTAGAA
NEBNext i706 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>ACGAATTC</u> GTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	GAATTCGT
NEBNext i707 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>AGCTTC</u> AGGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	CTGAAGCT
NEBNext i708 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>GCGCATT</u> AGTGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	TAATGCGC
NEBNext i709 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>CATAGCC</u> GGTGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	CGGCTATG
NEBNext i710 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>TTTCGCG</u> AGTGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	TCCGCGAA
NEBNext i711 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT-T <u>GCGCGAG</u> AGTGTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	TCTCGCGC
NEBNext i712 Primer	5'-CAAGCAGAAGACGGCATAACGAGA-T <u>TATCGCT</u> GTGACTGGAGTTCAGAC-GTGTGCTCTCCGATC*T-3'	AGCGATAG

^a NEBNext i500 Primer entsprechen den forward Primern; NEBNext i700 Primer den reverse Primern.

^b Die unterstrichenen Basen entsprechen dem Index Read.

Tabelle 7: Adaptersequenz für DBAS aus dem NEBNext Multiplex Oligos for Illumina (Dual Index Primer Set 1) Kit.

Adapter	Sequenz (5' to 3')
NEBNext Adaptor	5'-/5Phos/GAT CGG AAG AGC ACA CGT CTG AAC TCC AGT C/ideoxyU/A CAC TCT TTC CCT ACA CGA CGC TCT TCC GAT C*T-3'

Tabelle 8: Indexprimer für RRBS aus NEBNext Multiplex Oligos for Illumina (Methylated Adaptor, Index Primer Set 1) Kit.

Indexprimer ^a	Sequenz (5' to 3') ^b	Index-Read Sequenz
NEBNext Universal PCR Primer	5'-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC-s-T-3'	
NEBNext Index 1 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT- <u>CGTGAT</u> GTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	ATCACG
NEBNext Index 2 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT- <u>ACATCGG</u> TGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	CGATGT
NEBNext Index 3 Primer	5'-CAAGCAGAAGACGGCATAACGAGA- <u>TGCC</u> TAAAGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	TTAGGC
NEBNext Index 4 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT- <u>TGGTCA</u> GTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	TGACCA
NEBNext Index 5 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT- <u>CACTGT</u> GTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	ACAGTG
NEBNext Index 6 Primer	5'-CAAGCAGAAGACGGCATAACGAGA- <u>TATTGG</u> CGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	GCCAAT
NEBNext Index 7 Primer	5'-CAAGCAGAAGACGGCATAACGAGA- <u>TGATCT</u> GGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	CAGATC
NEBNext Index 8 Primer	5'-CAAGCAGAAGACGGCATAACGAGA- <u>TTCAAG</u> TGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	ACTTGA
NEBNext Index 9 Primer	5'-CAAGCAGAAGACGGCATAACGAGA- <u>TCTGAT</u> CGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	GATCAG
NEBNext Index 10 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT- <u>AAGCTA</u> GTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-	TAGCTT
NEBNext Index 11 Primer	5'-CAAGCAGAAGACGGCATAACGAGA- <u>TGTAGC</u> CGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	GGCTAC
NEBNext Index 12 Primer	5'-CAAGCAGAAGACGGCATAACGAGAT- <u>TACAAG</u> TGTGACTGGAGTTCAGACGT-GTGCTCTCCGATC-s-T-3'	CTTGTA

^a NEBNext Universal PCR Primer entspricht dem forward Primer; NEBNext Index Primer entsprechen den reverse Primern.

^b Die unterstrichenen Basen entsprechen dem Index Read.

Tabelle 9: Adaptersequenz für RRBS aus NEBNext Multiplex Oligos for Illumina (Methylated Adaptor, Index Primer Set 1) Kit.

Adapter	Sequenz (5' to 3') ^a
NEBNext Methylated Adaptor	5'-/PO ₄ /GAT /Me-dC/GG AAG AG/Me-dC/ A/Me-dC/A/Me-dC/ GT /Me-dC/TG AA/Me-dC /T/Me-dC//Me-dC /AGT/Me-dC//dU/A /Me-dC/A/Me-dC /T/Me-dC/T TT/Me-dC/ /Me-dC/ /Me-dC/T A/Me-dC/A /Me-dC/ GA /Me-dC/G/Me-dC/ T/Me-dC/TT/Me-dC//Me-dC/ GAT /Me-dC/-s-T-3'

^a PO₄ = Phosphat; Me = Methylgruppe.

Tabelle 10: Sonden für die TaqMan RT-PCR.

Gen	Assay ID	Assay Design	AL (bp) ^a	Dye	Transkripte
<i>hBRCA1</i>	HS01556193_m1	Probe spans exons	59	FAM-MGB	NM_007294.3; NM_00729.3; NM_007298.3; NM_007299.3; NM007300.3
<i>hBRCA2</i>	Hs00609073_m1	Probe spans exons	110	FAM-MGB	NM_000059.3
<i>hATM</i>	Hs00175892_m1	Probe spans exons	89	FAM-MGB	NM_000051.3
<i>hTP53</i>	Hs01034249_m1	Probe spans exons	108	FAM-MGB	NM_000546.5; NM_001126112.2; NM_001126113.2; NM_001126114.2; NM_001126115.1; NM_001126116.1; 001126117.1; NM_001126118.1; NM_001276695.1; NM_001276696.1; NM_001276697.1; NM_001276698.1; NM_001276699.1; NM_001276760.1; NM_001276761.1
<i>hMLH1</i>	Hs00979919_m1	Probe spans exons	104	FAM-MGB	NM_00249.3; NM_001167617.1; NM_001167618.1; NM_001167619.1; NM_001258271.1; NM_001258273.1; NM_001258274.1
<i>hRB1</i>	Hs01078066_m1	Probe spans exons	72	FAM-MGB	NM_000321.2
<i>hRAD51C</i>	Hs00427442_m1	Probe spans exons	142	FAM-MGB	NM_002876.3; NM_058216.2
<i>hPTEN</i>	Hs02621230_s1	Single exon	135	FAM-MGB	NM_000314.6; NM_001304717.2; NM_001304718.1
<i>hGAPDH</i>	Hs02786624_g1	Single exon	157	FAM-MGB	NM_001256799.2; NM_001289745.1; NM_001289746.1; NM_002046.5
<i>hACTB</i>	Hs01060665_g1	Probe spans exons	63	FAM-MGB	NM_001101.3

^a AL = Amplikon-Länge.

Tabelle 11: PCR-Primer und die korrespondierenden Sequenzierprimer für die Pyrosequenzierung.

Gen ^b	Primer	Sequenz (5' to 3')	CpG	Chromosomale Lage (bp) ^a	AL ^c	Typ	AT (°C) ^d
<i>h/pCAMTA1</i>	F	GTTTGGGGGTTTTTTAGGATT		1:7693591-7693780 (h)	189 bp	Open chromatin	59
	R*	CATAACAAAACTCCATCCTACTTAC		1:7636982-7637171 (p)		ENSR0000001165 (h)	
	S1 (h/p)	TGGTTTTGAGATATTGAGT	1				
	S2 (h/p)	AGGTGTTTTTTAGTAGGTT	4				
<i>hRBFOX3</i>	F	AGAGGAGATTGGGGTTTTT		17:77104329-77104572	243 bp	Open chromatin	57
	R*	CCTAACTCAACACTCCTTTAATTACC				ENSR00000132893	
	S1	GGGGTTTTTTAGAGGTTTATA	2				
	S2	GGAAGGTTGGTAGTAT	3				
<i>h/p PROSER2</i>	F (h)	GGGAGTTTAGTAGGATTTTGTTTAG		10:11911890-11912094 (h)	204 bp (h)	Open chromatin	58 (h)
	F (p)	GGTTTGGGAGTTTTGTAGGATTTTG		10: 12023450- 12023659 (p)	209 bp (p)	ENSR00000350501 (h)	59 (p)
	R* (h/p)	CCCCTACCCAAAACTCTCAAC					
	S2 (h/p)	GGGAGTAGGTGTTGGTTATTATTTA	8				
<i>pRTN4R</i>	F*	TGGGTTAGAATATGTTTTGGAGTTATGTA		22:18682523-18682772	249 bp	N/A	57
	R	ACCCATTACAATATTTCTAAACCCT			249 bp		
	S1	AAATAAAACCCCACT	2				
	S2	ATCCTAAAAATATACCTACAATC	2				

^a Die genomischen Koordinaten basieren auf Ensembl release 75 bzw. CHIMP2.1.4.

^b h = Human; p = Schimpanse (Pan troglodytes).

^c AL = Amplikon-Länge.

^d AT = Annealing-Temperatur.

* 5'-biotinyliert.

3. Kommerzielle Kit-Systeme

Tabelle 12: Kommerziell erworbene Kit-Systeme mit Herstellerangaben.

Kit-Systeme	Hersteller
Dneasy Blood and Tissue Kit	Qiagen, Hilden (GER)
EpiTect Fast Bisulfite Kit (50)	Qiagen, Hilden (GER)
EpiTect Fast 96 Bisulfite Kit	Qiagen, Hilden (GER)
High-Capacity RNA-to cDNA Kit	ThermoFisher Scientific, Waltham (USA)
High Sensitivity DNA Reagents	Agilent Technologies, Waldbronn (GER)
MinElute PCR Purification Kit (50)	Qiagen, Hilden (GER)
miRNeasy Mini Kit (50)	Qiagen, Hilden (GER)
MiSeq Reagent Kit v2 (500 cycles)	Illumina, San Diego (USA)
MiSeq Reagent Kit v3 (600 cycles)	Illumina, San Diego (USA)
NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1)	New England Biolabs, Ipswich (USA)
NEBNext Multiplex Oligos for Illumina (Methylated Adaptor, Index Primer Set 1)	New England Biolabs, Ipswich (USA)
Qubit dsDNA BR Assay Kit	ThermoFisher Scientific, Waltham (USA)
Qubit dsDNA HS Assay Kit	ThermoFisher Scientific, Waltham (USA)
TruSeq DNA LT Sample Prep Kit	Illumina, San Diego (USA)
TruSeqClusterGenerationKitv5	Illumina, San Diego, (USA)
TruSeqSBSKitv5	Illumina, San Diego (USA)

4. Reagenzien und Enzyme

Tabelle 13: Verwendete Reagenzien mit Angaben zum Hersteller.

Reagenzien	Hersteller
AMPure XP beads	Beckman Coulter, Krefeld (GER)
dATP Solution (100mM)	New England Biolabs, Ipswich (USA)
dNTP Mix (je 10 mM)	Roche Diagnostics, Mannheim (GER)
Gene Ruler Mix	Thermo Scientific, Waltham (USA)
HDGreen Plus DNA stain	INTAS Science Imaging, Göttingen (GER)
6 x Loading Dye	Thermo Scientific, Waltham (USA)
PhiX Control v3	Illumina, San Diego (USA)
PyroMark Gold Q96 CDT Reagents	Qiagen, Hilden (GER)
Streptavidin Sepharose High Performance	GE Healthcare, Uppsala (SWE)
TaqMan Fast Advanced Master Mix	Thermo Scientific, Waltham (USA)

Tabelle 14: Verwendete Enzyme mit Angaben zum Hersteller.

Enzym	Hersteller
FastStart Taq DNA Polymerase (5 U/ μ l)	Roche Diagnostics, Mannheim (GER)
Klenow fragment (3'→5' exo-) (5 U/ μ l)	New England Biolabs, Ipswich (USA)
MspI (20 U/ μ l)	New England Biolabs, Ipswich (USA)
Pfu Turbo C _x hotstart DNA Polymerase (2,5 U/ μ l)	Agilent, Waldbronn (GER)
T4 DNA Ligase (400 U/ μ l)	New England Biolabs, Ipswich (USA)
USER Enzyme (1 U/ μ l)	New England Biolabs, Ipswich (USA)

5. Chemikalien und Puffer/Lösungen

Tabelle 15: Chemikalien mit Angaben zum Hersteller.

Chemikalien	Hersteller
EDTA	Sigma-Aldrich, München (GER)
EDTA-Na ₂	Sigma-Aldrich, München (GER)
4 M Essigsäure	Roth, Karlsruhe (GER)
100 % EtOH	Roth, Karlsruhe (GER)
KHCO ₃	Sigma-Aldrich, München (GER)
5 M NaCl	AppliChem, Darmstadt (GER)
NaOH Plättchen	Roth, Karlsruhe (GER)
NH ₄ Cl	Sigma-Aldrich, München (GER)
PEG 8000	AppliChem, Darmstadt (GER)
Tris	AppliChem, Darmstadt (GER)

Tabelle 16: Verwendete Puffer und Lösungen mit Angaben zur Zusammensetzung.

Puffer/Lösungen	Zusammensetzung/Hersteller
50 x TAE	2M Tris ; 0,05M EDTA (pH 8) ; 100% Essigsäure
Denaturation Puffer	0,2 M NaOH
100 mM dNTP Solution (dATP, dCTP, dGTP, dTTP)	Invitrogen, Darmstadt (GER)
Elution Puffer	Qiagen, Hilden (GER)
10 x FastStart Taq PCR-Puffer (20 mM MgCl ₂)	Roche Diagnostics, Mannheim (GER)
HP3 Puffer (2 N NaOH)	Illumina, San Diego (USA)
10x Erythrozyten Lyse Puffer	155mM NH ₄ Cl; 10mM KHCO ₃ ; 0,1mM EDTA; pH 7,3-7,4
3 M NaAc (pH5,2)	Thermo Scientific, Waltham (USA)
10 x NEB Puffer 2	New England Biolabs, Ipswich (USA)

1 M PBS (pH 7,4)	Sigma-Aldrich, München (GER)
20 % PEG 8000/ 2,5 M NaCl	AppliChem, Darmstadt (GER)
PCR Grade Nucleotide Mix	Roche Diagnostics, Mannheim (GER)
Pfu Turbo C _x hotstart DNA Polymerase Puffer (10 x)	Agilent Technologies, Waldbronn (GER)
PyroMark Annealing Puffer	Qiagen, Hilden (GER)
PyroMark Binding Puffer	Qiagen, Hilden (GER)
T4-Ligase Puffer (10 x)	New England Biolabs, Ipswich (USA)
Wash Puffer (pH 7,6)	10 mM Tris-Acetat

6. Geräte

Tabelle 17: Verwendete Geräte mit Angaben zum Hersteller.

Geräte	Hersteller
AirClean 600 PCR Workstation UV-Bench	AirClean Systems, Raleigh (USA)
Agilent 2100 Bioanalyzer	Agilent Technologies, Waldbronn (GER)
Bandelin Sonorex RK 52 Ultraschallbad	Bandelin electronics, Berlin (GER)
Cluster Station	Illumina, San Diego (USA)
Delta Range PM 460	Mettler Toledo, Giessen (GER)
DNA Engine Tetrad 2 Thermal Cycler	Bio-Rad, München (GER)
EG4200-2NM Präzisionswaage	Kern&Sohn, Balingen (GER)
Elektrophoresis Power Supply	Bio-Rad, München (GER)
Eppendorf Masericycler nexus/pro	Eppendorf, Hamburg (GER)
Gel iX Imager	Intas Science Imager Instruments, Göttingen (GER)
GENE Mate Gelelektrophoresekamern	Starlab, Hamburg (GER)
Genome Analyzer II	Illumina, San Diego (USA)
Heraeus Multifuge 3SR+/ X1R/ X3	ThermoFisher Scientific, Waltham (USA)
Hettich Zentrifugen Mikro 20	Hettich Zentrifugen, Tuttlingen (GER)
PCR-Plattenzentrifuge Perfect Spin	VWR, Darmstadt (GER)
JANUS Automated Workstation	Perkin Elmer, Rodgau (GER)
Magnetic Particle Collector	ThermoFisher Scientific, Waltham, USA
MiSeq	Illumina, San Diego (USA)
NanoDrop 2000c Spectrophotometer	VWR, Darmstadt (GER)
pH-Meter inoLab	WTW, Weilheim (GER)
Pipetten Eppendorf Multichannel Plus	Eppendorf, Hamburg (GER)
Pipetten Eppendorf Research Plus (0,1 µl-1000 µl)	Eppendorf, Hamburg (GER)
POWER PAC 1000	Bio-Rad, München (GER)
PyroMark Q96 Capillary/Reagent Tips/Capillary Tip	Qiagen, Hilden (GER)

Holder	
PyroMark Q96 MD Pyrosequencer	Qiagen, Hilden (GER)
PyroMark Q96 MD Vacuum Workstation System	Qiagen, Hilden (GER)
Qubit Fluorometer	ThermoFisher Scientific, Waltham (USA)
SpeedVac Concentrator Plus	Eppendorf, Hamburg (GER)
TB2 Thermoblock	Analytik Jena, Jena (GER)
Tischzentrifuge Eppendorf Centrifuge 5424	Eppendorf, Hamburg (GER)
Tischzentrifuge Galaxy 16 DH	VWR, Darmstadt (GER)
TS-100 Thermoblock	VWR, Darmstadt (GER)
ViiA 7 Real-Time PCR System	ThermoFisher Scientific, Waltham (USA)
Vortex-Genie 2	Scientific Industries, New York (USA)

7. Software und Analysetools

Tabelle 18: Auflistung der wichtigsten Software- und Analysetools.

Software	Anwendung	Hersteller/Verfügbarkeit
BiSearch	Primer Design and Search Tool	http://bisearch.enzim.hu/?m=genompsearch
Bismark v0.14.3	Read alignment	http://www.bioinformatics.babraham.ac.uk/projects/bismark/
Bowtie2 v2.2.6	Read alignment	bowtie-bio.sourceforge.net/bowtie2/
Clustal Omega v1.2.2	Multiple sequence alignment	http://www.ebi.ac.uk/Tools/msa/clustalo/
Cutadapt v1.6	Adapter Trimming und Entfernen von ungewollten Sequenzen aller Art	https://cutadapt.readthedocs.io/en/stable/
Cyrillic v3	Berechnung des Heterozygotenrisiko	http://www.apbenison.com/about-cyrillic/
Ensembl genome browser	Genom-Datenbank für Vertebraten	http://www.ensembl.org/index.html
FastQC v0.11.2	Visualisierung von Sequenzierungsdaten zur Qualitätseinstufung	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
fastq-join ea-utils v1.1.2-537	Read joining	https://expressionanalysis.github.io/ea-utils/
GenSearchNGS v4.4	Visualisierung und Analyse der alignierten NGS Daten	http://www.phenosystems.com/www/index.php/products/gensearchngs
Illumina Experiment Manager	Generierung Sample Sheets für MiSeq Sequenzierung	Illumina, San Diego (USA)
LASAGNA	Search tool für Transkriptionsfaktor-Bindestellen	http://biogrid-lasagna.engr.uconn.edu/lasagna_search/

Microsoft-Office 2010	Datenauswertung, Textverarbeitung und Erstellen von Abbildungen	Microsoft Corporation, Redmond (USA)
PhenoGram	Erstellung der Phenogramme	http://visualization.ritchielab.psu.edu/phenograms/plot
PyroMark Q-CpG v1.0.9	Pyrosequenzierung	Qiagen, Hilden (GER)
PyroMark Assay Design v2.0	Primerdesign für bisulfit-konvertierte DNA	Qiagen, Hilden (GER)
QuantStudio Real-Time PCR Software	Analyse Expressionsdaten	ThermoFisher Scientific, Waltham (USA)
Reverse Complement	Konvertierung von Sequenzen	https://www.bioinformatics.org/sms/rev_comp.html
SPSS Version v23	Statistische und grafische Auswertung von DBS- und Pyrosequenzierungsdaten	IBM, Berlin (GER)
TrimGalore, v0.4.0	Adapter Trimming von FastQC Files	http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
UCSC Genome Browser	Genom-Datenbank	https://genome.ucsc.edu/

III METHODEN

1. Isolation genomischer DNA und RNA-Extraktion

1.1 DNA-Extraktion

1.1.1 Blutproben

Die bereitgestellte genomische DNA (gDNA) der Brustkrebspatienten aus Kohorte 2 und der Kontrollproben aus Kohorte 1 und 2 wurde mittels einer Standard Salting-out Methode extrahiert. Die Patientenproben aus Kohorte 1 wurden unter Anwendung des Puregene Blood Kits, entsprechend den Angaben des Herstellers, aus Blut isoliert. Die Konzentration und Qualität der gDNA wurde anschließend über ein NanoDrop 2000c Spektrophotometer (VWR, Darmstadt, Deutschland) bestimmt. Eine detaillierte Auflistung der untersuchten Patientenproben mit Angaben zu Herkunft, Alter und Kohorte finden sich im Materialteil in Tabelle 2.

1.1.2 Neuronale und nicht-neuronale Zellen

Bevor die gDNA aus dem präfrontalen Cortex isoliert wurde, wurde das Gewebe in neuronale und nicht-neuronale Zellen mittels Durchflusszytometrie separiert. Hierfür wurde ein sogenanntes FACS (fluorescence-activated cell sorting)-Gerät verwendet. Die Zellseparation und anschließende DNA-Isolation wurde von den Kollaborationspartnern aus München nach einem Protokoll von Matevoassian und Akbarian durchgeführt [229]. Eine detaillierte Auflistung der Gewebeproben mit Angaben zu Spezies, Geschlecht, Alter, Herkunftsort und Todesursache findet sich im Materialteil in Tabelle 4.

1.1.3 Präfrontales Cortexgewebe

Für die Validierung der RRBS-Daten und die damit verbundene Assay-Etablierung wurde DNA aus Gesamtgewebe des präfrontalen Cortex isoliert, welches von Menschen bzw. Schimpansen stammte (Tabelle 4). Hierfür wurde das DNeasy Blood & Tissue Kit (50) von Qiagen (Hilden, Deutschland) unter Zuhilfenahme des beiliegenden Handbuchs verwendet. Zu Beginn wurden 25-35 mg Gehirngewebe über Nacht (ÜN) lysiert; nachdem das Lysat auf die Säulchen aufgebracht worden ist, folgten verschiedene Waschschriffe. Die DNA wurde in zweimal 200 µl EB Puffer eluiert.

1.2 RNA-Extraktion

Vor dem Start der eigentlichen RNA-Extraktion wurde eine Lyse der Erythrozyten durchgeführt, um die Ausbeute der RNA zu erhöhen. Es wurden 2,5 ml Vollblut mit 45 ml 1x Lyse Puffer (Tabelle 16) in einem 50 ml Falcon vermennt und für 15 Minuten (min) bei Raumtemperatur (RT) auf einen Schüttler gestellt, sodass das Blut-Puffer-Gemisch durchgängig durchgemischt wurde. Anschließend erfolgte ein Zentrifugationsschritt für 5 min bei 4 °C und 500 x g. Hierbei setzte sich am Falconboden ein Pellet ab und der Überstand wurde vorsichtig abgekippt. Das Pellet wurde nun in 10 ml PBS gewaschen und noch einmal für 5 min bei 4 °C mit 500 x g zentrifugiert. Nach dem Abgießen des Überstands wurde das Pellet in 1 ml PBS gelöst und in ein 2 ml Gefäß überführt. Nach einem weiteren Zentrifugationsschritt (5 min; 4 °C; 500 x g) wurde der Überstand abgegossen. Die nachfolgenden Schritte wurden mit dem miRNeasy Mini Kit (Qiagen, Hilden, Deutschland), gemäß der Vorgaben des Herstellers (im beiliegenden Handbuch) durchgeführt. Die RNA wurde abschließend in 35 µl RNase-freiem Wasser eluiert. Die Isolation der RNA wurde anhand der in Tabelle 3 aufgelisteten Proben durchgeführt.

2. Natriumbisulfitbehandlung genomischer DNA

Es gibt bisher noch kein standardisiertes Verfahren, um die Methylierungs-Information über eine einfache Sequenzierung oder eine hybridisierungs-basierte Methode direkt aus der gDNA herauszulesen, da die Methylierung bei einer einfachen PCR oder dem Klonieren verloren geht. Es wurden bisher verschiedene Techniken entwickelt und angewandt, um diese Limitierung umgehen zu können; hierzu zählen beispielsweise ein Verdau der gDNA über ein methylierungs-sensitives Restriktionsenzym (HpaII, McrBC) oder eine Affinitätsaufreinigung mittels Antikörper. Die gängigste und einfachste Methode ist jedoch die Bisulfitkonvertierung, die unter anderem im Next-Generation-Sequencing (NGS)-Bereich großen Einsatz findet [230].

Die Verwendung von Natriumbisulfit, um unmethylierte Cytosine in Uracil zu konvertieren, wurde bereits 1970 von zwei unabhängigen Arbeitsgruppen beschrieben [231, 232]. Das grundlegende Prinzip basiert auf einer Desaminierungsreaktion am Cytosinring, bei der ein Ammonium-Ion (NH_4^+) abgespalten wird und die Umwandlung des Cytosins zu Uracil erfolgt. Methylierte Cytosine sind hingegen durch das Tragen einer Methyl-Gruppe (CH_3 -) geschützt und somit von dieser Reaktion ausgenommen. Im Detail kann dieser Vorgang untergliedert

werden in die Sulfonierung (Abbildung 15, 1), die hydrolytische Desaminierung (Abbildung 15, 2) und die alkalische Desulfonierung (Abbildung 15, 3). Bei der Sulfonierung erfolgt eine nukleophile Addition des Hydrogensulfits (HSO_3^-) an das C-6 Atom des Cytosins, wobei sich ein instabiles Cytosinsulfonat-Derivat bildet. Im nächsten Schritt kommt es aufgrund der wässrigen Umgebung zu einer hydrolytischen Desaminierung am C-4 Atom; hierbei wird die NH_2 -Gruppe in Form von NH_4^+ abgespalten. Im dritten Schritt, der alkalischen Desulfonierung, wird HSO_3^- freigesetzt; es bildet sich erneut eine 5,6-Doppelbindung aus und Uracil entsteht [233].

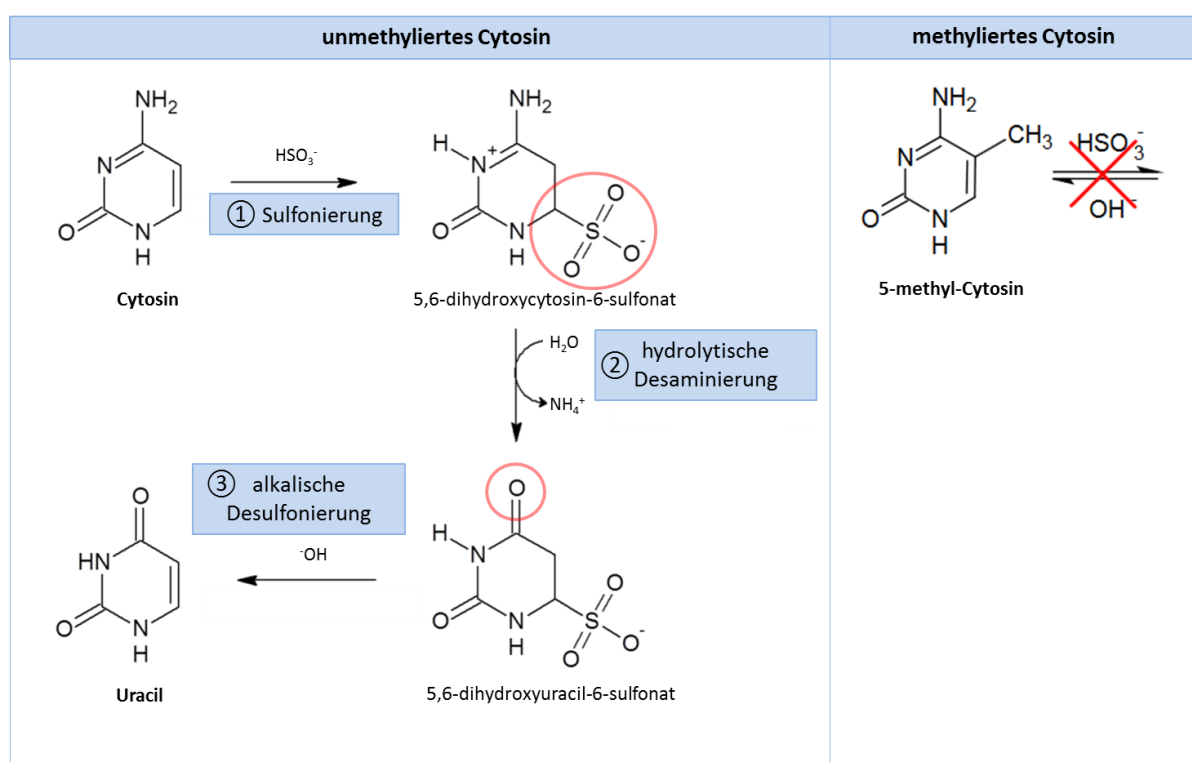


Abbildung 15: Chemische Reaktion der Bisulfit-vermittelten Desaminierung. Das grundlegende Prinzip basiert auf einer reversiblen Bindung des Hydrogensulfits (HSO_3^-) und der Abspaltung eines Ammonium-Ions (NH_4^+) bis hin zur Umwandlung zu Uracil. Links ist die Konvertierung eines unmethylierten Cytosins in drei Reaktionsschritten dargestellt (1. Sulfonierung, 2. hydrolytische Desaminierung, 3. alkalische Desulfonierung). Auf der rechten Seite ist das durch die Methyl-Gruppe (CH_3), geschützte Cytosin abgebildet. (Abb. modifiziert nach [Int. 7])

2.1 Brustkrebspatientenproben

Für die Bisulfitkonvertierung der gDNA der Brustkrebspatientinnen wurde das EpiTect Fast 96 Bisulfite Kit von Qiagen (Hilden, Deutschland) verwendet und laut der Angaben des Herstellers, wie im beiliegenden Handbuch beschrieben, verfahren. Dieses Kit wurde

ausgewählt, da die Konvertierung so auf einer Platte mit 96 Probenpositionen durchgeführt werden kann und zudem alle Patientenproben sowie Kontrollproben der Kohorte 1 bzw. 2 in einem Ansatz konvertiert werden können. Für die Reaktion wurden je Probe $\sim 1 \mu\text{g}$ gDNA eingesetzt. Die genomische bisulfit-konvertierte DNA (gbisDNA) wurde in $70 \mu\text{l}$ EB Puffer eluiert und kann bis zu drei Jahre ohne Qualitätsverluste bei $-20 \text{ }^\circ\text{C}$ gelagert werden.

2.2 gDNA der neuronalen und nicht-neuronalen Zellen für die Datenvalidierung

Die Konvertierung der gDNA, welche aus den neuronalen und nicht-neuronalen Zellen isoliert wurde, wurde mittels des EpiTect Fast Bisulfite Kits (50) von Qiagen (Hilden, Germany) durchgeführt. Bei diesem Kit handelt es sich um eine kleinere Ausgabe des in III 2.1 beschriebenen Kits und ermöglicht die Konvertierung weniger Proben in einem Ansatz. Für die Reaktion wurde je Probe $\sim 1 \mu\text{g}$ gDNA eingesetzt. Die gbisDNA wurde in zweimal $15 \mu\text{l}$ EB Puffer eluiert.

2.3 gDNA aus dem präfrontalen Cortex-Gesamtgewebe für die Datenvalidierung

Die Natriumbisulfitbehandlung erfolgte mit dem EpiTect Fast Bisulfite Kit (50) von Qiagen (Hilden, Deutschland). Aufgrund des fortgeschrittenen Alters des Gewebes war die DNA Ausbeute sehr gering ($10 - 100 \text{ ng}$). Deshalb wurden $\sim 500 \text{ ng}$ in die Reaktion eingesetzt und nach dem Ansatz für „low concentration samples“ verfahren. Die DNA wurde in zweimal $15 \mu\text{l}$ EB Puffer eluiert.

3. Next-Generation Sequencing

3.1 Deep Bisulfite Amplicon Sequencing

Das Deep Bisulfite Amplicon Sequencing (DBAS) ist eine zielgerichtete NGS-basierte Technik. Im Gegensatz zu einer genomweiten Sequenzierung ist diese Methode für die Analyse von spezifischen Targetregionen ausgelegt. Mit der von uns entwickelten Technik ist es möglich, 48 Proben und 12 verschiedene Amplikons je Probe in einem einzigen MiSeq-Lauf zu analysieren. Die Limitierung liegt hierbei ausschließlich in der Anzahl an Indexkombinationen und der Kapazität des Sequenziergeräts. Die Coverage liegt durchschnittlich zwischen 5.000 und 25.000 Reads je Probe; deshalb eignet sich diese Anwendung besonders, um seltene Epimutationen bzw. geringe Methylierungsunterschiede detektieren zu können. Das DBAS

wurde für die Methylierungsanalyse der Brustkrebspatientinnen angewandt. Zu Beginn wurden die Proben bisulfit-konvertiert, wie unter III 2.1 beschrieben, und im nächsten Schritt die Zielregionen über eine PCR amplifiziert. Es folgen das sogenannte A-Tailing, die Adapter-Ligation und eine Barcoding-PCR. Abschließend wird die fertige Library mit dem MiSeq von Illumina sequenziert (Abbildung 16).

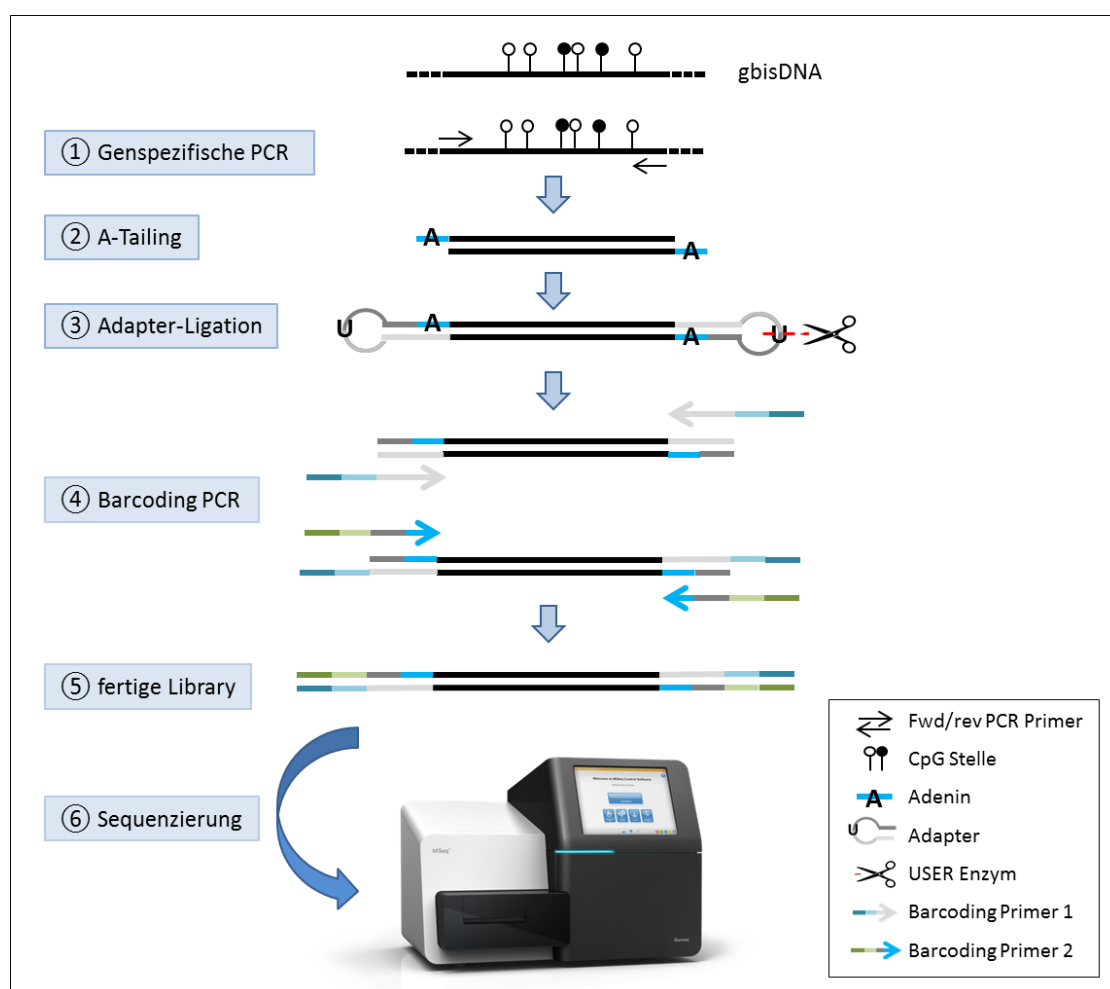


Abbildung 16: Schematische Darstellung des Ablaufs der Probenpräparation beim DBAS. Die gbisDNA wird für eine zielgerichtete PCR eingesetzt. Daran schließt sich das A-tailing, die Adapter-Ligation und die Barcoding PCR an. Die fertige Library wird abschließend mit dem MiSeq sequenziert.

3.1.1 Genspezifische PCR

Da es sich beim DBAS um eine zielgerichtete Analyse bestimmter genomischer Regionen handelt, geht der Library-Präparation eine genspezifische PCR mit der gbisDNA voraus, um die gewünschten Targets spezifisch anzureichern.

Durch die Bisulfitkonvertierung wurden alle unmethylierten Cytosine zu Uracil umgewandelt, wohingegen die methylierten Cytosine erhalten bleiben. Nach der PCR-Reaktion erscheinen nun alle Uracile, durch eine komplementäre Bindung mit Adenin, im forward Strang (original top strand; OT) als Thymine (Abbildung 17).

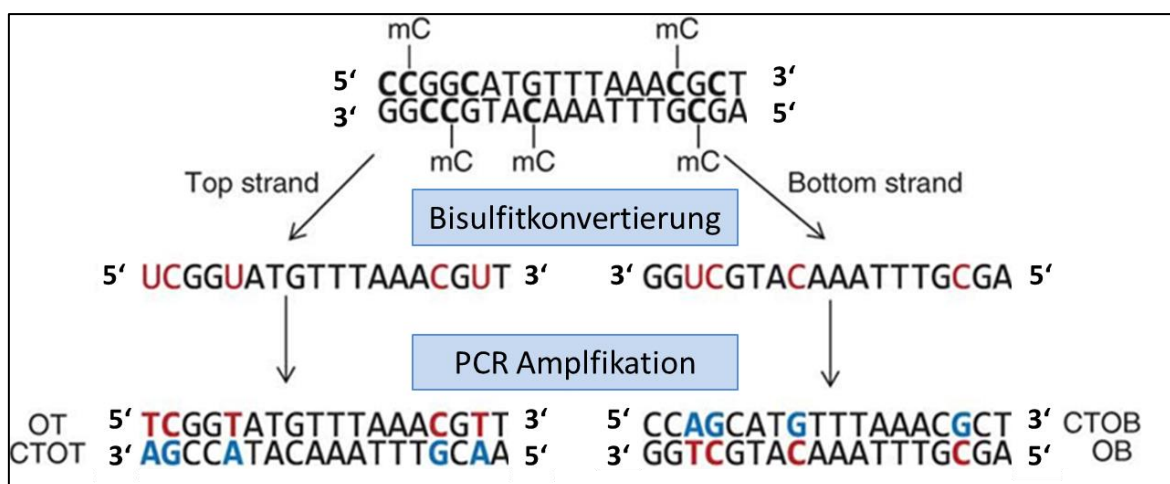


Abbildung 17: Beispielhafte Sequenzen nach einer Amplifikation der gBisDNA. Durch die Umwandlung aller unmethylierten Cytosine in Uracil und das Bestehen der methylierten Cytosine als Cytosin erhält man nach einer PCR nicht wie üblich zwei zueinander komplementäre Stränge, sondern zweimal zwei zueinander komplementäre Stränge. OT, original top Strang; CTOT, zum original top Strang komplementärer Strang; OB, original bottom Strang; CTOB zum original bottom Strang komplementärer Strang (Abb. modifiziert nach [234]).

Für das Designen der PCR-Primer wurde die PyroMark Assay Design 2.0 Software von Qiagen (Hilden, Deutschland) verwendet. Da infolge der Bisulfitkonvertierung der forward (OT) und der reverse (OB) Strang nicht mehr komplementär zueinander sind, müssen die Primer für jeden Strang spezifisch designiert werden. Auch sollte darauf geachtet werden, dass die Primer nicht über einer CpG-Stelle binden, da es so zu einer allelespezifischen Amplifikation des methylierten bzw. unmethylierten Alleles kommen kann. Alle PCRs, im Rahmen des DBAS, wurden nach folgendem Reaktionsansatz von 50 µl angefertigt:

2 µl	gBisDNA (~ 20 ng/µl)
5 µl	10x PCR Reaction Puffer (20 mM MgCl ₂)
1 µl	dNTP-Mix (je 10 mM)
2 µl	forward Primer (10 µM)
2 µl	reverse Primer (10 µM)
0,4 µl	FastStart Taq DNA Polymerase (5 U/µl)
39,6 µl	H ₂ O

Das PCR-Programm wurde wie folgt programmiert:

Initiale Denaturierung	95 °C	5 min	
Denaturierung	95 °C	30 s	} x 40 Zyklen
Annealing	X °C	30 s	
Elongation	72 °C	1 min	
Finale Elongation	72 °C	10 min	
Pause	8 °C	Pause	

Die Primersequenzen sowie die spezifischen Annealing-Temperaturen sind im Materialteil in Tabelle 5 aufgeführt. Für die Überprüfung der PCR wurden jeweils 4 µl Produkt mit 6 µl 1x LD (Thermo Scientific, Waltham, USA) vermischt und mittels eines 1,5% Agarose-Gels in einem 1x TRIS-Acetat-EDTA-Elektrophoresepuffer (TAE, Tabelle 16) bei 170 V aufgetrennt.

3.1.2 Aufreinigung mittels magnetischen Beads

Für die nachfolgende Quantifizierung der PCR-Produkte müssen die Rückstände der vorangegangenen PCR, wie z.B. Enzyme oder freie dNTPs, aber vor allem Primer-Dimere entfernt werden, um eine exakte Messung der Produkte gewährleisten zu können.

Bei den AMPure XP beads (Beckman Coulter, Krefeld, Deutschland) handelt es sich um paramagnetische Polystyrolkugeln, welche mit einer Schicht aus Magnetit ummantelt sind. Die Magnetit-Oberfläche ist zudem mit Carboxylmolekülen überzogen. Diese Moleküle vermitteln letztendlich auch die reversible Bindung zwischen der DNA und den Kugeln (Abbildung 18, A). Die AMPure XP beads finden nicht nur Anwendung bei der Aufreinigung von Reaktionen, um unerwünschte Rückstände zu entfernen, sondern dienen auch gleichzeitig dazu, Fragmente gewünschter Länge selektieren zu können. Dafür entscheidend ist die Wahl des Verhältnisses von DNA zu Beads. Wählt man eine niedrige Ratio, binden zuerst die längeren Fragmente; erhöht man jedoch das Volumen der Beads, ist es möglich, auch kleinere Fragmente aus dem Überstand zu isolieren. Die Zugabe von 20% PEG (8000 g/mol); 2,5 M NaCl erleichtert die Bindung der negativ geladenen DNA an die Carboxylmoleküle auf der Oberfläche der Kugeln und determiniert ebenfalls, neben der Wahl des Beads zu DNA Verhältnisses, die Bindung der Fragmente gewünschter Länge (Abbildung 18, B).

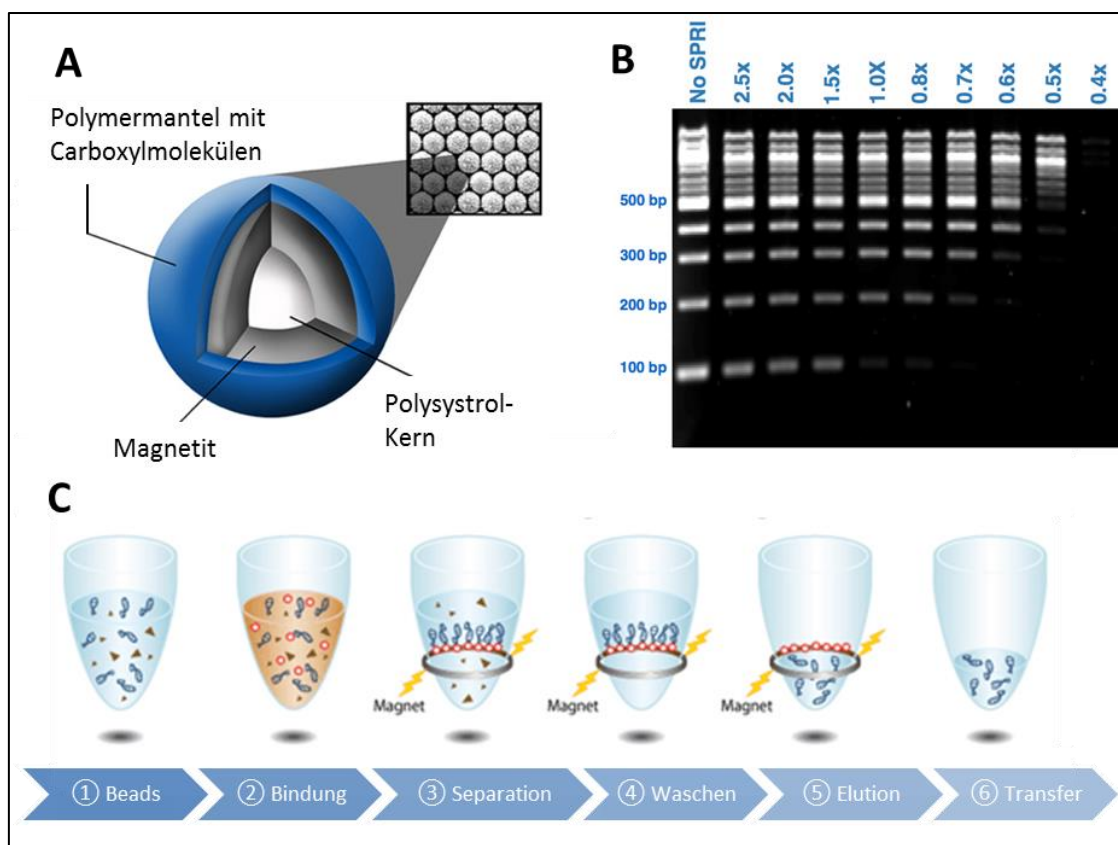


Abbildung 18: Aufreinigung mittels SPRI (Solid Phase Reversible Immobilisation) beads. **A)** Aufbau eines paramagnetischen Polystyrolkugelchens. **B)** Das Gelbild zeigt die Fragmentselektion je nach Verhältnis von DNA zu Beads. Je niedriger die Ratio, desto größer sind die isolierten Fragmente. **C)** Übersicht über den Ablauf der einzelnen Schritte bei der Bead-Aufreinigung. **1)** Zugabe der Beads zur DNA-Lösung. **2)** Bindung der DNA an die Beads. **3)** Separation mittels Magnet. **4)** Waschschritt mit EtOH **5)** Elution der DNA von den Kugelchen. **6)** Transfer der DNA. (Abb. entnommen aus [Int. 8] und modifiziert)

3.1.2.1 Manuelle Aufreinigung

Zu Beginn wurden für ein DNA:Beads Verhältnis von 1:0,6 jeweils 45 µl PCR-Produkt mit 27 µl AMPure XP beads vermengt. Es ist empfehlenswert, die Beads vor ihrer Zugabe auf RT zu erwärmen, dadurch wird die Bindung der DNA an die Kugelchen begünstigt. Nach einer 15-minütigen Inkubation bei RT, wurden die einzelnen Proben für mindestens 5 min in den Magnetständer (ThermoFisher Scientific, Waltham, USA) gestellt. Nachdem eine klare Phasentrennung erfolgte, wurde der Überstand abgenommen. Im Anschluss wurde mit je 200 µl frisch angesetztem 75% EtOH gewaschen. Als nächstes wurden die Beads mit geöffneten Deckel der Reaktionsgefäße für ca. 10 min stehen gelassen, um den restlichen EtOH verdampfen zu lassen. Jedoch sollten die Kugelchen nicht zu lange getrocknet werden, da sich die reversible Bindung der DNA an die Beads andernfalls nicht mehr gut auflösen

lässt. Der perfekte Grad der Trocknung ist erreicht, wenn sich in der Schicht der Beads kleine Risse bilden. Abschließend wurde die DNA durch Zugabe von 20 µl EB Puffer (Qiagen, Hilden, Deutschland) eluiert. Die Kügelchen wurden ebenfalls mit Hilfe des Magneten aus dem Eluat entfernt. Der Ablauf der Bead-Aufreinigung ist schematisch in Abbildung 18, C dargestellt.

3.1.2.2 Automatisierte Aufreinigung

Bei einem großen Ansatz mit vielen unterschiedlichen Genen und Proben empfiehlt es sich, von der manuellen Aufreinigung auf einen automatisierten Ansatz zu wechseln. Dafür wurde die oben beschriebene Technik für den JANUS Pipettierroboter (Perkin Elmer, Rodgau, Deutschland) angepasst und etabliert.

Zu Beginn wurden die 96-well Platten und Wannen mit den Reagenzien vorbereitet. Bei der Spaltenwanne wurde die dritte Spalte mit den Beads, die bereits auf RT aufgewärmt sind, befüllt. Das Volumen variiert je nach Anzahl der Proben. Je Probe wurden 27 µl Beads zugegeben. Das Endvolumen berechnet sich deshalb wie folgt: $(27 \mu\text{l} * \text{Anzahl der Proben}) + 2 \text{ ml Überschuss}$. Die zweite Wanne wurde generell mit 50 ml 75% EtOH bestückt. Zudem wurde eine 96-well Platte mit 50 µl EB Puffer je Position befüllt. Der Puffer wurde ausschließlich in die Positionen vorgelegt, in denen sich auf der korrespondierenden PCR-Platte ein PCR-Produkt befindet. Des Weiteren wurden die festinstallierten Spitzenboxen mit frischen Spitzen (8 bzw. 96) versehen. Im Anschluss wurden die Platten auf dem JANUS Roboter arrangiert, wie in folgendem Schema dargestellt (Abbildung 19).

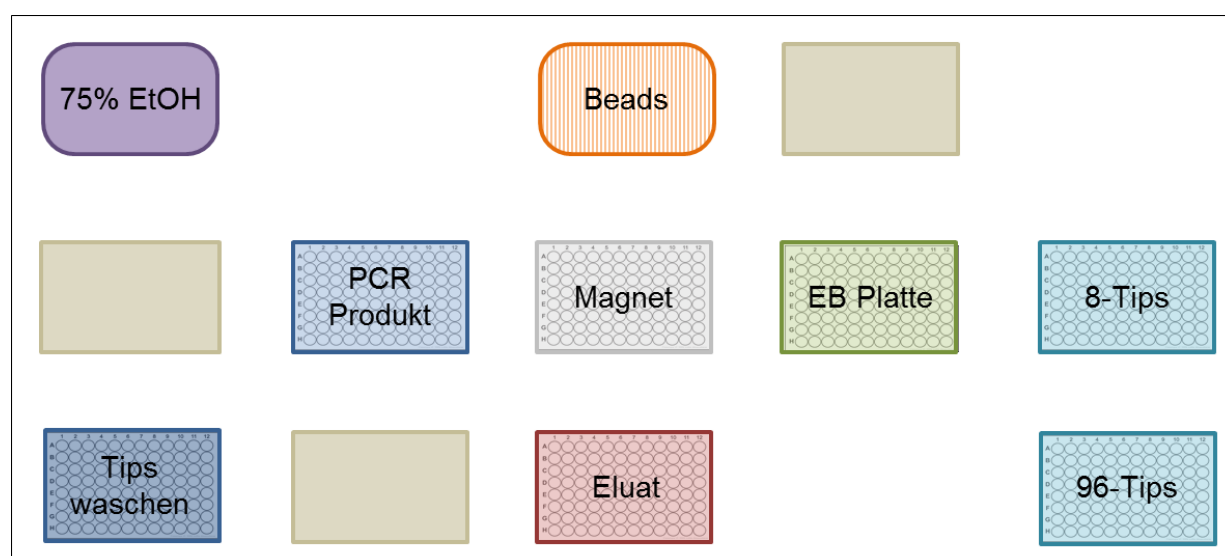


Abbildung 19: Schematische Darstellung der Platten- bzw. Reagenzien-Anordnung auf dem JANUS Pipettierroboter. Die beige Boxen sind Plätze für Platten bzw. Reagenzien, die in dieser Aufreinigung nicht besetzt werden müssen und somit leer bleiben. In der oberen Reihe befinden sich die Wanne mit 75% EtOH und die Spaltenwanne mit den Beads. Die mittlere Reihe beinhaltet 96-well Platten mit PCR-Produkt und EB Puffer, zudem den festinstallierten Magneten und eine Spitzenbox, die mit acht Spitzen besetzt werden muss. In der unteren Reihe muss eine leere 96-well Platte für das Eluat bereitgestellt werden; zusätzlich findet sich hier eine Spitzenbox mit 96 Spitzen und eine festinstallierte Waschstation, um die Spitzen zwischen den einzelnen Pipettierschritten zu reinigen.

Zu Beginn wurde die Wasserpumpe für die Spitzenwaschstation eingeschaltet, der Roboter initialisiert und das Programm gestartet. Die Schritte des Clean-up-Programmes lassen sich wie folgt auflisten: **1.** Aufnahme von acht Tips. **2.** Zugabe der Beads mit den acht Tips in die erste Spalte des PCR-Produktes (je Position 27 μ l). Mischen der Beads mit dem PCR-Produkt in Spalte 1. **3.** Spitzen waschen. **4.** Zugabe der Beads mit den acht Tips in die zweite Spalte des PCR-Produktes (je well 27 μ l). Mischen der Beads mit dem PCR-Produkt in Spalte 2. **5.** Wiederholen der Schritte **2.** und **3.**, bis alle Proben mit Beads versehen sind. **6.** Inkubation für 15 min. **7.** Transfer der PCR-Produkt-Platte zum Magneten. **8.** Inkubation für 5 min. **9.** Tauschen der 8-Tips mit 96-Tips. **10.** Abnahme des Überstands wird. **11.** Waschen mit je 200 μ l 75% EtOH. **12.** Trocknen der Beads für 10 min. **13.** Zugabe von 20 μ l EB Puffers je well. **14.** Mischen und Ablösen der DNA von den Beads. **15.** Inkubation auf dem Magneten für 5 min. **16.** Überführung des Eluats in eine leere 96-well Platte.

Zwischen den Schritten **9.-15.** erfolgt immer ein Spitzenwaschschritt an der Waschstation, sobald die Spitzen mit den Proben oder einem Reagenz in Berührung gekommen sind.

3.1.3 Quantifizierung der PCR-Produkte

Für die nachfolgende Verdünnung und das Poolen der PCR-Produkte werden die Proben via Qubit Fluorometer (ThermoFisher Scientific, Waltham, USA) auf ihre Konzentration hin vermessen. Es wird das Qubit dsDNA BR Assay Kit (ThermoFisher Scientific, Waltham, USA) verwendet, da die Konzentration erwartungsgemäß zwischen 2 und 1000 ng liegen sollte. Das Qubit Fluorometer benutzt für die Konzentrationsbestimmung einen Fluoreszenzfarbstoff, der spezifisch an doppelsträngige (ds) DNA bindet. Dieser Farbstoff emittiert nur, wenn er an die dsDNA gebunden ist. Im Vergleich zu anderen UV absorbierenden Geräten wie beispielsweise dem Nanodrop, ermöglicht diese Technik eine genauere Quantifizierung der PCR-Produkte.

Zu Beginn wurden die Reagenzien für 30 min auf RT aufgewärmt, damit sich der Fluoreszenzfarbstoff vollständig verflüssigt. Je Probe wurde eine Lösung bestehend aus 1 µl Quant-iT Reagenz und 199 µl Quant-iT Puffer angefertigt. Für die Proben wurden dann 198 µl dieser Mischung mit 2 µl PCR-Produkt versehen, gevortext und für 2 min im Dunkeln bei RT inkubiert. Die Standards, welche für die Kalibrierung des Geräts notwendig sind, wurden mit 190 µl Quant-iT Mix und 10 µl Standard #1 bzw. #2 gemischt, gevortext und ebenfalls für 2 min im Dunkeln inkubiert. Nachdem die Kalibrierung des Geräts abgeschlossen ist, konnten die Proben vermessen werden. Proben, deren Konzentration unter 2 ng liegt, werden erneut mit dem sensitiveren Qubit dsDNA HS Assay Kit (ThermoFisher Scientific, Waltham, USA) vermessen. Das Verfahren ist äquivalent zu dem bereits beschriebenen Qubit dsDNA BR Assay Kit.

3.1.4 Library-Präparation

3.1.4.1 Herstellung der Pools

3.1.4.1.1 Verdünnen der PCR-Produkte

Das anschließende Poolen der PCR-Produkte erfordert, dass die einzelnen Proben auf die gleiche Konzentration von je 0,2 ng/µl, in Abhängigkeit der Assay-Anzahl je Pool, eingestellt werden. Es wurden folgende Berechnungen, ausgehend von 3 µl einzusetzendem PCR-Produkt, angestellt:

$$x \text{ Assays} * 0,2 \frac{\text{ng}}{\mu\text{l}} = \text{finale Pool Konzentration}$$

$$\text{Proben Konzentration/finale Pool Konzentration} = \text{Grad der Verdünnung}$$

$$(\text{Grad der Verdünnung}-1) * 3 \mu\text{l} = \text{H}_2\text{O Volumen}$$

Für den Sonderfall, dass die Konzentration der PCR-Produkte unterhalb der errechneten finalen Pool-Konzentration liegt, wurden folgende Berechnungen durchgeführt:

$$0,2 \frac{\text{ng}}{\mu\text{l}} / C_{\text{Amplikon}} = x$$

$$3 \mu\text{l} * n \text{ Assays} = y$$

$$x * y = \text{Volumen PCR-Produkt}$$

$$\text{finales Volumen } \mu\text{l} - \text{Volumen PCR-Produkt} = \text{Volumen H}_2\text{O}$$

Das finale Volumen wurde so bestimmt, dass die Konzentration der Probe höher ist als die finale Pool-Konzentration.

3.1.4.1.2 Pooling der verdünnten PCR-Produkte

Pro Indexkombination wurden 3 µl des zugehörigen PCR-Produkts zu einem Pool vermischt. Bei der Kombination der Proben innerhalb eines Index muss beachtet werden, dass man je Amplikon/Gen nur eine Probe auswählt. Andernfalls können die Proben nicht mehr unterschieden werden. Die Unterscheidung der Amplikons erfolgt über das Alignment der jeweiligen Probe zum zugehörigen Genabschnitt. Es können bis zu zwölf verschiedene Amplikons/Gene in einem Pool kombiniert werden. Die Summe der Indexkombinationen ist über die kommerziell erwerblichen Barcoding-Primer und die Größe der Sequenzierkassette limitiert. In diesem Versuchsaufbau wurden maximal 48 Indexkombinationen mit bis zu zwölf Genen benutzt.

3.1.4.2 Aufreinigung mittels magnetischen Beads

Die Pools wurden vor dem nachfolgenden A-Tailing und der Adapter-Ligation einem weiteren Waschschrift unterzogen. Diese Aufreinigung trägt dazu bei, Primer-Dimere, die im ersten Waschschrift (3.1.2) nicht vollständig entfernt wurden, aus der Lösung zu entnehmen. Diese würden andernfalls auch als Target für die Adapter bei der Ligation dienen.

Der Clean-up wird ebenfalls in einem DNA:Beads Verhältnis von 1:0,6 (15 µl:9 µl), wie bereits unter IV 3.1.2.1 beschrieben, durchgeführt. Jedoch wurde der Waschschrift mit 200 µl 75% EtOH ein weiteres Mal wiederholt und die DNA in 12,5 µl EB Puffer eluiert.

3.1.4.3 Modifikation der Fragmentenden

Unter dem sogenannten A-Tailing versteht man das Hinzufügen eines zusätzlichen Adenins an das 3' Ende des dsPCR-Produkts. Diese Anwendung ist notwendig, um die Fragmente für die anschließende Adapter-Ligation zu präparieren. Innerhalb der PCR wurde bereits standardmäßig durch die Taq Polymerase ein zusätzliches Adenin angefügt, jedoch ist dieses durch das Kühlen bzw. Einfrieren und die weiteren Behandlungen der PCR-Produkte wieder abgespalten worden. Der Reaktionsansatz (20 µl) setzte sich wie folgt zusammen:

5 µl	Pool (0,2 ng/µl)
2 µl	NEB Puffer 2 (10x)
1 µl	dATP (1 mM)
0,3 µl	Klenow Fragment, 3' → 5' Exo- (5 U/µl)
11,7 µl	H ₂ O

Die Reaktion wurde unter folgenden Bedingungen durchgeführt:

30 °C	20 min
37 °C	20 min
8 °C	Pause

Es muss beachtet werden, dass der Cyclerdeckel bei der A-tailing Reaktion nicht geheizt wird, da das Enzym andernfalls nicht optimal arbeitet. Abschließend wurde das Klenow Fragment (New England Biolabs, Ipswich, USA) bei 75 °C für 20 min inaktiviert. Es ist des Weiteren wichtig, dass an dieser Stelle keine Pause gemacht wird bzw. die Proben weder gekühlt noch eingefroren werden, da es sonst zur erneuten Abspaltung des angehefteten Adenins kommt. An dieser Stelle muss direkt mit der Adapter-Ligation fortgefahren werden.

3.1.4.4 Ligation der Adapter

Die Ligation der Adapter ist neben dem A-Tailing innerhalb dieses Verfahrens einer der entscheidenden Schritte, da über die Adapter das Annealing der Indexprimer bei der finalen Amplifikation erfolgt. Sollte eine zu geringe Menge an Adaptern an die Fragmente ligiert werden, kommt es zu einer verminderten Ausbeute. Der Reaktionsansatz von 30 µl setzte sich wie folgt zusammen:

20 µl	Pool (modifiziert)
3 µl	T4 DNA Ligase Reaction Puffer (10x)
0,5 µl	NEBNext Adaptor for Illumina (15 µM)
0,5 µl	T4 DNA Ligase (400 U/µl)
6 µl	H ₂ O

Die Inkubation erfolgte ÜN in einem Cycler bei 16 °C. Auch hier ist es wichtig, dass der Deckel nicht geheizt wird, da das Enzym andernfalls nicht optimal arbeitet. Nach der Ligation wurde auch in diesem Schritt eine Enzyminaktivierung bei 65 °C für 10 min durchgeführt.

Da die Adapter von NEB, im Gegensatz zu der offenen Y-Struktur der Illumina Adapter, eine Haarnadel-Struktur aufweisen, muss abschließend noch eine Behandlung mittels USER (Uracil-Specific Excision Reagent) Enzym angeschlossen werden. Dieses Enzym generiert eine

offene Stelle an einer Uracil Position und spaltet somit die Haarnadel-Struktur der Adapter auf. Bei dem USER Enzym handelt es sich um eine Mischung aus einer Uracil DNA Glykosylase (UDG) und einer DNA Glykosylase-Lyase Endonuklease VIII. Über die UDG wird das Entfernen der Uracil Base katalysiert, indem eine apyrimidinische Stelle gebildet wird, wohingegen das Phosphodiester Rückgrat erhalten bleibt. Durch die Lyase Aktivität der Endonuklease VIII wird ein Bruch an der 3' und 5' Stelle der apyrimidinischen Site im Phosphodiester Rückgrat hervorgerufen, so dass die basen-freie Deoxyribose abgespalten wird.

Für diese Reaktion wurden jeweils 3 µl zu den Proben gegeben und bei 37 °C für 15 min im Cycler ohne geheizten Deckel inkubiert.

3.1.4.5 Aufreinigung der Ligation via magnetischen Beads

Vor der finalen Amplifikation der Fragmente muss die Ligation einem Waschschrift unterzogen werden. Es müssen alle Rückstände aus der vorangegangenen Adapter-Ligation entfernt werden, da rückbleibende Adapter ebenfalls als Target für die Barcoding-Primer fungieren und es so zu einer starken Anreicherung an Primer-Adapter-Dimeren kommen würde.

Die Aufreinigung erfolgte auch hier über die Ampure XP Beads, wie bereits unter III 3.1.2.1 beschrieben. Das Verhältnis DNA:Beads wurde jedoch von 1:0,6 auf 1:0,95 (33 µl:31,35 µl) erhöht. Die DNA wurde in 20 µl EB Puffer eluiert.

3.1.4.6 PCR zur Überprüfung der Ligations-Effizienz

Dieser Schritt ist optional und muss nicht bei jedem Versuchsansatz durchgeführt werden. Im Fall, dass ein neues Genset getestet wird, ist es empfehlenswert, diesen Schritt einzufügen. Durch diese PCR kann die Effizienz der Ligation überprüft werden, um somit die optimale Zyklanzahl für die finale Barcoding-PCR festzulegen. Dadurch kann zum einen eine Überamplifikation und zum anderen eine zu geringe Ausbeute vermieden werden. Für diese PCR können die im NEBNext Multiplex Oligos for Illumina (Tabelle 6) bereitgestellten Barcoding Primer verwendet werden. Es ist jedoch kostengünstiger, Primer mit dieser Sequenz bei einem anderen Hersteller zu kaufen, da es sich hier nur um einen Test handelt. Der PCR-Ansatz (50 µl) setzte sich wie folgt zusammen:

3 µl	Pool (Adapter-ligiert)
5 µl	10x PCR Reaction Puffer (20 mM MgCl ₂)
0,5 µl	dNTP-Mix (je 10 mM)
1 µl	forward Primer (10 µM)
1 µl	reverse Primer (10 µM)
1 µl	FastStart Taq DNA Polymerase (5 U/µl)
38,5 µl	H ₂ O

Der Reaktionsansatz wurde auf 4 PCR-Tubes zu je 10 µl aufgeteilt und jeweils die Zyklenzahl zwischen 16 und 22 Zyklen variiert. Das PCR-Programm wurde wie folgt programmiert:

Initiale Denaturierung	95 °C	2 min	} x 16 - 22 Zyklen
Denaturierung	95 °C	30 s	
Annealing	65 °C	30 s	
Elongation	72 °C	45 s	
Finale Elongation	72 °C	7 min	
Pause	8 °C	Pause	

Um die optimale Zyklenzahl zu finden, wurde eine Gelelektrophorese, wie bereits unter III 3.1.1 beschrieben, durchgeführt.

3.1.4.7 Finale Barcoding-PCR

Als Abschluss der Probenpräparation werden die Pools amplifiziert werden. Dieser Schritt dient sowohl der Vervielfältigung der Fragmente als auch dem Barcoding der einzelnen Pools. Hierfür werden die bereits unter III 3.1.4.6 erwähnten Indexprimer aus dem NEBNext Multiplex Oligos for Illumina Kit (Tabelle 6) verwendet. Bei einer geringen Anzahl (2-12) an Pools sollte das beiliegende Handbuch oder der Illumina Experiment Manager zu Rate gezogen werden. Die Illumina Sequenziergeräte, wie beispielsweise der MiSeq, sind mit einem roten, für die Detektion der Basen A/C, und einem grünen Laser/LED, für die Detektion der Basen G/T, ausgestattet. Dies kann bei einer nicht balancierten Auswahl der Indexprimer zu einem Sequenzierausfall der Index-Sequenz führen. Der Reaktionsansatz von 50 µl setzte sich wie folgt zusammen:

3 µl	Pool (Adapter-ligiert)
5 µl	PfuTurbo C _x reaction Puffer (10x)
1 µl	dNTP-Mix (je 10 mM)
1 µl	Index Primer i500 (10 µM)
1 µl	Index Primer i700 (10 µM)
1 µl	PfuTurbo C _x hotstart DNA Polymerase (2,5 U/µl)
38 µl	H ₂ O

Das PCR-Programm wurde wie folgt programmiert:

Initiale Denaturierung	92 °C	2 min	} x 21 Zyklen
Denaturierung	92 °C	30 s	
Annealing	65 °C	30 s	
Elongation	68 °C	45 s	
Finale Elongation	68 °C	7 min	
Pause	8 °C	Pause	

Für die Überprüfung der PCR wurden eine Gelelektrophorese, wie bereits unter III 3.1.1 beschrieben, durchgeführt.

3.1.4.8 Aufreinigung der finalen PCR mittels magnetischen Beads

Bevor zur Sequenzierung übergegangen werden kann, müssen die Rückstände aus der vorangegangenen PCR entfernt werden. Es ist besonders wichtig, dass die Primer bzw. Primer-Dimere rückstandslos entfernt werden, da über die Barcoding-Primer die Bindung der Fragmente an die Flusszelle vermittelt wird; deshalb können auch Primer-Rückstände dort binden und es würde zu einer verminderten Anzahl an Amplikon-Reads kommen.

Die Aufreinigung erfolgte wiederum über die Ampure XP Beads, wie bereits unter III 3.1.2.1 beschrieben. Das Verhältnis DNA:Beads wurde jedoch auf 1:0,9 (50 µl:45 µl) erhöht. Die DNA wurde in 40 µl EB Puffer eluiert. Die Aufreinigung wurde auf dieselbe Weise ein weiteres Mal wiederholt. Das Verhältnis DNA:Beads blieb gleich (40 µl:35,55 µl) und die DNA wurde in 25 µl EB Puffer eluiert.

3.1.4.9 Quantifizierung der finalen PCR

Die Vermessung der Amplifikate, auf ihre Größe und Konzentration hin, wurde mit dem Bioanalyzer von Agilent (Waldbronn, Deutschland) durchgeführt. Da die Fragmentlängen unter 1000 bp liegen und die Pools erwartungsgemäß eine Konzentration zwischen 1 und 5 ng/µl aufweisen, wurde das High Sensitivity DNA Analysis Kit (Agilent, Waldbronn, Deutschland) verwendet und gemäß der Angaben des Herstellers verfahren. Nachfolgend ist ein Beispiel für einen aufgereinigten Pool nach der Barcoding-PCR dargestellt (Abbildung 20).

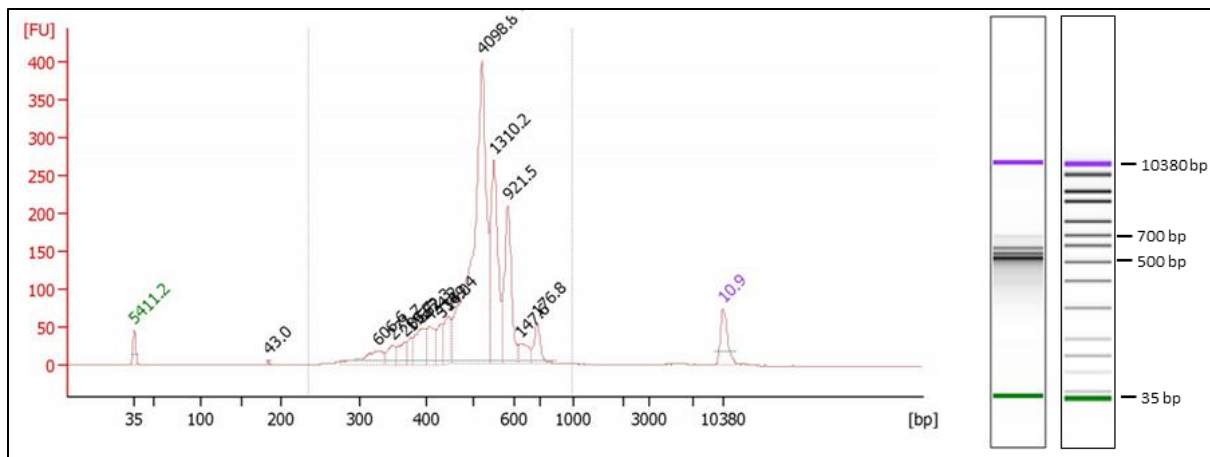


Abbildung 20: Elektropherogramm und Gelbild eines finalen Pools beim DBAS. Auf der linken Seite ist ein Elektropherogramm exemplarisch für einen aufgereinigten finalen Pool dargestellt. Die Fragmente mit einer Länge von 35 bp und 10380 bp zeigen die Größenstandards an. Es sind keine Primer-Dimer bzw. Primer-Adapter-Dimere mehr enthalten. Diese hätten erwartungsgemäß einen Peak bei 180 bp gezeigt. Die Peaks zwischen 300 bp und 700 bp entsprechen den einzelnen Amplikons, die im Pool enthalten sind. Auf der rechten Seite ist der aufgereinigte Pool als Gelbild und der Größenstandard dargestellt.

3.1.4.10 Verdünnung und finales Pooling

Die einzelnen Pools wurde jeweils auf 4 nM, wie im MiSeq Handbuch angegeben, verdünnt, da die Sequenzierung mit der v3 600 Zyklen Kassette (Illumina, San Diego, USA) erfolgt. Es wurden folgende Berechnungen angestellt, ausgehend von 3 µl einzusetzendem PCR-Produkt: $(3 * \text{Konzentration Pool} / 4 \text{ nM}) - 3 = \text{Volumen EB Puffer}$. Nachdem alle Pools auf eine Endkonzentration von 4 nM eingestellt waren, erfolgte das finale Pooling. Hierfür wurden 3 µl je Pool (4 nM) in einem 1,5 ml Reaktionsgefäß zusammengefügt und gemischt. Danach wurde noch einmal eine Quantifizierung mittels Bioanalyzer durchgeführt, wie unter III 3.1.4.9 beschrieben, um die Konzentration des finalen Pools zu überprüfen. Die Messung erfolgte in Triplikaten, um einen exakten Messwert zu erhalten. Für die anschließende Denaturierung und Verdünnung wurde der Mittelwert aus diesen drei Werten genommen.

3.1.5 Sequenzierung via MiSeq

Die Sequenzierung der Libraries wurde mit dem MiSeq von Illumina (San Diego, USA) durchgeführt. Dieses Gerät basiert auf der „Sequencing-by-Synthesis“ (SBS) -Technologie. Innerhalb eines Zyklus werden alle vier Nukleotide zugegeben, diese sind zur Unterscheidung der eingebauten Nukleotide jeweils mit einem unterschiedlichen Fluoreszenzfarbstoff

markiert. Die Farbgruppe fungiert gleichzeitig als Terminator, dadurch kann innerhalb eines jeden Zyklus nur ein markiertes Desoxynukleosidtriphosphat (dNTP) eingebaut werden. Nach der Detektion der Lichtsignale werden die Farbgruppen enzymatisch abgespalten und ein neuer Zyklus beginnt (Abbildung 21). Der MiSeq ermöglicht bei einer Laufzeit von 4-55 h einen maximalen Output von 15 GB und bis zu 25 Mio. Reads per Lauf. Die maximale Read-Länge beläuft sich aktuell auf 2 x 300 bp und ist somit bestens geeignet für das „targeted gene sequencing“.

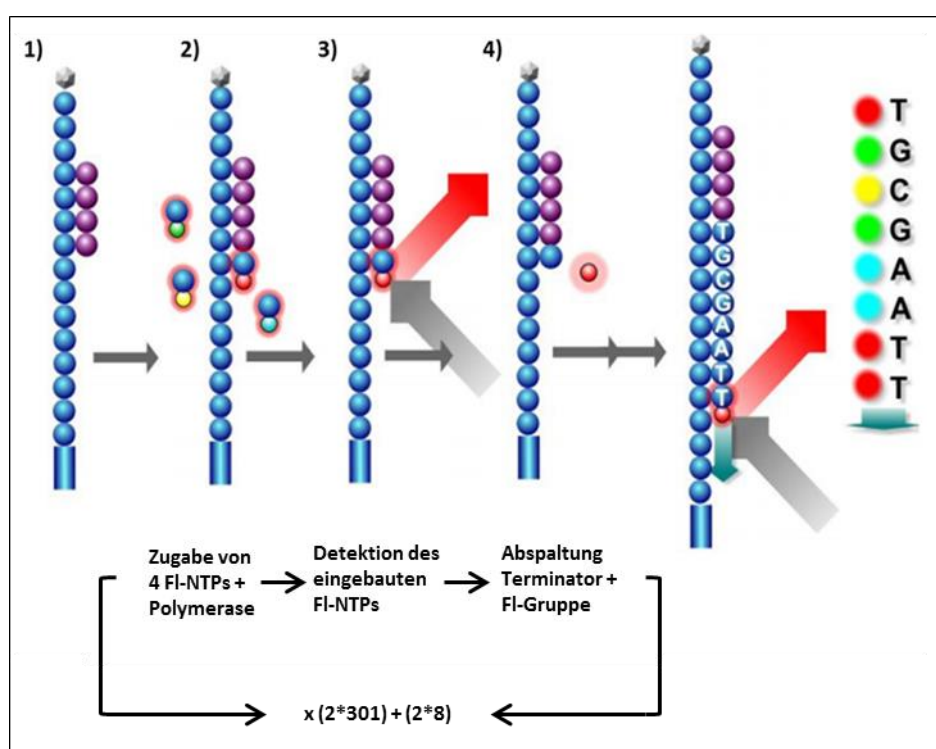


Abbildung 21: Schematische Übersicht über einen NGS-Zyklus bei der SBS-Technologie. Dargestellt sind die einzelnen Schritte eines NGS-Sequenzierzyklus. **1)** Zugabe aller vier fluoreszenzfarbstoffmarkierten (FI) NTPs. **2)** Komplementärer Einbau eines FI-NTPs an den DNA-Strang. **3)** Detektion des eingebauten FI-NTPs. **4)** Abspaltung des Terminators und der Farbgruppe des eingebauten NTPs. Wiederholung der Schritte **1) – 4)** bis zum Ende der Sequenzierreaktion. (Abb. entnommen von [Int. 9] und modifiziert)

Die Denaturierung und Verdünnung des 4 nM Pools erfolgte nach dem „MiSeq System Denature and Dilute Libraries Guide“ von Illumina [Int. 10]. Es wurden 5 µl Library mit 5 µl 0,2 N NaOH in einem Verhältnis von 1:1 verdünnt und für 5 min bei RT inkubiert. Es folgte eine 1:100 Verdünnung durch Zugabe des vorgekühlten HT1 Puffers (990 µl). Die denaturierte Library wurde final auf 12 pM eingestellt und auf Eis gelagert. Bei bisulfite-

behandelten Libraries, handelt es sich um sogenannte „low diversity libraries“; deshalb muss die fehlende Basendiversität durch das Zufügen von PhiX (Illumina, San Diego, USA) balanciert werden. In diesem Fall ist die Diversität nicht nur durch das geringe Vorhandensein der Base C/G eingeschränkt, sondern auch durch das Bestehen der Library aus Amplikons. Um diesem Ungleichgewicht entgegenzuwirken, setzte sich die finale Library aus 80% Pool und 20% PhiX zusammen. Das PhiX wurde ebenfalls mit 0,2 N NaOH denaturiert und mittels HT1 Puffer auf 12 pM Endkonzentration eingestellt. Von dieser Mischung werden 600 µl auf die Kassette geladen. Am MiSeq selbst musste ein sogenanntes „Sample Sheet“ generiert werden, in dem die Kombinationen der Indexsequenzen per Pool angegeben werden, damit die Reads nach erfolgtem Lauf den einzelnen Pools zugeordnet werden können. Die paired-end Sequenzierung erfolgte mit der v3 600 Zyklen Kassette.

3.2 Reduced Representation Bisulfite Sequencing

Im Vergleich zu der oben genannten DBAS-Technik handelt es sich beim Reduced Representation Bisulfite Sequencing (RRBS) um eine ungerichtete Methode. Das bedeutet, dass diese Technik nicht nur auf ausgewählte Bereiche im Genom beschränkt ist, sondern ein für das gesamte Epigenom repräsentativer Teil untersucht werden kann. Das RRBS stellt eine kostengünstige und effiziente Technik dar, welche sich auch besonders gut für Proben mit einer niedrigen DNA-Menge eignet, was oftmals bei klinischen Proben der Fall ist. Zudem werden neben CpG-Inseln und Promotorbereichen auch beispielsweise Exons, 3' untranslatierte Bereiche und repetitive Elemente abgedeckt. Basierend auf den genannten Vorteilen wurde das RRBS zur differentiellen Methylierungsanalyse zwischen Mensch und Schimpanse in den neuronalen und nicht-neuronalen Zellen eingesetzt. Um bei diesem Verfahren die Komplexität des Genoms herabzusetzen, wird die DNA zu Beginn durch einen Restriktionsverdau mit MspI in unterschiedlich lange Fragmente zerteilt. Daran schließen End-repair, A-tailing, Adapter-Ligation und die Bisulfitkonvertierung an. Abschließend erfolgt eine Barcoding-PCR, sowie die Sequenzierung mit dem Genome Analyzer II (GAII) von Illumina (San Diego, USA) (Abbildung 22).

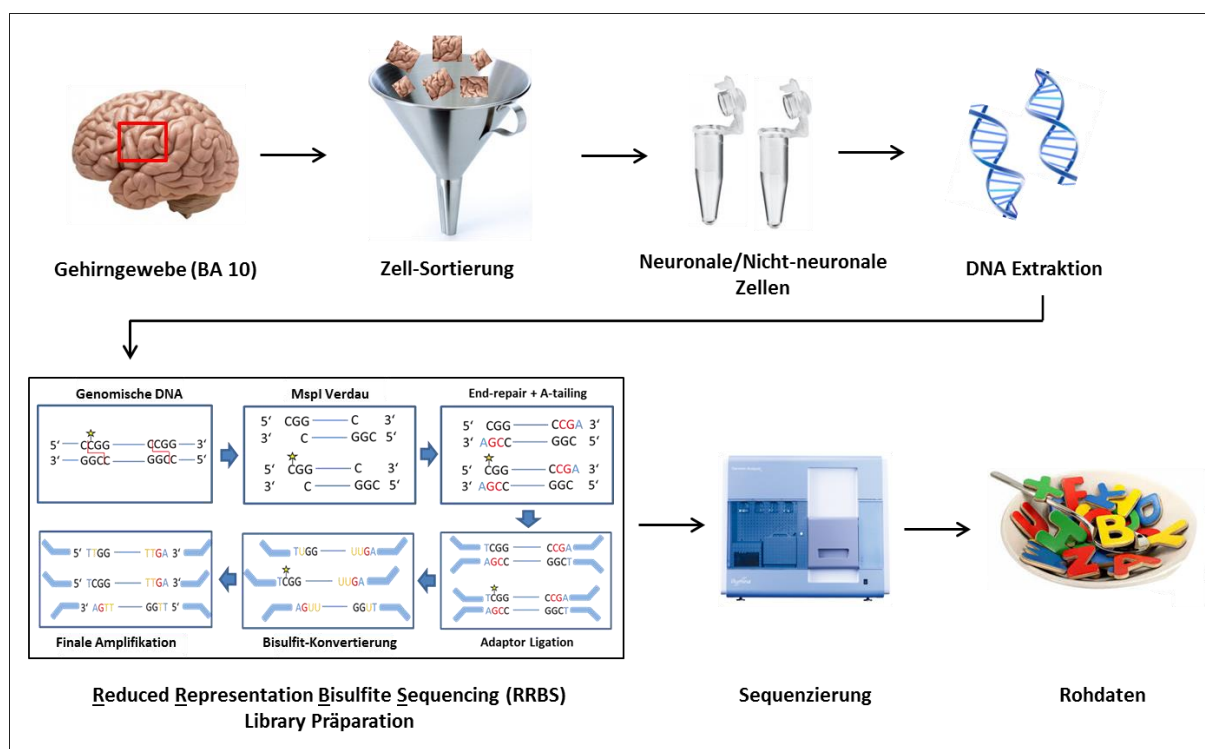


Abbildung 22: Schematische Darstellung des Arbeitsablaufs zur Generierung der RRBS-Daten. Zu Beginn wird das Gehirngewebe aus dem präfrontalen Cortex (Brodmann-Areal 10, BA10) in neuronale und nicht-neuronale Zellen sortiert und die gDNA extrahiert. Anschließend erfolgt die Proben-Präparation mittels der RRBS Methode. Die Sequenzierung wird mit dem GAI durchgeführt.

3.2.1 MspI Restriktionsverdau

Der Restriktionsverdau mit einer methylierungs-sensitiven Endonuklease dient in diesem Versuchsansatz nicht nur rein zur Fragmentierung der gDNA. Im Gegensatz zu einer Zerkleinerung der gDNA mittels Ultraschall, bringt das spezifische Schneidemuster von MspI (5'-C[^]CGG-3') in diesem Fall einen Vorteil mit sich. Es kommt somit zu einer Anreicherung von Fragmenten, die überwiegend CpG-Dinukleotide und somit Informationen über die DNA-Methylierung enthalten. Diese werden in den nachfolgenden Schritten speziell angereichert. Bei MspI (New England Biolabs, Ipswich, USA) handelt es sich um ein Isoschizomer von HpaII; jedoch ist MspI – im Gegensatz zu HpaII – in der Lage, die Phosphodiesterbindung des innen gelegenen Cs zu hydrolysieren, auch wenn dieses eine Methylgruppe aufweist. Der Reaktionsansatz von 30 µl setzte sich wie folgt zusammen:

x µl	gDNA (200 ng)
3 µl	NEB Puffer 2 (10x)
1 µl	MspI (20 U/µl)
x µl	H ₂ O

Das Volumen der gDNA und des Wassers wurde anhand der Konzentration der gDNA je Probe individuell errechnet. Je Reaktion wurden 200 ng gDNA eingesetzt. Die Inkubation erfolgte im Thermoblock bei 37 °C ÜN ohne geheizten Deckel.

3.2.2 Modifikation der Fragmentenden

Durch den Restriktionsverdau mit MspI entstehen an beiden Enden des Fragments 3'-Überhänge. Diese müssen für die nachfolgende Adapter-Ligation ergänzt und mit einem zusätzlichen Adenin versehen werden. Die Auffüllung der Lücken erfolgt durch das sogenannte Klenow-Fragment (New England Biolabs, Ipswich, USA); dieses Enzym besitzt sowohl eine 5' → 3' Polymerase-Aktivität als auch eine 3' → 5' Exonuklease-Aktivität.

Für die Reaktion wurde ein dNTP-Mix (New England Biolabs, Ipswich, USA) eingesetzt, der die 10-fache Menge an dATPs enthält, um die Wahrscheinlichkeit für das Anhängen eines Adenins am 3'-Ende zu erhöhen. Die Verwendung von zwei verschiedenen Temperaturen führt zu optimalen Konditionen für das Enzym und trägt ebenfalls zur Erhöhung der Effizienz bei. Der Reaktionsansatz von 32 µl setzte sich wie folgt zusammen:

30 µl	DNA (MspI verdaut)
1 µl	Klenow Fragment, 3' → 5' Exo- (5 U/µl)
1 µl	dNTP-Mix (10 mM dATP, 1 mM dCTP, 1 mM dGTP)

Die Inkubation erfolgte im Cycler (2.6) bei 30 °C und anschließend bei 37°C für je 20 min ohne geheizten Deckel.

3.2.3 DNA-Aufreinigung mittels magnetischen Beads

Der bead clean-up erfolgte wie bereits unter III 3.1.2.1 beschrieben. Das DNA:beads Verhältnis belief sich dabei auf 1:2 (32 µl:64 µl). Der Waschschrift mit 200 µl 75% EtOH erfolgte zweimal. Die DNA wurde in 20 µl EB eluiert.

3.2.4 Multiplexed Adapter-Ligation

Im Vergleich zu der unter III 3.1.4.4 beschriebenen Adapter-Ligation innerhalb des DBAS Protokolls wird beim RRBS das NEBNext Multiplex Oligos for Illumina (Methylated Adaptor, Index Primers Set1) (Tabelle 9) verwendet. Neben der Verwendung eines „Single Index“ je Probe, ist der entscheidende Unterschied, dass es sich hier um methylierte Adapter handelt.

Diese Modifikation ist notwendig, da die Adapter-Ligation vor der Bisulfitkonvertierung erfolgt. Durch die Konvertierung würde es andernfalls zur Abänderung der Basensequenz der Adaptern führen. Die hätte zur Konsequenz, dass der Index- bzw. Universalprimer bei der finalen PCR nicht binden kann. Der Reaktionsansatz von 30 µl für die Ligation setzte sich wie folgt zusammen:

20 µl	DNA (modifiziert)
3 µl	T4 DNA Ligase Reaction Puffer (10x)
0,9 µl	NEBNext Adaptor for Illumina, methylated (15 µM)
1 µl	T4 DNA Ligase (400 U/µl)
5,1 µl	H ₂ O

Die Proben werden bei 16 °C ÜN im Thermocycler ohne geheizten Deckel inkubiert. Nach der Ligation wurde auch in diesem Schritt eine Enzymaktivierung bei 65 °C für 10 min durchgeführt. Wie bereits unter III 3.1.4.4 des DBAS Protokolls beschrieben, liegen die NEB Adapter (methyliert bzw. unmethyliert) in einer Hairpin-Struktur vor und müssen für die nachfolgenden Schritte aufgespalten werden. Für diese Reaktion wurden jeweils 3 µl zu den Proben gegeben und bei 37 °C für 15 min im Cycler ohne geheizten Deckel inkubiert.

3.2.5 DNA-Aufreinigung und Determination der Fragmentlängen

Die AMPure XP beads finden nicht nur Anwendung bei der Aufreinigung von Reaktionen von unerwünschten Rückständen, sondern sie dienen auch gleichzeitig dazu, die Fragmente nach gewünschter Länge selektieren zu können, wie bereits unter III 3.1.2 genauer erklärt (Abbildung 18). Zu Beginn wurde jede Probe (33 µl) mit H₂O auf ein Endvolumen von 360 µl aufgefüllt und versetzt mit 720 µl 20% PEG (8000 g/mol); 2,5 M NaCl (Tabelle 16). Anschließend wurden jeweils 66 µl beads (DNA:Beads; 1:2) zu den Proben gegeben und durch auf- und abpipettieren gemischt. Das weitere Verfahren verhielt sich, wie unter III 3.1.2.1 beschrieben. Der Waschschrift erfolgte einmalig mit 1 ml 75% EtOH. Die DNA wurde in 25 µl EB-Puffer von den Beads isoliert.

3.2.6 Ligations-Effizienz PCR

Die Effizienz der Adapter-Ligation muss mittels PCR überprüft werden, bevor man die nachfolgende Natriumbisulfitbehandlung anschließt. Die Primer können, wie bereits unter III 3.1.4.6 im DBAS Protokoll erklärt, direkt aus dem NEBNext Kit verwendet werden, oder aber,

als kostengünstigere Variante, bei einer anderen Firma gekauft werden. Der Reaktionsansatz von 50 μl setzte sich hierfür wie folgt zusammen:

3 μl	DNA (Adapter-ligiert)
5 μl	10x PCR Reaction Puffer (20 mM MgCl_2)
0,5 μl	dNTP-Mix (je 10 mM)
1 μl	forward Primer (10 μM)
1 μl	reverse Primer (10 μM)
1 μl	FastStart Taq DNA Polymerase (5 U/ μl)
38,5 μl	H_2O

Der Reaktionsansatz von 50 μl pro Probe wurde auf vier PCR-Gefäße zu je 10 μl verteilt. Die Amplifikation der DNA-Fragmente erfolgte bei allen Ansätzen unter den gleichen Bedingungen, jedoch variierte die Wiederholung der Zyklen zwischen 10 und 19:

Initiale Denaturierung	95 °C	2 min	} x 10 - 19 Zyklen
Denaturierung	95 °C	30 s	
Annealing	65 °C	30 s	
Elongation	72 °C	45 s	
Finale Elongation	72 °C	7 min	
Pause	8 °C	Pause	

Nach erfolgtem Lauf wurde die Qualität der Adapter-Ligation mittels Gelelektrophorese überprüft, wie unter III 3.1.1 beschrieben.

3.2.7 Natriumbisulfitbehandlung

Die Bisulfitkonvertierung wurde mit dem EpiTect Fast Bisulfite Kit (50) von Qiagen (Hilden, Deutschland) durchgeführt; dabei wurde laut den Angaben des Herstellers, wie im beiliegenden Handbuch beschrieben, verfahren. Die DNA wurde in zweimal 20 μl eluiert.

3.2.8 Aufreinigung mittels Säulchen

Die Aufreinigung der Bisulfit-behandelten Proben ist optional, jedoch empfehlenswert, um die Ausbeute bei der finalen Amplifikation zu erhöhen. Es bleiben meist Rückstände nach der Konvertierung zurück, welche als Inhibitoren auf die Taq bzw. Pfu Polymerase wirken können. Hierfür wurde das MinElute PCR Purification Kit von Qiagen (Hilden, Deutschland) verwendet und nach den Angaben des Herstellers, wie im beiliegenden Handbuch

beschrieben, verfahren. Die Proben wurden jeweils mit dem 5-fachen Volumen an PB Puffer (185 µl) vermengt. Zusätzlich wurden 10 µl 3 M NaAC (pH 5) zugegeben, um ein saures Milieu zu generieren. Die DNA wurde in einem Volumen von 30 µl EB Puffer von den Säulchen gelöst.

3.2.9 Bestimmung der optimalen Zyklenzahl

Vor der Amplifikation der finalen RBBS-Library wird die optimale Zyklenzahl mittels einer PCR in einem kleineren Ansatz (50 µl) festgelegt. Wie bereits unter III 3.1.4.6 beschrieben, ist dieser Schritt wichtig, um einer Überamplifikation bzw. einer zu geringen Ausbeute entgegenzuwirken. Auch hier können die Primer eines anderen Herstellers verwendet werden. Der PCR-Ansatz setzte sich wie folgt zusammen:

3 µl	bisDNA
5 µl	10x PCR Reaction Puffer (20 mM MgCl ₂)
0,5 µl	dNTP-Mix (je 10 mM)
1 µl	forward Primer (10 µM)
1 µl	reverse Primer (10 µM)
1 µl	FastStart Taq DNA Polymerase (5 U/µl)
38,5 µl	H ₂ O

Mit dem Reaktionsansatz von 50 µl wird ebenso verfahren, wie unter III 3.1.4.6 bereits beschrieben. Jedoch werden in diesem Fall Zyklen zwischen 13 und 23 gewählt.

Initiale Denaturierung	95 °C	2 min	} x 13 - 23 Zyklen
Denaturierung	95 °C	30 s	
Annealing	65 °C	30 s	
Elongation	72 °C	45 s	
Finale Elongation	72 °C	7 min	
Pause	8 °C	Pause	

Die Visualisierung der Fragmente erfolgt nach dem Lauf mittels Gelelektrophorese, wie bereits unter III 3.1.1 beschrieben.

3.2.10 Finale Amplifikation der DNA-Fragmente

Nach der Festlegung der optimalen Zyklenzahl wurde ein großer Reaktionsansatz von 200 µl für jede Probe angefertigt, um die Menge der einzelnen Fragmente für die abschließende Sequenzierung zu erhöhen. Bei dieser PCR wurde je Probe ein Index- und ein

Universalprimer (Tabelle 8) verwendet, wie bereits unter III 3.2.4 erwähnt. Es ist auch hier wichtig, auf die richtige Wahl der Indexprimer zu achten. Eine detaillierte Erklärung findet sich unter III 3.1.4.7. Der Reaktionsansatz setzt sich wie folgt zusammen:

26 µl	bisDNA
20 µl	PfuTurbo C _x reaction Puffer (10x)
2 µl	dNTP-Mix (je 10 mM)
4 µl	Index Primer (10 µM)
4 µl	Universal Primer (10 µM)
4 µl	PfuTurbo C _x hotstart DNA Polymerase (2,5 U/µl)
140 µl	H ₂ O

Der Reaktionsansatz jeder Probe wird auf vier PCR-Tubes verteilt (je 50 µl) und anschließend unter nachfolgenden Bedingungen amplifiziert. Das PCR-Programm wurde wie folgt programmiert:

Initiale Denaturierung	95 °C	2 min	} x 19-22 Zyklen
Denaturierung	95 °C	30 s	
Annealing	65 °C	30 s	
Elongation	68 °C	45 s	
Finale Elongation	68 °C	7 min	
Pause	8 °C	Pause	

Die Zyklenanzahl schwankt je nach Probe zwischen 19 und 22 und wurde entsprechend der „optimalen Zyklen“ - PCR individuell für jede Probe angepasst.

3.2.11 DNA-Aufreinigung mittels magnetischen Beads

Im Anschluss an die finale PCR wurden die korrespondierenden Proben in jeweils einem 1,5 ml Reaktionsgefäß vereint und 184 µl Beads zugegeben. Das weitere Verfahren verhielt sich wie unter III 3.1.2.1 beschrieben. Der Waschschrift erfolgte mit 1 ml 75% EtOH. Die DNA wurde in 40 µl EB-Puffer von den Beads isoliert. Dieser Vorgang wurde ein weiteres Mal auf die gleiche Weise durchgeführt, jedoch wurden, angepasst an die eingesetzte DNA-Menge (40 µl), in diesem Schritt nur 38 µl an Beads eingesetzt. Der Waschschrift erfolgte mit 200 µl 75 % EtOH. Die DNA wurde in 25 µl EB Puffer eluiert.

3.2.12 Quantifizierung der RRBS Library

Die Vermessung der Amplifikate, auf ihre Größe und Konzentration hin, wurde mit dem Bioanalyzer von Agilent durchgeführt. Es wird das High Sensitivity DNA Analysis Kit verwendet, wie unter III 3.1.4.9 beschrieben, und nach den Angaben des Herstellers verfahren. Im Folgenden ist ein Beispiel für eine fertige RRBS Library dargestellt (Abbildung 23).

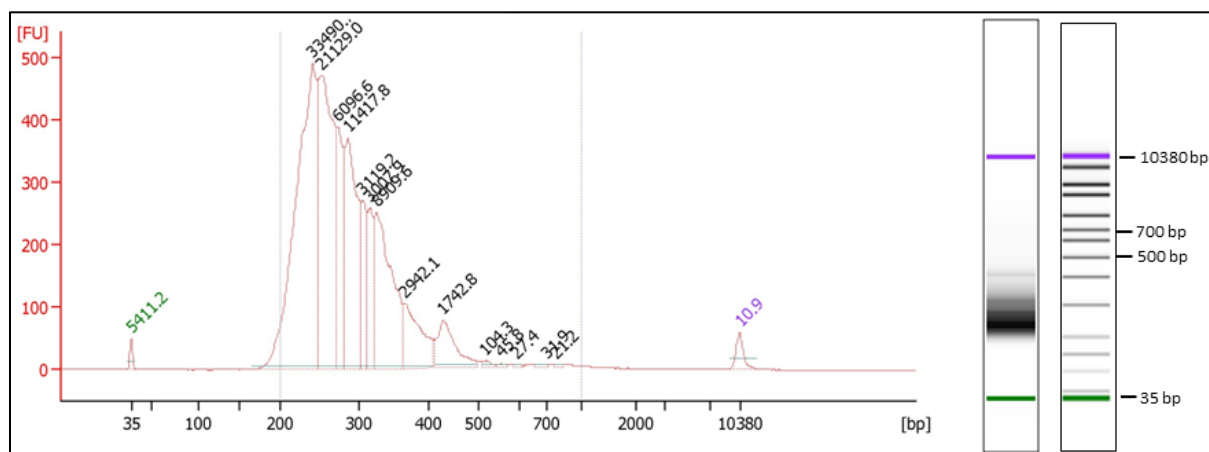


Abbildung 23: Elektropherogramm und Gelbild einer finalen RRBS Library. Auf der linken Seite ist ein Elektropherogramm exemplarisch für einen aufgereinigten finalen Pool dargestellt. Die Fragmente mit einer Länge von 35 bp und 10380 bp zeigen die Größenstandards an. Es sind keine Primer-Dimer bzw. Primer-Adapter-Dimere mehr enthalten. Diese hätten erwartungsgemäß einen Peak bei 180 bp gezeigt. Die Peaks zwischen 150 bp und 500 bp entsprechen den einzelnen Fragmenten. Auf der rechten Seite ist der aufgereinigte Pool als Gelbild und der Größenstandard dargestellt.

3.2.13 Sequenzierung der RRBS Libraries

Die Sequenzierung der Libraries wurde mit dem GAI von Illumina durchgeführt. Dieses Gerät basiert, wie auch, der beim DBAS verwendete MiSeq, auf der SBS-Technologie. Detaillierte Informationen können unter III 3.1.5 nachgelesen werden.

Zu Beginn wurden die einzelnen Libraries auf eine Konzentration von 10 nmol/L in einem Volumen von 10 µl verdünnt und im gleichen Verhältnis von je 0,67 µl pro Probe gemischt. Zusätzlich wurden 2 µl 10 nM PhiX hinzugefügt. Das führte zu einem Verhältnis 5:1 von Library zu PhiX. Das PhiX ist nötig, wie unter III 3.1.5 beschrieben, um für einen Farbbalance bei der Sequenzierung zu sorgen, da es sich bei RRBS-Libraries ebenfalls um „low diversity libraries“ handelt. Nach der Denaturierung der dsDNA mit einer 2 N NaOH Lösung, wurden

die Proben auf eine Endkonzentration von 8 pM eingestellt. Die genauen Angaben für die Denaturierung bzw. Verdünnung und die jeweils benötigten Puffer können im MiSeq System Denature and Dilute Libraries Guide von Illumina eingesehen werden. Für die Herstellung der Reagenzien zur Generierung der Cluster wurde das TruSeq Cluster Generation Kitv5 von Illumina verwendet und nach dem „Reagent Preparation Guide For Single-Read Runs“ verfahren [Int. 11]. Die Bildung der Cluster auf der Flusszelle erfolgte mit der Cluster Station von Illumina. Für die anschließende Sequenzierung wurde das TruSeq SBS Kitv5 Kit von Illumina eingesetzt und nach dem „Reagent Preparation Guide“ für den Genome Analyzer II verfahren [Int. 12]. Bei dieser Art des Sequenzierlaufs handelte es sich um das single-end sequencing mit 72 Zyklen zur Sequenzierung der einzelnen Fragmente und einer anschließenden Indexsequenzierung von sechs Zyklen. Beim single-end sequencing wird im Vergleich zum paired-end sequencing ausschließlich der forward Strang der Fragmente sequenziert. Nachdem die Cluster generiert wurden, wird die DNA denaturiert und die, durch die Amplifikation der Fragmente entstandenen, reversen Stränge abgespalten und von der Flusszelle gespült. Beim paired-end sequencing wird vor der Zugabe von Primern, NTPs und Polymerase ein weiterer Amplifikationsschritt eingefügt. Dadurch liegen neben den forward Strängen auch die reversen Stränge für die Sequenzierung vor.

4. Datenanalyse und Statistik

4.1 Auswertung der NGS-Daten

4.1.1 DBAS

Die bioinformatische Analyse wurde mit Hilfe von Frau Dr. Silke Appenzeller (Institut für Biochemie und Molekularbiologie, Universität Würzburg) durchgeführt. Nach dem Demultiplexen erfolgte eine erste Beurteilung der Read-Qualität mit dem Programm FastQC, v0.11.2 [Int. 13] Im nächsten Schritt wurden Reads von unzureichender Qualität und Adaptersequenzen mit dem Programm TrimGalore, v0.4.0 [Int. 14] und Cutadapt, v1.6 [Int. 15] aus dem Datensatz eliminiert [235]. Das anschließende „Joining“ der paired-end Reads wurde mit der fastq-join Option des Programms ea-utils, v1.1.2-537 [Int. 16] durchgeführt. Die Reads wurden dann gegen die Sequenzen des humanen Referenzgenoms hg19 UCSC unter Zuhilfenahme von Bismark, v0.14.3 [Int. 17] und Bowtie, v2.2.6 [Int. 18] aligniert [236, 237]. Für die weitere Bearbeitung der alignierten Reads wurde das Programm SAMtools, v1 verwendet [Int. 19] [238]. Reads, welche eine unvollständige Bisulfitkonvertierung

aufwiesen, wurden an dieser Stelle entfernt. Die Berechnung der Methylierungswerte, das sogenannte „methylation calling“, wurde mit dem `bismark_methylation_extractor` ausgeführt. Für die Visualisierung und die weitere Analysen, zur Detektion von aberrant methylierten Allelen, Einzel-CpG-Fehlern und Sequenzveränderungen, wurde das Programm `GensearchNGS` (Tabelle 18) eingesetzt. Reads, die weniger als 5 CpG-Stellen abdeckten und/oder kürzer als 100 bp waren, wurden von der Analyse ausgeschlossen.

4.1.2 RRBS

4.1.2.1 Datenanalyse

Die bioinformatische Analyse der RRBS-Daten wurde in Zusammenarbeit mit Christian Remmele (Institut für Bioinformatik, Universität Würzburg) angefertigt. Zu Beginn wurden die Adaptersequenzen und Reads mit einer niedrigen Qualität aus dem Datensatz mittels dem Tool `TrimGalore`, v0.4.0 [Int. 14] von `Cutadapt`, v1.6 [Int. 15] und `FastQC`, v0.11.3 [Int. 13] aus dem Datensatz eliminiert [235, 236]. Als Referenz für das sogenannte „Mapping“ der Reads wurde für die menschlichen Reads das humane Genom hg19 [Int. 20] und für den Schimpansen das Schimpansengenom panTro4 [Int. 21] verwendet. Die sogenannten „curated reads“ wurden mittels `Bismark`, v0.14.5 unter Verwendung der Standardparameter kartiert [236]. In die Berechnung der CpG-Methylierung flossen ausschließlich die sogenannten „uniquely mapped reads“ ein. Zur Vereinfachung der weiteren Analysen wurden die Positionen der CpG-Stellen des Schimpansengenoms via `liftover` [Int. 22, Int. 23] auf das humane Genom transferiert [239]. Für die Korrespondenzanalyse wurden die 10.000 variabelsten Methylierungsloci, über alle Libraries hinweg betrachtet, über die `cca` Funktion [240] des R Pakets `vegan` [Int. 24] extrahiert [241]. Der Plot wurde basierend auf den beiden Achsen mit dem höchsten Informationsgehalt erstellt. Das `DSS` Paket wurde für das Testen der differentiellen Methylierung auf Einzel-Basen-Level verwendet [242, 243]. Für jede der vier DMR (differenziell methylierten Region) Kontraste (intraspezifisch Mensch/Schimpanse, interspezifisch neuronal/nicht-neuronal) wurden nur CpGs berücksichtigt, die von mindestens drei Reads in mindestens zwei von drei Proben je Gruppe abgedeckt wurden. Für die Kalkulation der durchschnittlichen Methylierungswerte und der Methylierung je CpG-Stelle, sowie der Durchführung des Wald-Tests, wurde der function `DML Test` angewandt. Die kalkulierten p-Werte wurden auf multiples Testen hin korrigiert [244]. Die Ermittlung der DMRs basierte auf den `DMLtest` Ergebnissen; diese wurden durch das sogenannte „calling“

der callDMR Funktion innerhalb des DSS Pakets berechnet [242, 243]. Als DMRs wurden Regionen nur dann klassifiziert, wenn sie eine Mindestlänge von 100 bp aufwiesen, mindestens zwei CpG-Stellen beinhalteten und mindestens 50% der CpG-Stellen eine statistisch signifikant unterschiedliche Methylierung zeigten. Benachbarte DMRs, die weniger als 100 bp voneinander entfernt lagen, wurden zu einer DMR zusammengefasst. Die DMRs wurden gemäß dem DSS Paket gerankt.

Die Annotation der Gene wurde dem humanen Genom hg19 via BioMart entnommen [245]. Der Promotorbereich wurde als 1,5 kb upstream und 500 bp downstream des Transkriptionsstartpunkts definiert. Mittels des R-Pakets GenomicRanges wurden die overlapping, preceding und following Gen- und Promotor-Annotationen durchgeführt [246].

4.1.2.2 Enrichmentanalysen

Das Hintergrundmodell für die Methylierungsloci wurde über die Extraktion aller CpG-Stellen erstellt, die im humanen bzw. im Schimpansengenom enthalten sind. Das DMR-Hintergrundmodell zur Berechnung des Enrichments innerhalb jedes Kontrasts umfasste alle potentiellen DMRs, basierend auf den getesteten Methylierungsloci. Jeder Hintergrund wurde mit den gleichen DSS Parametern wie auch die berechneten DMRs ermittelt. Schwankungen in der genomischen Verteilung der Methylierungsloci, welche beispielsweise durch die CpG-Verteilung, technische Schwankungen durch den MspI-Verdau oder verwendete Parametereinstellungen hervorgerufen werden können, wurden ebenfalls berücksichtigt.

Zur Errechnung des Enrichments wurden die Koordinaten der Zentromer-, Telomer- und Subtelomerregionen auf dem humanen und Schimpansengenom extrahiert [247-249]. Die Anzahl der gemessenen Methylierungsloci und der Methylierungsloci des Hintergrundmodells wurde via Fisher-Test berechnet, um eine potentielle An- oder Abreicherung der gemessenen Methylierungsloci in Zentromer-, Telomer- oder Subtelomerregionen zu identifizieren [250]. Für die regionale Enrichmentanalyse der DMRs, wurde die Überlappung der DMRs mit drei verschiedenen Regionen für den intra- bzw. interspezifischen Kontrast kalkuliert. Der Fisher-Test wurde mit dem DMR-Hintergrundmodell durchgeführt. P-Werte $<0,05$ werden als signifikant erachtet. Eine Odds ratio <1 bzw. >1 weist auf eine Anreicherung bzw. Abreicherung der DMRs hin. Die Analyse

des Enrichments von Genen, welche mit neuropsychiatrischen Erkrankungen oder humanspezifischer Gehirnregulation assoziiert sind, erfolgte mit den DMR-überlappenden Genen und Promotoren (1,5 kb upstream) der vier Kontraste. Zudem wurden die Gene und Promotoren (1,5 kb upstream) extrahiert, welche mit dem Hintergrundmodell übereinstimmen. Diese beiden Gensets wurden mit Genlisten abgeglichen, die mit humanspezifischer Gehirnexression [251-254], humanspezifischen Histonmethylierungssignaturen [255] oder mit neuropsychiatrischen Erkrankungen assoziiert werden (138 Gene für Mentale Retardierung (Anhang-Tabelle 3), 203 Autismogene (Autism database AutDB, 2017) [Int. 25] [256], 408 Gene für Bipolare Störungen (Genetic Database for Bipolar Disorder BDgene, 2016) [Int. 26], 351 Schizophreniegene [257], 129 Alzheimergene (AlzGene database, 2017) [Int. 27] [258]). Die signifikante An- bzw. Abreicherung der übereinstimmenden Gene wurde mittels Fisher-Test berechnet. Nach der Korrektur für multiples Testen wurden p-Werte $<0,05$ als signifikant erachtet und Odds ratios <1 bzw. >1 als angereichert bzw. abgereichert bewertet. Für die Erstellung der gene ontology (GO) Analyse (mittels des R-Pakets RDAVIDWebService) wurden die mit den DMRs und dem Hintergrundmodell übereinstimmenden Gene/Promotoren verwendet [259]. Die Berechnung erfolgte mit den GO-Annotationen GOTERM_BP_FAT, GOTERM_CC_FAT und GOTERM_MF_FAT unter Einsatz der gemessenen DMRs als Vordergrundmodell und der Gene des Hintergrundmodells als Hintergrund.

4.2 Statistische Analyse

Das IMB SPSS Programm, v23 wurde sowohl für die deskriptiven als auch für die bivariaten statistischen Berechnungen genutzt. Bei den vergleichenden Gruppenanalysen wurde zuerst ein Test auf Normalverteilung durchgeführt. Zeigten die Werte innerhalb einer Gruppe eine Normalverteilung, wurde der parametrische (T-Test) bzw. für nicht normalverteilten Gruppen ein nicht-parametrischer Test (Mann-Whitney U Test) angewandt. Für die Korrelationsanalysen basiert die Art der Berechnung ebenfalls auf der Verteilung. Bei normalverteilten Werten erfolgte die Analyse über Pearson, bei nicht normalverteilten Werten über Spearman. Um eine signifikante Beziehung zwischen zwei kategorialen Variablen zu testen, kam der Chi-Quadrat Unabhängigkeitstest zum Einsatz. P-Werte $<0,05$ wurden als signifikant angesehen.

5. Datenvalidierung via Pyrosequenzierung

5.1 Genspezifische PCR

Das Designen der PCR- und Sequenzier-Primer erfolgte mit der PyroMark Assay Design 2.0 Software von Qiagen. Bei der Pyrosequenzierung wird entweder der forward oder reverse Strang sequenziert. Demzufolge wird an den Primer des zu sequenzierenden Stranges das 5'-Ende mit einem zusätzlichen Biotin versehen. Über diese Modifikation wird der gewünschte Strang mittels Waschstation selektiert. Für die Etablierung der Primer wurde generell eine Gradienten-PCR angefertigt, um die optimale Primer Annealing-Temperatur und die erforderliche Anzahl an PCR-Zyklen ermitteln zu können. Hierfür wurde ein Viertel (12,5 µl) des unter III 3.1.1 aufgeführten PCR-Ansatzes verwendet und ein Gradient von 50 – 60 °C für die Annealing-Temperatur gesetzt. Es wurden die unter III 3.1.1 aufgeführten PCR-Bedingungen verwendet. Die genspezifischen Annealing-Temperaturen sind in Tabelle 11 aufgeführt.

5.2 Pyrosequenzierung

Bei der Pyrosequenzierung handelt es sich um eine Methode, die der Sequenzierung von DNA-Fragmenten dient; sie basiert ebenfalls auf der SBS-Technologie.

In einer Kaskade von enzymatischen Reaktionen wird innerhalb der Polymerisation der DNA sichtbares Licht generiert, welches proportional zur Anzahl an eingebauten Nukleotiden steht.

Zu Beginn der Sequenzierung erfolgt die Hybridisierung des Sequenzier-Primers an die ssDNA. Es wird sowohl ein Enzym-Mix, bestehend aus dem Klenow Fragment der DNA Polymerase I, ATP-Sulfurylase, Luciferase und Apyrase, als auch ein Substrat-Mix zugegeben, bestehend aus Adenosinphosphosulfat (APS) und Luciferin. Erfolgt der Einbau eines dNTPs an den DNA-Strang über das Klenow Fragment, wird Pyrophosphat (PP_i) abgespalten. Dieses reagiert mit APS über die ATP-Sulfurylase zu Adenosintriphosphat (ATP). Das ATP wiederum katalysiert die Umwandlung von Luciferin zu Oxyluciferin durch die Luciferase. Die Intensität des emittierten Lichtsignals wird zeitgleich von einem Sensor gemessen und steht in direkter Abhängigkeit zur Menge an generiertem ATP und somit zur Anzahl an eingebauten Nukleotiden (Abbildung 24). Pro Sequenzierzyklus wird nur eine Sorte dNTPs zugegeben. Nach erfolgter Detektion des Signals werden die nicht eingebauten Nukleotide und das

überschüssige ATP mittels Apyrase abgebaut; eine neue Runde erfolgt nach dem gleichen Prinzip.

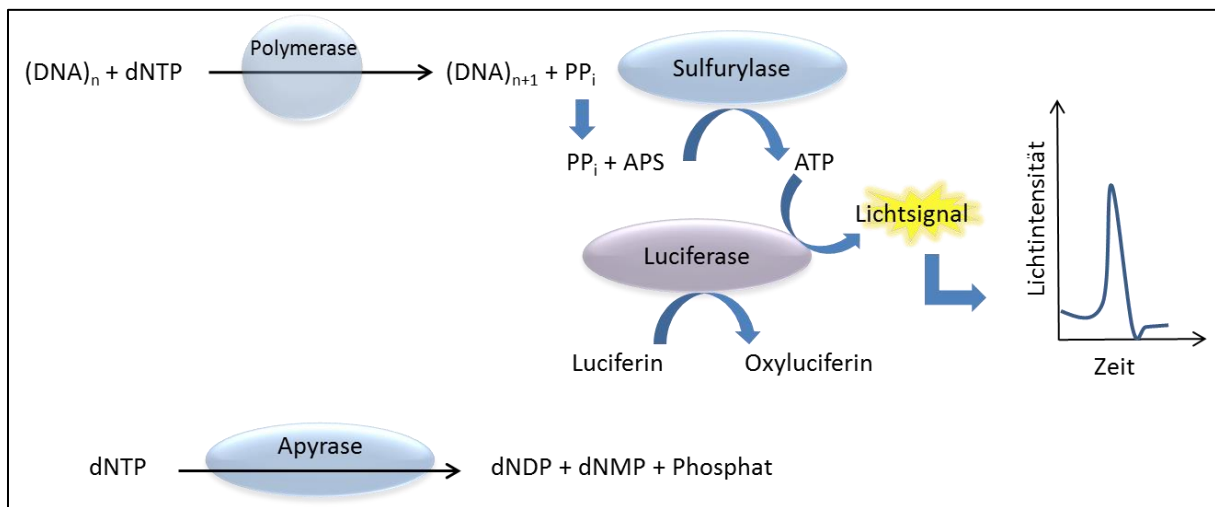


Abbildung 24: Übersicht über die enzymatische Reaktionskaskade bei der Pyrosequenzierung. Durch den Einbau eines dNTPs kommt es zur Abspaltung des Pyrophosphats (PP_i); dieses reagiert mit Adenosinphosphosulfat (APS) über die Sulfurylase zu Adenosintriphosphat (ATP). Durch das ATP wird die Umwandlung von Luciferin zu Oxyluciferin katalysiert. Das emittierte Lichtsignal wird detektiert. Die Apyrase sorgt für den Abbau der nicht-eingebauten dNTPs. (Abb. erstellt nach dem Vorbild von [Int. 28])

Zu den Amplifikaten wurden jeweils 2 μ l Streptavidin Sepharose HP beads (Qiagen, Hilden, Deutschland), 40 μ l Binding-Puffer (Qiagen, Hilden, Deutschland) und 28 μ l H_2O hinzugegeben und für 5 min geschüttelt. Dieser Schritt wird auf einer 96-well Platte durchgeführt. Parallel wurden auf einer separaten Sequenzierplatte, pro Well 11,5 μ l Annealing-Puffer (Qiagen, Hilden, Deutschland) und 0,5 μ l des für die jeweilige Probe korrespondierenden Sequenzierprimers zugegeben. Die Aufbereitung der Proben erfolgte mit der PyroMark Q96 Workstation (Qiagen, Hilden, Deutschland). Die DNA bindet über die Biotin-Modifikation an die Sepharose Beads. Diese werden dann mittels Vakuum an den Waschkopf angesaugt. Zuerst erfolgt ein Waschschrift in 75% EtOH, dann wurde der Kopf in den Denaturierungs-Puffer gehalten, um die ssDNA Fragmente zu erhalten. Abschließend erfolgte ein Waschschrift. Es ist wichtig zu beachten, dass vor jedem Eintauchen in eine neue Lösung jegliche Flüssigkeit aus dem Waschkopf vollständig entfernt wird, vor allem der EtOH sollte möglichst vollständig abgefließen sein, da es andernfalls zu möglichen Störungen der enzymatischen Reaktionen beim Sequenzieren kommen kann. Sind alle Schritte an der

Waschstation erfolgt, wird der Kopf möglichst nah über die Sequenzierplatte und erst nach Abschalten des Vakuums direkt in die Platte gehalten. Anschließend wurde die Sequenzierplatte für 2 min auf den Heizblock bei 80°C gestellt. Während dieser Schritt erfolgt, können die Küvetten befüllt werden. Die jeweilige Menge an Enzym und Substrat wird von der Software vorgegeben; diese wird anhand der eingespeicherten Assays für die jeweiligen Gene berechnet. Die Platte sollte vor der Sequenzierung etwas abkühlen; währenddessen kann der Dispensations-Test durchgeführt werden.

6. Genexpression

Die Genexpressionsanalyse erfolgte nach dem Prinzip der Real-Time TaqMan Technologie. Diese ist im Vergleich zur SYBR Green-basierten Real-Time PCR wesentlich sensitiver und spezifischer. Bei der TaqMan Technologie interkaliert der Fluoreszenzfarbstoff, anders als bei der SYBR Green Methode, nicht in die dsDNA, sondern wird direkt bei der Amplifikation eingebaut. Die TaqMan Sonden sind am 5'-Ende mit einem Reporterfarbstoff und am 3'-Ende mit einem sogenannten Quencher-molekül versehen. Mittels des Fluoreszenz-Resonanz-Energietransfers (FRET) unterdrückt der Quencherfarbstoff das Fluoreszenzsignal des Reportermoleküls. Solange die Sonde nicht an das cDNA-Template hydrolysiert, befinden sich Reporter und Quencher in räumlicher Nähe und das Fluoreszenzsignal bleibt „stillgelegt“ (Abbildung 25, 1). Während der Annealingphase lagert sich die Sonde an die komplementäre Sequenz an, welche zwischen dem forward und dem reverse Primer lokalisiert ist (Abbildung 25, 2). Durch die 5'→3' Exonukleaseaktivität der Taq-Polymerase kommt es zur Aufspaltung der Sonde und somit zur Trennung zwischen Reporter- und Quencher-molekül. Dadurch wird das Fluoreszenzsignal aktiviert (Abbildung 25, 3). Dieser Vorgang wiederholt sich in jedem Zyklus und die Fluoreszenzintensität steigt in direkter Abhängigkeit zur Zielsequenzkonzentration.

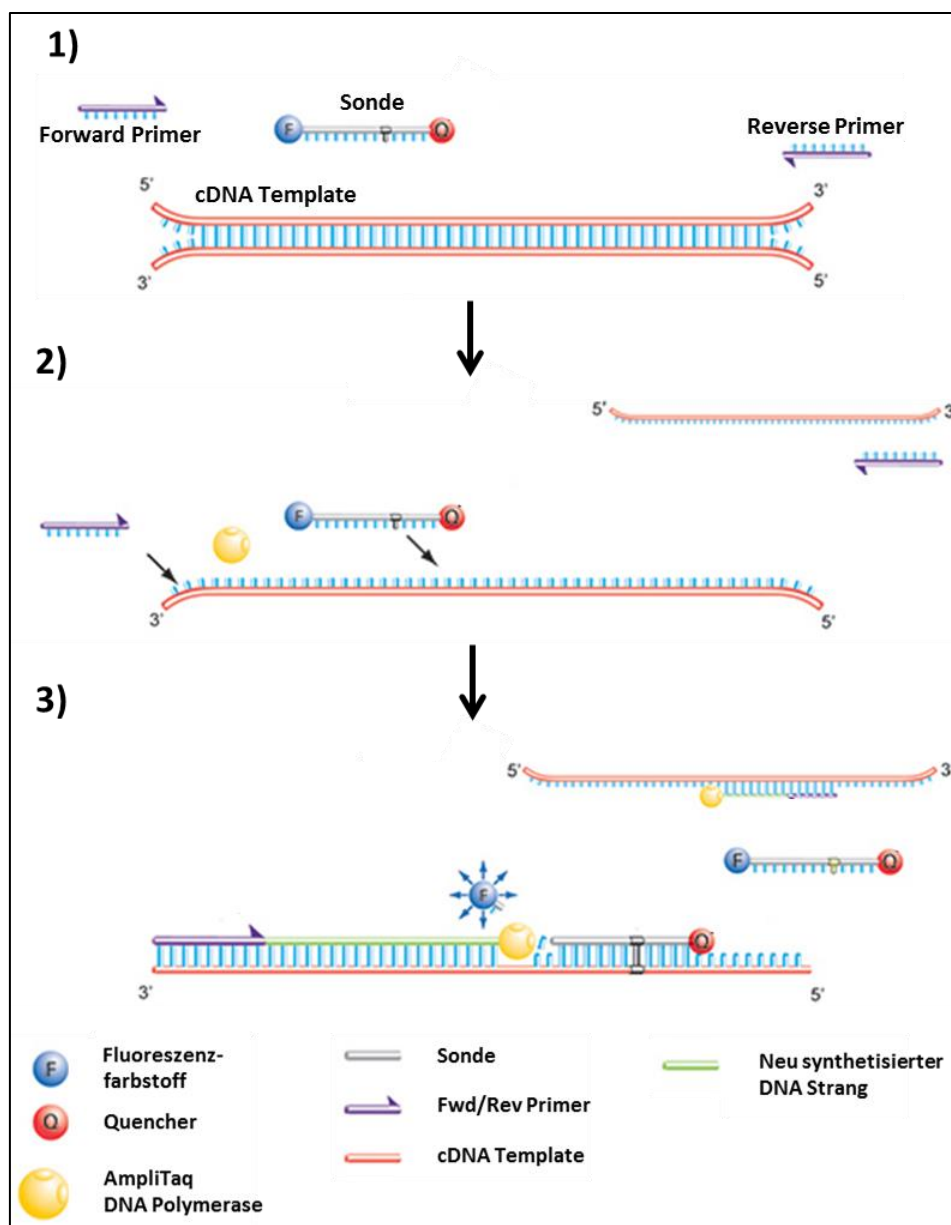


Abbildung 25: Prinzip der Real-Time TaqMan Technologie. **1)** Reaktionsgemisch bestehend aus forward (Fwd) und reverse (Rev) Primern, sowie cDNA Template und Sonde, welche mit einem Quencher- (Q) und Fluoreszenzfarbstoffmolekül (F) versehen ist. **2)** Annealing der Sonde und Primer jeweils komplementär zur Zielsequenz. **3)** Elongationsphase durch die Taq-Polymerase. Es kommt zur Spaltung der Sonde und der Aktivierung des Fluoreszenzsignals (Abb. entnommen von [Int. 29] und modifiziert).

Bei der Genexpressionsanalyse wurden die in Tabelle 3 aufgelisteten Patientenproben untersucht. Die RNA-Extraktion erfolgte wie unter III 1.2 beschrieben. Die cDNA wurde mit dem High Capacity DNA-to-cDNA Kit (ThermoFisher Scientific, Waltham, USA) unter dem Einsatz von 1 µg RNA synthetisiert. Der Reaktionsansatz und die Reaktionsbedingungen wurden dem beiliegenden Handbuch entnommen. Die TaqMan Real-Time PCR wurde mit

dem TaqMan Fast Advanced Master Mix (ThermoFisher Scientific, Waltham, USA), 50 ng cDNA pro Reaktion und den kommerziell erwerblichen TaqMan Assays von ThermoScientific (Tabelle 10) nach den Angaben des Herstellers durchgeführt. Das ABI Viiia7 System (ThermoFisher Scientific, Waltham, USA) wurde für den Lauf der RT¹ PCR genutzt. Jede Probe und negative Kontrolle wurde in technischen Triplikaten analysiert. *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) und *ACTB* (actin beta) wurden als Referenzgene für die Normalisierung verwendet. Die Evaluierung der Schmelz- und Amplifikationskurven wurde unter Zuhilfenahme der QuantStudio Real-Time PCR Software v1.2.4 (ThermoFisher Scientific, Waltham, USA) nach der $\Delta\Delta C_t$ Methode durchgeführt.

IV ERGEBNISSE

1. Genomweite Methylierungsanalysen im präfrontalen Cortex via RRBS

Für die Erstellung eines Methylierungsprofils neuronaler und nicht-neuronaler Zellen im präfrontalen Cortex des Menschen und Schimpansen, wurde eine genomweite differenzielle Methylierungsanalyse durchgeführt. Es wurden je Spezies drei Individuen untersucht, von denen präfrontales Cortex Gewebe (Brodmann-Areal 10, BA10) nach dem Tod entnommen wurde. Anschließend wurde das Gesamtgewebe in neuronale (N+) und nicht-neuronale (N-) Zellen aufgetrennt. Die daraus gewonnene DNA wurde dann via Reduced Representation Bisulfite Sequencing (RRBS) präpariert und abschließend mit dem Genome Analyzer II (GAII) sequenziert. Nach der bioinformatischen Auswertung wurden ausgewählte Gene mittels Pyrosequenzierung überprüft, um die Methylierungswerte des RRBS-Experiments zu validieren.

1.1 Analyse der RRBS-Daten

Der gesamte RRBS-Datensatz mit 12 Libraries umfasste rund 149,5 Millionen (Mio.) sequenzierter Reads (5,9-19,6 Mio. Reads pro Library). Nach Ausschluss der sogenannten „uncrated“ (Reads mit einer Länge von weniger als 16 bp), „unmapped“ und „multi mapped“ Reads verblieben 90,4 Mio. sogenannte „uniquely mapped“ Reads (48,4% im menschlichen und 78,1% im Schimpansen Genom) für die anschließende Methylierungsanalyse (Abbildung 26, links). Die „uniquely mapped“ Reads der menschlichen Libraries bzw. des Schimpansen deckten zwischen 2,2 und 2,9 Mio. bzw. 2,0 und 2,4 Mio. Methylierungsloci ab. Nach dem Ausschluss der Loci, die von weniger als drei Reads in mindestens zwei Libraries pro Gruppe abgedeckt wurden, flossen 1,3 Mio. (49,1%) menschliche und 1,1 Mio. (49,5%) Schimpansen-Methylierungsloci in die intra- bzw. interspezifische Analyse zur Untersuchung der differentiellen Methylierung ein (Abbildung 26, rechts).

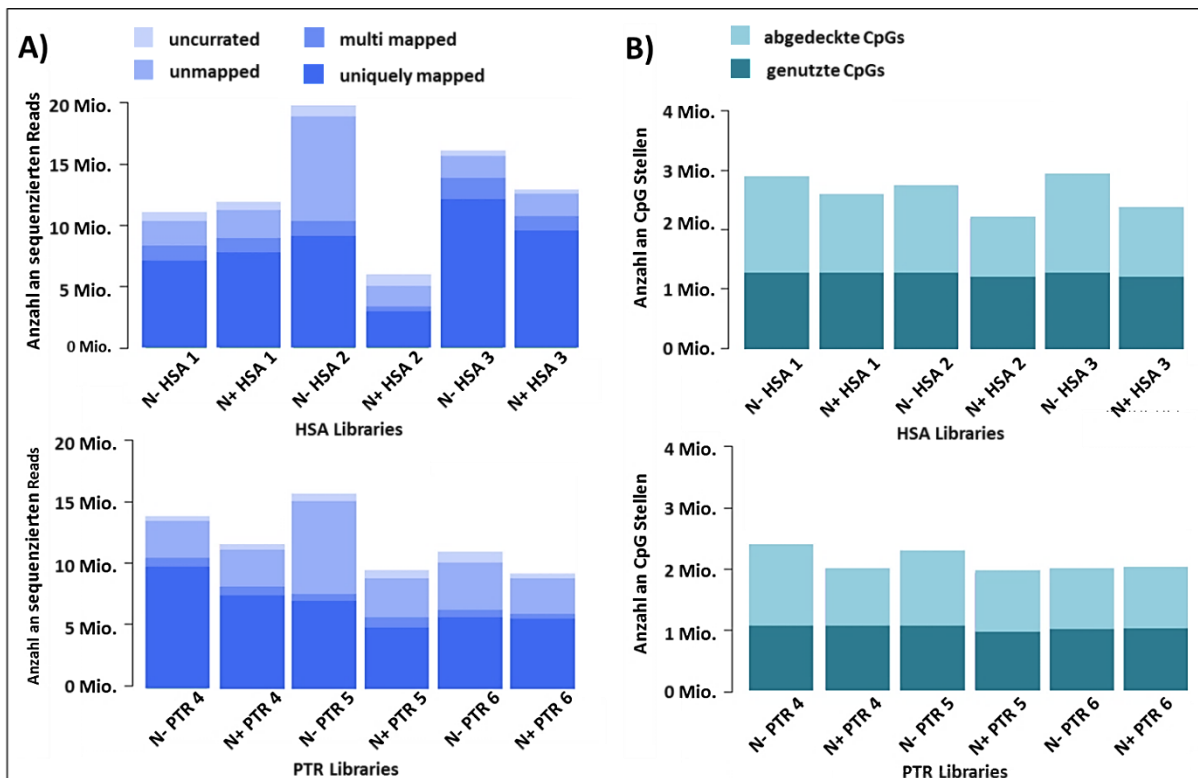


Abbildung 26: Anzahl an sequenzierten Reads und abgedeckten CpG-Stellen je Library. **A)** Das obere Diagramm zeigt die Anzahl an sequenzierten Reads für alle menschlichen (HSA) neuronalen (N+) und nicht-neuronalen (N-) Libraries. Die untere Grafik repräsentiert die sequenzierten Schimpansen (PTR) Libraries. Nur die „uniquely mapped“ Reads sind in die nachfolgende Analyse eingeflossen. Der Farbcode kann der Legende entnommen werden. **B)** Die helltürkisfarbenen Balken der oberen Grafik stellen die Anzahl der CpG-Stelle, die über die „uniquely mapped“ Reads abgedeckt werden, je HSA Library der N+ und N-, dar. Die dunkeltürkisfarbenen Balken geben die Anzahl an CpGs an, welche durch mindestens drei Reads in mindestens zwei Libraries je Zelltyp und Spezies abgedeckt werden. Diese Fraktionen wurden in die anschließende differenzielle Methylierungsanalyse eingesetzt. Das untere Diagramm zeigt die CpG-Stellen für den PTR.

Zur Überprüfung der Qualität der Daten wurde eine Korrespondenzanalyse basierend auf den 10.000 variabelsten Methylierungsloci angefertigt. Der höchste Methylierungsunterschied kristallisierte sich zwischen den Zelltypen (Achse 1, 21%) heraus, gefolgt vom Speziesunterschied (Achse 2, 11%). Es konnte eine deutliche Clusterbildung der biologischen Replikate beobachtet werden (Abbildung 27).

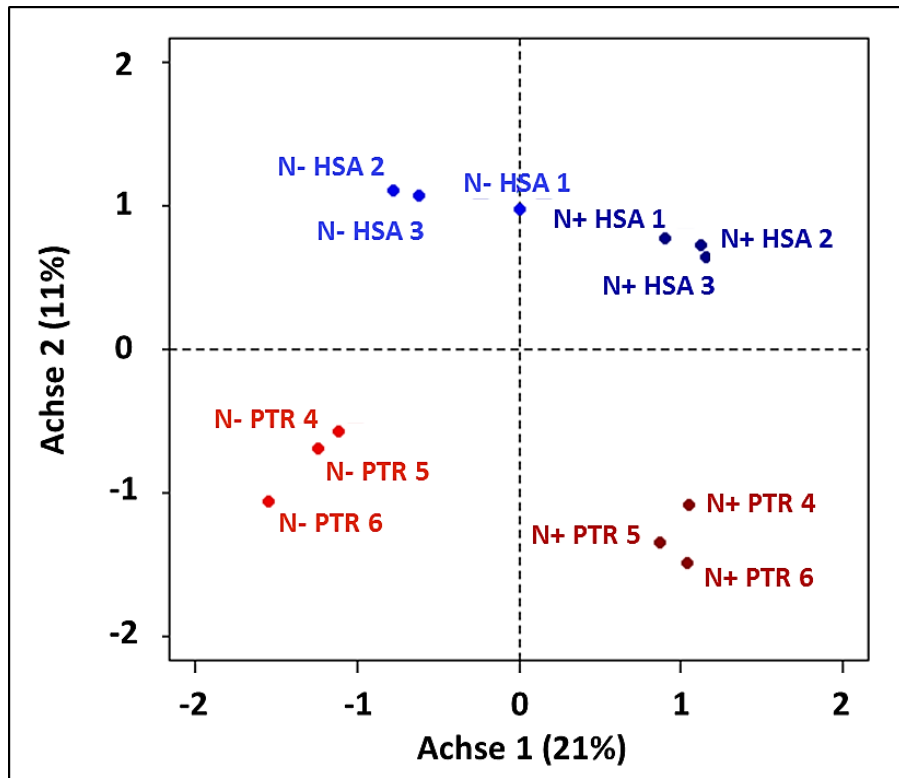


Abbildung 27: Korrespondenzanalyse basierend auf den 10.000 variabelsten Methylierungsloci. Die einzelnen Proben werden durch die Punkte repräsentiert. Die Distanz zwischen den Punkten entspricht einer Qui-Quadrat-Distanz (χ^2). Auf der X-Achse ist der Methylierungsunterschied zwischen den neuronalen (N+) und nicht-neuronalen (N-) Zellen und auf der Y-Achse der Unterschied zwischen Mensch (HSA) und Schimpanse (PTR) dargestellt.

Der Hauptteil der getesteten Methylierungsloci (N+ und N- je Spezies kombiniert) liegt auf Genomebene (Abbildung 28, A) innerhalb der CpG-Inseln (71% im Mensch, 11% im Schimpansen), gefolgt von „Open sea“ (16% im Mensch, 12% im Schimpansen) und „CpG shores“ (11% im Mensch und Schimpansen). Auf Genebene (Abbildung 28, B) sind die meisten Methylierungsloci im Genkörper (42% im Mensch, 32% im Schimpansen) und intergenisch (17% im Mensch, 32% im Schimpansen) lokalisiert. Am zweithäufigsten liegen die CpGs im Promotorbereich (13% im Mensch, 19% im Schimpansen) und 5' UTR (26% im Mensch, 16% im Schimpansen). Die Verteilung der abgedeckten CpG-Stellen verhält sich sowohl im Genom- als auch im Genmodell bei beiden Spezies ähnlich.

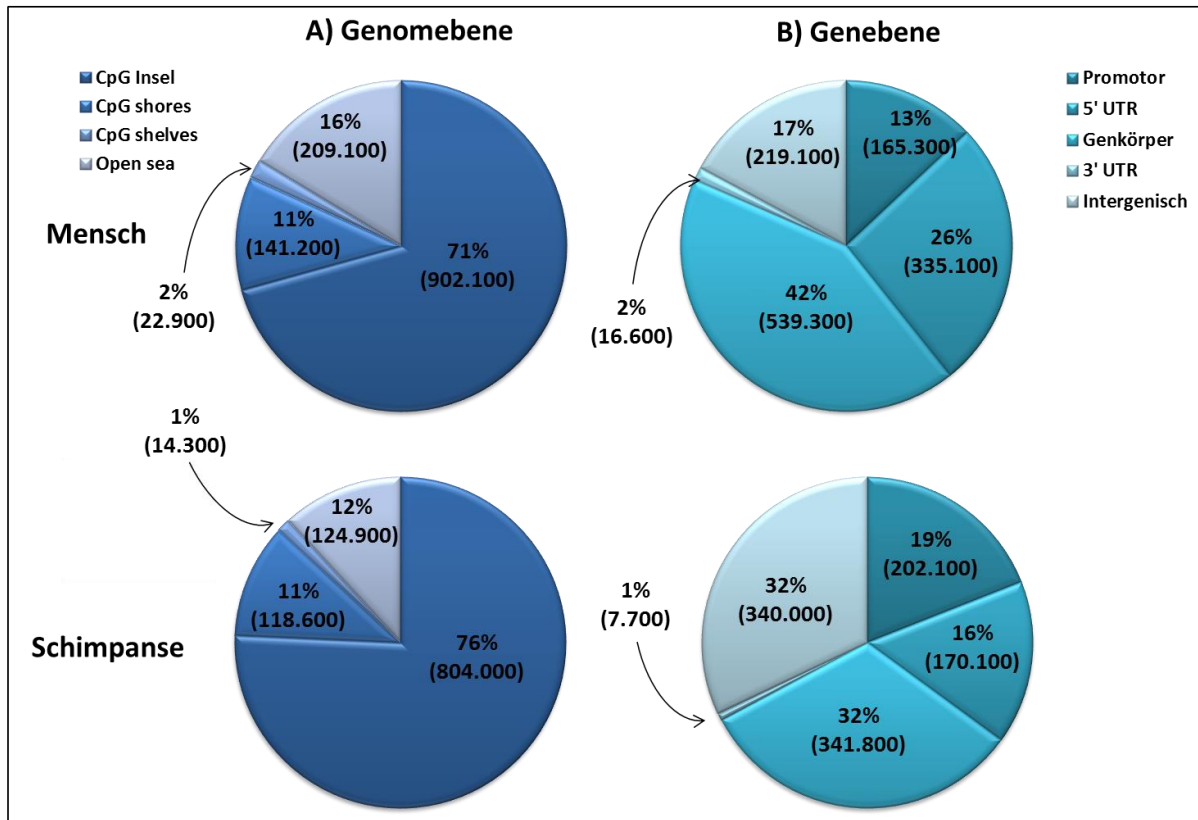


Abbildung 28: Verteilung der abgedeckten CpG-Stellen auf Genom- und Genebene. Abbildung A) zeigt die Verteilung in einem Genomregionenmodell und Abbildung B) in einem Genregionenmodell jeweils für den Menschen (oben) und den Schimpansen (unten). Je Spezies wurden die N+ und N- Proben kombiniert. Beide Spezies zeigen eine ähnliche Verteilung der abgedeckten CpG-Stellen für beide Modelle. Neben den Prozentzahlen ist in Klammern die jeweilige Anzahl an CpG-Stellen angegeben.

Der Test auf differenzielle Methylierung im intraspezifischen Vergleich (N+ vs. N-) ergab nach der Korrektur für multiples Testen 34.182 signifikante Loci im Menschen und 20.393 signifikante Loci im Schimpansen. Der interspezifische Vergleich (HSA vs. PTR) ergab 4.778 signifikante Loci in den neuronalen Zellen und 22.264 in den nicht-neuronalen Zellen.

1.1.1 Intra- und interspezifische differenziell methylierte Regionen

Als Ausgangspunkt für alle nachfolgenden Analysen wurde der RRBS-Datensatz auf differenziell methylierte Regionen (DMRs) hin untersucht. Bei den DMRs handelt es sich entweder um Regionen, die innerhalb eines Zelltyps zwischen den Arten (HSA, Mensch; PTR, Schimpanse) oder innerhalb einer Art zwischen den Zelltypen (N+, neuronale Zellen; N-, nicht-neuronale Zellen) unterschiedlich stark methyliert sind. Dieser Einteilung zufolge erhält

man vier unterschiedliche Kontraste: HSA: N+ vs. N-; PTR: N+ vs. N-; N+: HSA vs. PTR; N+: HSA vs. PTR.

Für den intraspezifischen Vergleich wurden dreimal mehr DMRs im Menschen (n=1635) als im Schimpansen (n=496) ermittelt. Im Schimpansen war die Anzahl der hyper- und hypomethylierten neuronalen DMRs stark ausgeglichen. Im Menschen dagegen waren über 90% der DMRs neuronal hypomethyliert und nur ~10% neuronal hypermethyliert (Abbildung 29,links). Der Interspeziesvergleich zwischen Mensch und Schimpanse in neuronalen bzw. nicht-neuronalen Zellen ergab in Summe 762 DMRs. In den nicht-neuronalen Zellen wurden sechsmal mehr DMRs (n=666) detektiert als in den neuronalen Zellen (n=96). Bei den neuronalen Zellen war die Anzahl an human hypermethylierten Regionen (n=56) schwach erhöht im Vergleich zu den hypomethylierten Regionen (n=40). Im Falle der nicht-neuronalen Zellen hingegen ergaben sich 16 Mal mehr hypermethylierte DMRs (n=628) im Menschen verglichen zum Schimpansen (n=38) (Abbildung 29, rechts).

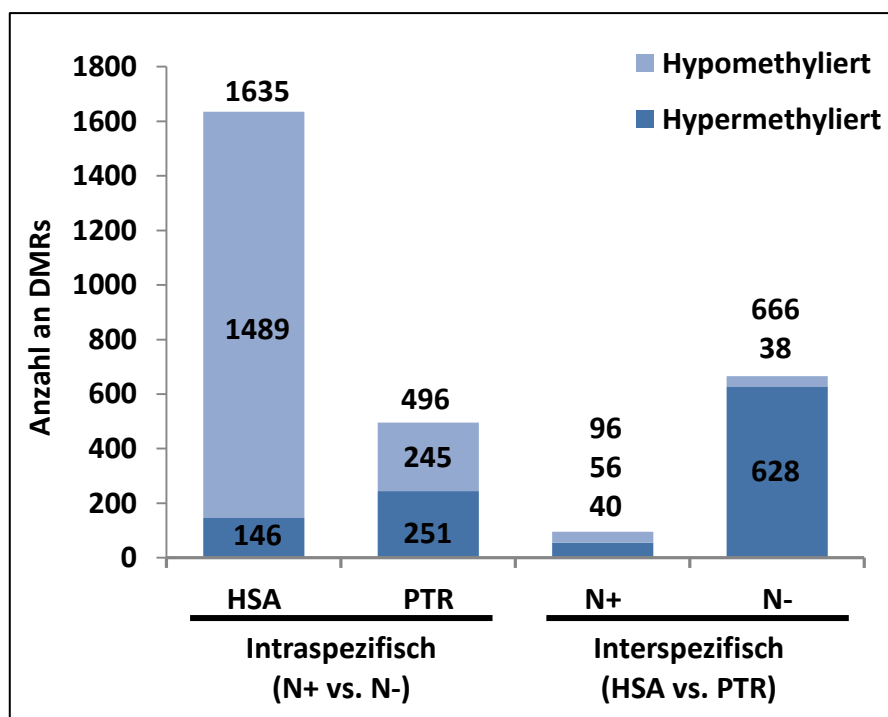
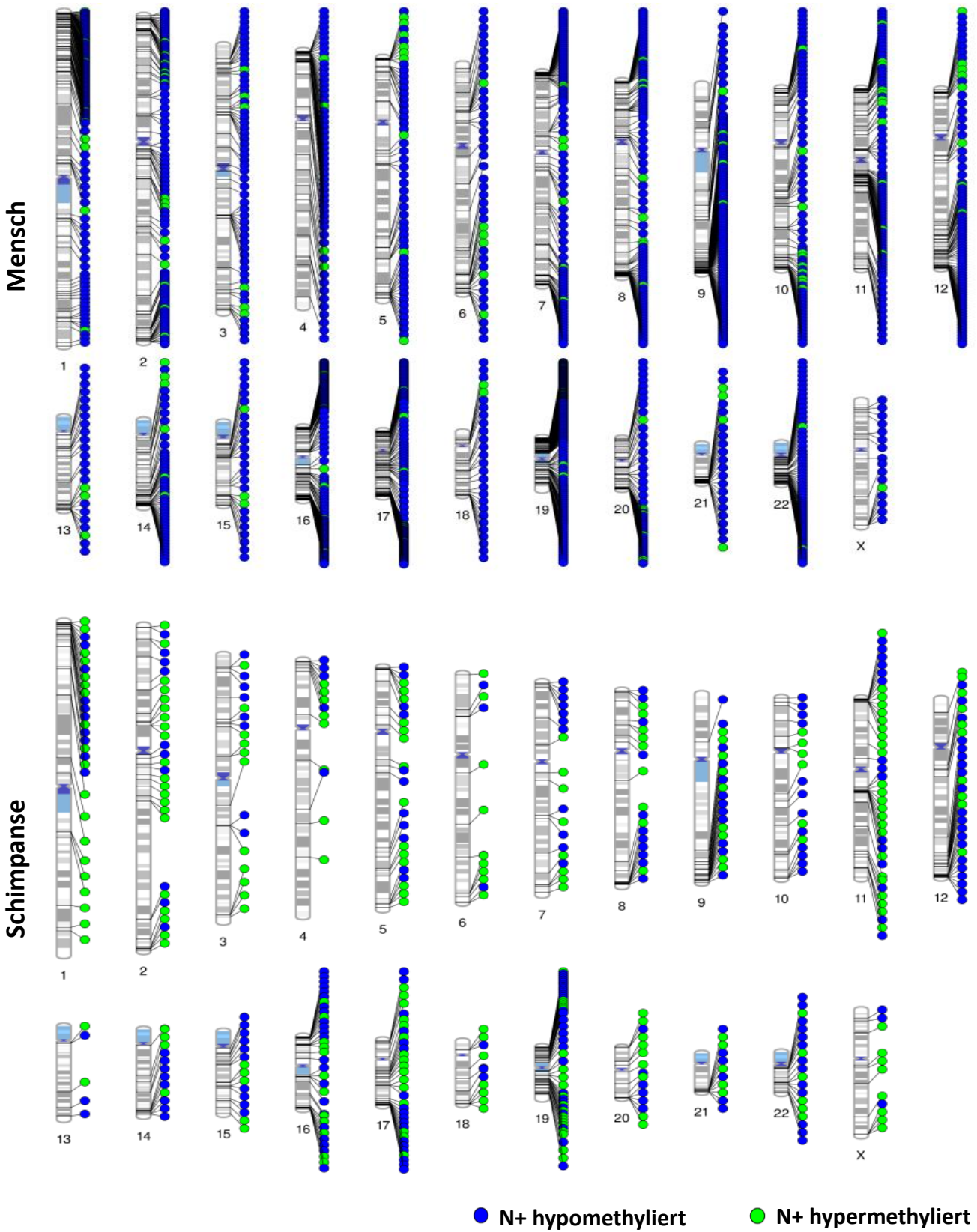


Abbildung 29: Anzahl an Intra- und interspezifischen DMRs. Angegeben ist die jeweilige Anzahl an DMRs je Kontrast sowohl für den Intraspezies- (links) als auch Interspezies-Vergleich (rechts). Die Zahlen geben die Anzahl für die Differenzierung der DMRs in hyper- und hypomethylierte DMRs an. Diese Einteilung bezieht sich im Intraspeziesvergleich auf die neuronalen Zellen und im Interspeziesvergleich auf den Menschen. (HSA, Mensch; PTR, Schimpanse; N+, neuronale Zellen; N-, nicht-neuronale Zellen)

A)

Intraspezifisch
neuronal/nicht-neuronale Zellen



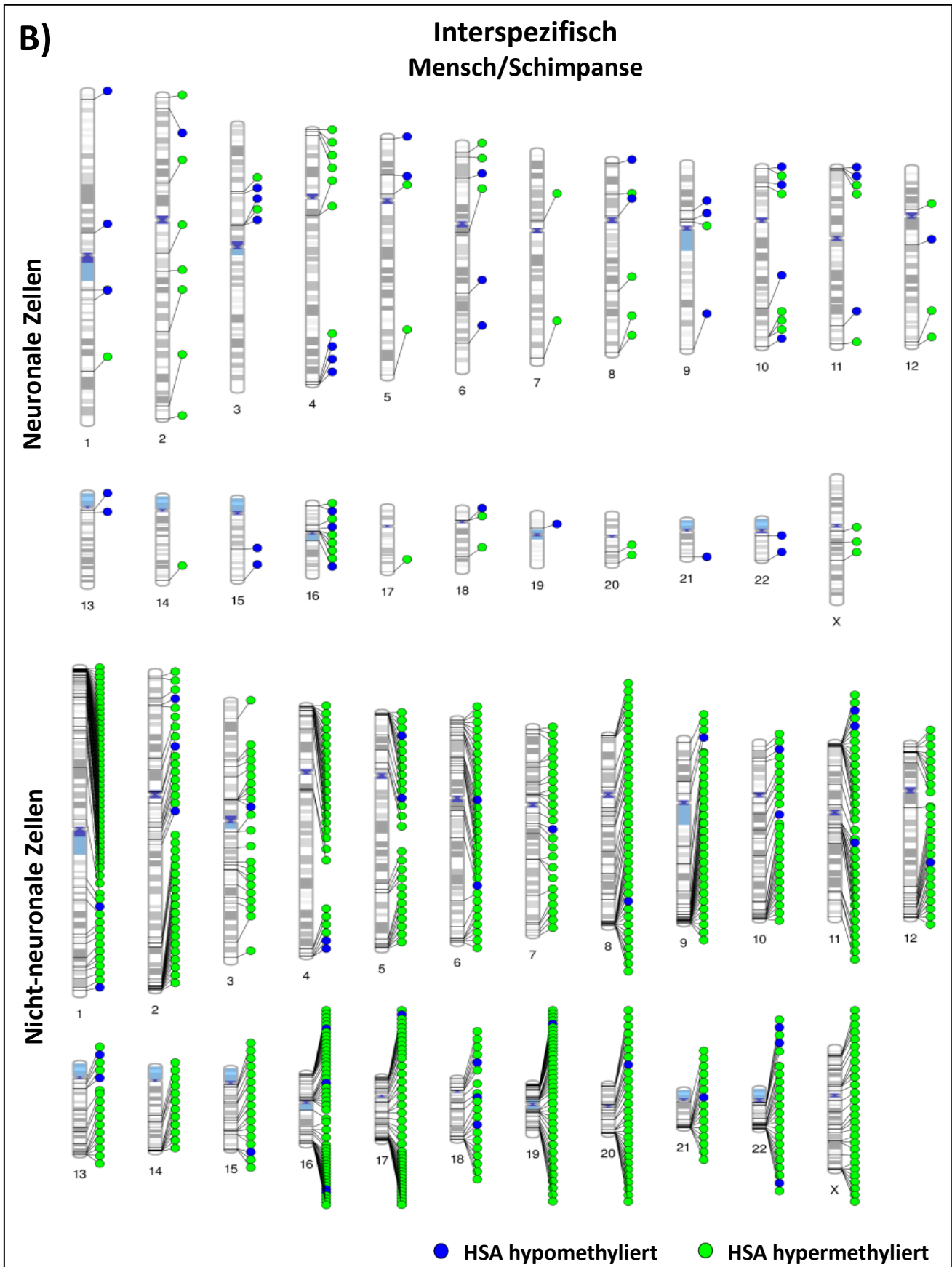


Abbildung 30: Chromosomale Verteilung der intra- und interspezifischen DMRs. Die DMRs sind in allen vier Kontrasten über das ganze Genom hinweg verteilt. Da in dieser Studie nur weibliche Individuen untersucht wurden, ist das Y-Chromosom nicht dargestellt. **A)** In blau sind die neuronal (N+) hypomethylierten und in grün die hypermethylierten DMRs dargestellt. **B)** In Blau sind die human (HSA) hypomethylierten und in Grün die hypermethylierten DMRs abgebildet.

In Abbildung 30 ist die chromosomale Verteilung der intraspezifischen DMRs (Abbildung 30, A) und der interspezifischen DMRs (Abbildung 30, B) dargestellt. Die DMRs sind in allen vier Kontrasten über das ganze Genom verteilt. Da in dieser Studie ausschließlich weibliche Individuen untersucht wurden, ist das Y-Chromosom nicht gezeigt. Durch die Verwendung der RRBS Methode erfolgt eine nicht-zufällige Anreicherung von CpG-reichen Fragmenten, da im ersten Schritt ein Verdau mit MspI durchgeführt wird (Abbildung 31). Dennoch zeigte eine regionale Analyse, unter Berücksichtigung der nicht-zufälligen Verteilung der angereicherten CpG-Stellen, eine signifikante Anreicherung ($p = 3,8E-46$, Faktor 3,0 im Menschen; $p=1,3E-125$, Faktor 4,9 im Schimpansen; Fisher-Test) an differenziell methylierten CpGs in den Subtelomerregionen.

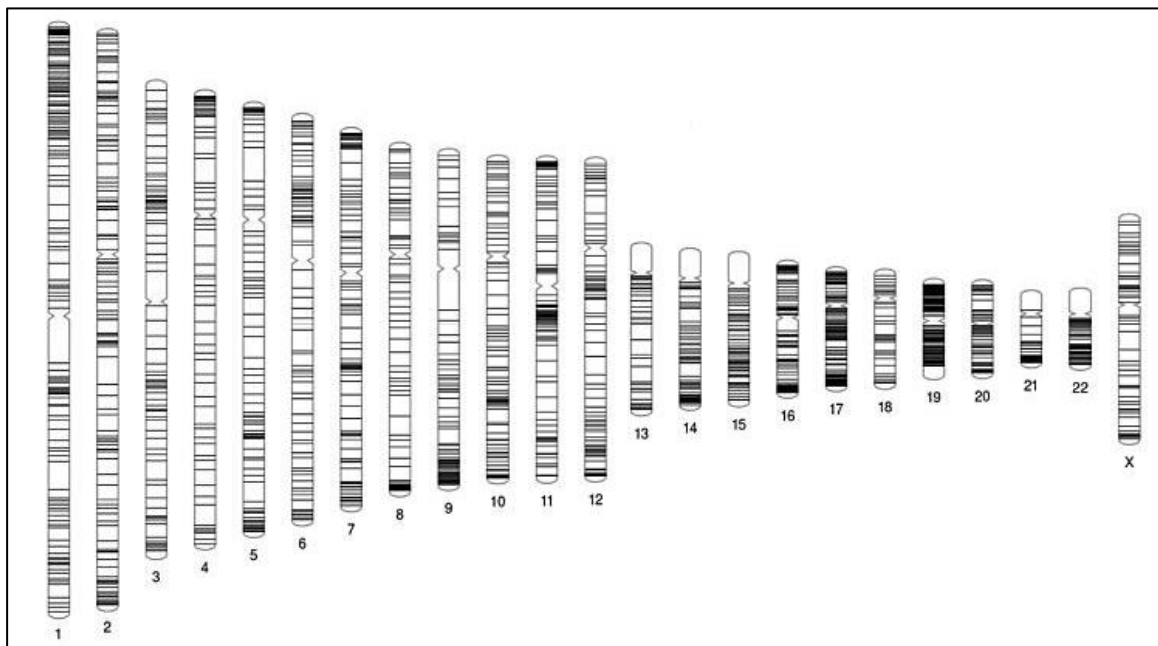


Abbildung 31: Chromosomale Verteilung der mittels RRBS abgedeckten CpG-Stellen. Als schwarze dünne Linie ist nur jede tausendste CpG-Stelle dargestellt, die mittels RRBS abgedeckt wurden. Die dicken schwarzen Linien bzw. Blöcke verweisen auf Regionen im Genom, in denen die sequenzierten CpG-Stellen angereichert vorliegen.

1.1.2 Gene Ontology Term Analysen

Um einen besseren Überblick über die funktionelle und biologische Rolle der Gene zu erlangen, die in den DMRs lokalisiert sind, wurde eine sogenannte „Gene ontology“ (GO) Analyse durchgeführt. Hierbei werden die Genprodukte bestimmten „Terms“ innerhalb der drei Ontologien: „biological process“, „cellular component“ und „molecular function“ zugeordnet. Neben der Klassifizierung der einzelnen Genprodukte wird zudem getestet, ob bestimmte „Terms“ innerhalb eines Datensets besonders häufig abgedeckt werden und somit angereichert sind. Für die GO-Analyse und das Enrichment wurden die Gene und Promotoren der 2.313 intraspezifischen und 762 interspezifischen DMRs verwendet.

Der Intraspeziesvergleich im Menschen (N+ vs. N-) ergab insgesamt 52 signifikant angereicherte (FDR adjusted p-value < 0,05) „Terms“ in der Ontologie „biological process“, 31 „Terms“ in „cellular component“ und 19 „Terms“ in „molecular function“. Die meisten „Terms“ der Ontologie „biological process“ waren assoziiert mit synaptischen Signalen und Neurogenese (Abbildung 32, A). Die intraspezifischen DMRs im Schimpansen wiesen deutlich weniger angereicherte GO „Terms“, die nach der Korrektur für multiples Testen einen signifikanten p-Wert (FDR adjusted p-value) hatten, auf. In der Ontologie „biological process“ ergab die Analyse 7 „Terms“, in „cellular component“ 12 und in „molecular function“ 1 signifikant angereicherten Term (Abbildung 32, B).

Die GO-Analyse ergab keine signifikant angereicherten „Terms“ (FDR adjusted p-value) für die interspezifischen Kontraste (HSA vs. PTR). Eine komplette Auflistung aller signifikant angereicherten „Terms“ der drei Ontologien befindet sich im Anhang (Anhang-Tabelle 1, Anhang-Tabelle 2).

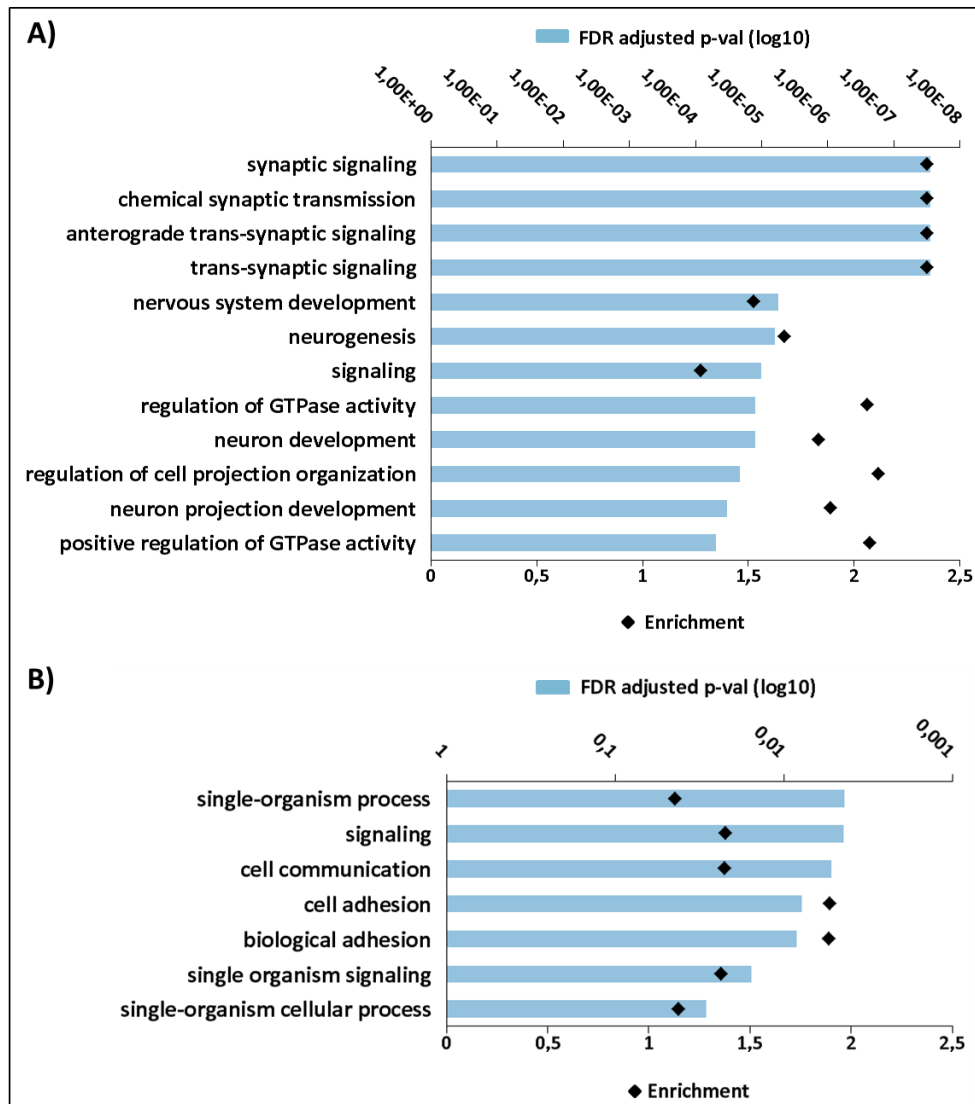


Abbildung 32: Signifikant angereicherte „Terms“ der Ontologie „biological process“. Die Grafiken zeigen die Top 12 und 7 „Terms“ der signifikant angereicherten GO „Terms“ für den Intraspezies-Vergleich im Menschen (A) bzw. im Schimpansen (B). Die Balken stellen den FDR adjusted p-value und die Rauten den Grad der Anreicherung dar.

Gene, die in die Neurogenese involviert sind, wie beispielsweise *CELSR1*, *GDI1*, *IL1RAPL1*, *OPCML*, *OPHN1* und *SHANK3*, waren in den neuronalen und nicht-neuronalen DMRs (HSA vs. PTR) zweifach angereichert (Tabelle 19). Bei diesen Genen könnte es sich um epigenetische Regulatoren handeln, welche in die menschliche Gehirnevolution involviert waren, weshalb sie somit gute Kandidatengene für nachfolgende Analysen darstellen.

Tabelle 19: Mit der Neurogenese assoziierte Gene, die mit interspezifischen DMRs überlappen. Die obere Hälfte der Tabelle zeigt die neuronalen DMRs; die untere Hälfte die nicht-neuronalen DMRs. Angegeben sind die chromosomale Position der DMRs, die Länge und Anzahl der abgedeckten CpG-Stellen (# CG), den Gennamen des DMR-überlappenden Gens und die durchschnittliche Methylierung dieser Region in Prozent.

Zelltyp	Chr.	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
							Mensch	Schimpanse
Neuronal	chr9	139389718	139393758	4041	6	NOTCH1	36	85
	chr10	135010300	135010673	374	6	KNDC1	7	37
	chr11	132662857	132663105	249	7	OPCML ⁺	84	10
	chr12	57591001	57591112	112	5	LRP1	17	70
	chr19	17838761	17838874	114	5	MAP1S	6	55
	chr20	60511870	60516006	4137	5	CDH4 ⁺	69	37
	chrX	67352685	67352815	131	11	OPHN1*	70	29
Nicht-neuronal	chr1	3011508	3011608	101	8	PRDM16	60	23
	chr1	3028919	3029108	190	12		87	4
	chr1	3038277	3038394	118	9		65	14
	chr1	3148336	3148455	120	5		91	49
	chr1	3103003	3103272	270	11		73	38
	chr1	6170013	6170141	129	12	CHD5	61	11
	chr1	10764500	10764665	166	12	CASZ1	64	26
	chr1	10811734	10813837	2104	10		77	32
	chr1	11561458	11562143	686	26	PTCHD2	53	13
	chr1	19983375	19983497	123	7	MINOS;NBL1	55	9
	chr1	22918367	22918520	154	3	EPHA8	87	28
	chr1	177001756	177001904	149	7	ASTN1	89	5
	chr1	205631992	205632102	111	9	SLC45A3	87	29
	chr1	243637718	243637844	127	10	SDCCAG8	89	42
	chr1	245851577	245851741	165	13	KIF26B	12	60
	chr2	121532650	121532965	316	2	GLI2	95	56
	chr2	206572534	206572645	112	5	NRP2	79	21
	chr3	129317949	129318104	156	12	PLXND1	54	10
	chr3	184377554	184377665	112	5	EIF2B5	67	24
	chr5	137593424	137593714	291	18	GFRA3	67	18
	chr6	35778755	35778892	138	4	LHFPL5	83	49
	chr7	42004666	42004966	301	14	GLI3	69	24
	chr7	44279728	44279853	126	3	CAMK2B	81	26
	chr7	150871359	150872590	1232	30	GBX1	68	28
	chr7	156800542	156800664	123	3	MNX1	42	2
	chr8	1877950	1879119	1170	9	ARHGEF10	72	15
	chr8	1882222	1885011	2790	3		89	11
	chr8	1788181	1788341	161	3		95	18
	chr8	68113651	68113763	113	4	ARFGEF1	93	25
	chr8	141454885	141455166	282	4	TRAPPC9	75	27
	chr9	133748333	133748450	118	5	ABL1	73	29
	chr10	73222997	73223132	136	3	CDH23	39	1
	chr11	31840149	31840575	427	4	PAX6	70	26
chr11	61448209	61448360	152	6	DAGLA	13	41	
chr11	61545421	61546428	1008	5	MYRF	56	17	
chr11	76848977	76849143	167	15	MYO7A	66	31	
chr11	113283246	113283362	117	6	DRD2	59	4	
chr11	132662753	132663105	353	17	OPCML ⁺	75	4	
chr15	33360195	33360385	191	11	FMN1	65	7	
chr15	68112105	68112229	125	3	SKOR1	66	22	

chr16	67918866	67919928	1063	5	<i>NRN1L</i>	67	5
chr17	1918769	1918869	101	5	<i>RTN4RL1</i>	93	45
chr17	42989090	42989653	564	4	<i>GFAP</i>	76	32
chr17	71391468	71393610	2143	7	<i>SDK2</i>	86	60
chr17	79045375	79045513	139	8	<i>BAIAP2</i>	69	6
chr17	79022690	79023511	822	5		69	19
chr19	35800743	35801231	489	12	<i>MAG</i>	85	34
chr19	41025452	41025556	105	14	<i>SPTBN4</i>	68	20
chr20	60277050	60277173	124	5	<i>CDH4⁺</i>	61	14
chr22	20229564	20229670	107	7	<i>RTN4R</i>	8	53
chr22	43553694	43555961	2268	4	<i>TSPO</i>	75	21
chr22	46847905	46848027	123	6	<i>CELSR1</i>	89	45
chr22	46315247	46318508	3262	4	<i>WNT7B</i>	84	27
chr22	51169861	51172813	2953	4	<i>SHANK3</i>	83	21
chrX	28632961	28633099	139	5	<i>IL1RAPL1</i>	71	20
chrX	67352685	67352923	239	26	<i>OPHN1*</i>	68	12
chrX	153043548	153044190	643	9	<i>PLXNB3</i>	62	13
chrX	153665659	153665766	108	19	<i>GDI1</i>	55	10
chrX	153673108	153673230	123	7		60	11

* Abdeckung der gleichen Region im selben Gen.

= Abdeckung einer benachbarten Genregion im selben Gen.

1.1.3 DMRs und neuropsychiatrische Erkrankungen

Um weitere Gene zu identifizieren, die einen möglichen Einfluss auf die Gehirn-Evolution des Menschen hatten, wurde eine Kandidatengenanalyse durchgeführt mit Genen, die mit der Entstehung von neuropsychiatrischen Erkrankungen assoziiert sind.

Bei diesem Abgleich wurden die Gene und Promotoren, die mit den DMRs überlappen, mit Kandidatengenen für Alzheimer (AlzGene database, 2017) [Int. 27] [258], Autismus (Autism database AutDB, 2015) [Int. 25] [256], Bipolaren Störungen (Genetic Database for Bipolar Disorder BDgene, 2016) [Int. 26], Mentaler Retardierung (MR) (Anhang-Tabelle 3) und Schizophrenie [257] auf An- bzw. Abreicherung getestet (Tabelle 20). Eine Odds ratio >1 bzw. <1 entspricht einer An- bzw. Abreicherung.

Tabelle 20: Enrichmentanalyse der DMRs in Genlisten, die mit neuropsychiatrischen Erkrankungen assoziiert sind. Die oberen Zahlen geben den p-Wert an, die mittleren stehen für die Odds ratio und die unteren stellen die Anzahl an DMR-überlappenden Genen dar. Signifikante p-Werte (Fisher Test; $p < 0,05$) sind mit einem Stern markiert.

	Intraspezifisch		Interspezifisch	
	neuronal/nicht-neuronal		Mensch/Schimpanse	
	Mensch	Schimpanse	neuronal	nicht-neuronal
Alzheimer	0,007*	1	1	0,311
	2,73	0,50	0	0
	13	1	0	0
Autismus	6,1E-05*	0,002*	0,612	0,109
	2,95	3,52	1,57	2,09
	27	12	1	9
Bipolare Störung	1,4E-04*	0,005*	1	0,612
	2,24	2,49	0,84	1,21
	39	16	1	10
Mentale Retardierung	0,002*	0,456	0,146	0,081
	2,73	1,68	4,89	2,42
	17	4	2	7
Schizophrenie	0,612	0,338	1	1
	1,19	1,49	0,93	0,91
	20	9	1	7

Im intraspezifischen Vergleich waren Gene, die mit Alzheimer (Fisher-Test; $p=0,007$; Odds ratio: 2,73), Autismus (Fisher-Test; $p=6,1E-05$; Odds ratio: 2,95), Bipolaren Störungen (Fisher-Test; $p=1,4E-04$; Odds ratio: 2,24) und MR (Fisher-Test; $p=0,002$; Odds ratio: 2,73) assoziiert sind, in den humanen DMRs signifikant angereichert. Die Gene, die mit Alzheimer und MR assoziiert sind, zeigten keine signifikante Übereinstimmung im Vergleich zu Autismus (Fisher-Test; $p=0,002$; Odds ratio: 3,52) und Bipolaren Störungen (Fisher-Test; $p=0,005$; Odds ratio: 2,49) im Schimpansen. Bezüglich der Schizophrenie-assoziierten Gene ergaben sich Übereinstimmungen, jedoch nicht signifikant angereichert (Tabelle 20, linke Seite).

Es ergaben sich keine Übereinstimmungen mit den Alzheimer-assoziierten Gene sind in den interspezifischen DMRs der neuronalen und nicht-neuronalen Zellen. Dahingegen überlappten Gene, welche mit Autismus, Bipolaren Störungen, MR oder Schizophrenie assoziiert sind mit den interspezifischen DMRs; jedoch ohne signifikante Anreicherung. Die Enrichmentanalyse der MR-assoziierten Gene in den nicht-neuronalen DMRs wies auf eine leichte Tendenz für eine Anreicherung der MR-Gene in den nicht-neuronalen DMRs hin (Fisher-Test; $p=0,081$; Odds ratio: 2,42); in Summe stimmten 7 MR-assoziierte Gene in den

nicht-neuronalen DMRs überein und 2 MR-assoziierte Gene in den neuronalen DMRs (Tabelle 20, rechte Seite).

Alle Gene, die mit MR in Verbindung gebracht werden und mit den interspezifischen DMRs in neuronalen und nicht-neuronalen Zellen übereinstimmen, zeigen eine Hypermethylierung im Menschen im Vergleich zum Schimpansen (Tabelle 21). Die Hypermethylierung fand sich in allen Genen, mit Ausnahme von *GDI1* (Promotorbereich), im Genkörper, was normalerweise mit einer genaktivierenden Wirkung assoziiert ist [260]. Das auf dem X-Chromosom gelegene *OPHN1* Gen lag sowohl in den neuronalen als auch in den nicht-neuronalen Zellen im Menschen hypermethyliert vor – im Vergleich zum Schimpansen. Zwei weitere auf dem X-Chromosom gelegene Gene, *IL1RAPL1* und *GDI1*, waren hypermethyliert in humanen nicht-neuronalen Zellen. *MYT1L* zeigte eine Hypermethylierung in der CTCF Bindregion (Chr2: 1.817.284-1.818.177) in menschlichen und Schimpansenneuronen. Der restliche Genkörper war in neuronalen und nicht-neuronalen Zellen in beiden Spezies hypomethyliert.

Beim Intraspezies-Vergleich zeigten alle MR-Gene sowohl im Menschen als auch im Schimpansen eine neuronale Hypomethylierung im Vergleich zu den nicht-neuronalen Zellen – bis auf eine Ausnahme im Schimpansen. Es stimmten ebenfalls Gene zwischen Mensch und Schimpanse überein, jedoch handelte es sich nicht um die gleichen Regionen, wodurch ein Vergleich der Methylierungswerte nicht möglich ist. Alle Gene/DMRs, die mit neuropsychiatrischen Krankheiten assoziiert sind, können der Tabelle im Anhang entnommen werden (Anhang-Tabelle 4).

Tabelle 21: Inter- und intraspezifische DMRs assoziiert mit Mentaler Retardierung. Die obere Hälfte der Tabelle zeigt den Interspezies-Vergleich und die untere Hälfte den Intraspezies-Vergleich. Angegeben sind die chromosomale Lage der DMRs, die Länge (bp) und Anzahl der abgedeckten CpG-Stellen (# CG), die Gennamen des DMR-überlappenden Gens und die durchschnittliche Methylierung dieser Region in Prozent.

Zelltyp	Chr.	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
							Mensch	Schimpanse
Nicht-neuronal	chr8	141454885	141455166	282	4	<i>TRAPPC9</i>	75	27
	chr8	120779789	120779900	112	4	<i>TAF2</i>	58	13
	chrX	28632961	28633099	139	5	<i>IL1RAPL1</i>	71	20
	chrX	67352685	67352923	239	26	<i>OPHN1*</i>	68	12
	chrX	153665659	153665766	108	19	<i>GDI1</i>	55	10
	chr2	100217832	100218105	274	11	<i>AFF3</i>	69	25

	chr17	71358981	71359087	107	5	<i>SDK2</i>	85	44
	chr17	71391468	71393610	2143	7		86	60
Neuronal	chr2	1817284	1818177	894	10	<i>MYT1L</i>	69	32
	chrX	67352685	67352815	131	11	<i>OPHN1*</i>	70	29

Spezies							Neuronal	Nicht-Neuronal
Mensch	chr2	100217971	100218105	135	10	<i>AFF3</i>	14	65
	chr11	126873871	126873975	105	5	<i>KIRREL3⁼</i>	8	42
	chr2	1828372	1828483	112	11	<i>MYT1L⁼</i>	1	66
	chr2	2297359	2300439	3081	4		9	61
	chr2	2295414	2295587	174	11		7	60
	chr8	141107065	141108104	1040	7	<i>TRAPPC9</i>	2	43
	chr8	141231039	141231144	106	8		23	89
	chr22	37022212	37022374	163	8	<i>CACNG2</i>	21	82
	chr22	36973023	36973376	354	6		11	80
	chr16	89255474	89255720	247	6	<i>CDH15</i>	21	78
	chr7	154065566	154072194	6629	11	<i>DPP6</i>	16	72
	chr14	102450175	102450398	224	5	<i>DYNC1H1</i>	3	69
	chr20	62120339	62120485	147	9	<i>EEF1A2</i>	11	80
	chr1	1739109	1739239	131	9	<i>GNB1</i>	9	62
	chr9	140052457	140052647	191	8	<i>GRIN1</i>	2	36
	chr9	140036019	140036167	149	8		3	64
	chr2	241720020	241721873	1854	13	<i>KIF1A</i>	10	66
	chr2	241721974	241722201	228	7		22	80
	chr19	54684569	54684846	278	8	<i>MBOAT7</i>	12	70
	chr4	119275340	119275472	133	9	<i>PRSS12</i>	5	65
chr1	110706542	110706680	139	4	<i>SLC6A17</i>	13	68	
chr1	44359940	44360114	175	6	<i>ST3GAL3</i>	7	85	
Schimpanse	chrX	28632996	28633099	104	5	<i>IL1RAPL1</i>	67	21
	chr11	126422411	126426956	4546	4	<i>KIRREL3⁼</i>	24	80
	chr2	1818160	1818263	104	5	<i>MYT1L⁼</i>	14	95

* Abdeckung der gleichen Region im selben Gen.

⁼ Abdeckung einer benachbarten Genregion im selben Gen.

1.1.4 DMRs und human-spezifische Expressions- und Histonmethylierungs-Signaturen

Des Weiteren wurden die DMRs mit Kandidatengenlisten abgeglichen, bestehend aus 2.357 Protein-kodierenden Genen, welche ein humanes gehirnspezifisches Expressionsmuster aufweisen [251-254], und 441 Genen [255], die in der Nähe von human-spezifischen Histonmethylierungs-Signaturen lokalisiert sind.

Für die human-spezifisch exprimierte Gene ergaben sich Übereinstimmungen mit allen Kontrasten; im Menschen waren diese signifikant angereichert (Tabelle 22, links). Der Vergleich zwischen den Genen, die nahe human-spezifischer Histonmethylierungs-Signaturen lokalisiert sind, und den DMR-assoziierten Genen, zeigte eine signifikante Anreicherung in den interspezifischen DMRs (Fisher Test; N+: p=0,007, Odds ratio: 5,70; N-: p=0,001, Odds ratio: 2,82) (Tabelle 22, links) und in den humanen intraspezifischen DMRs (Fisher Test; p=1,4E-04, Odds ratio: 2,35) (Tabelle 22, rechts).

Tabelle 22: DMR-Abgleich mit human-spezifisch regulierten Genen. Die oberen Zahlen geben den p-Wert an, die mittleren stehen für die Odds ratio und die unteren stellen die Anzahl an DMR-überlappenden Genen dar. Signifikante p-Werte (Fisher Test; p<0,05) sind zusätzlich mit einem Stern markiert.

	Intraspezifisch		Interspezifisch	
	neuronal/nicht-neuronal		Mensch/Schimpanse	
	Mensch	Schimpanse	neuronal	nicht-neuronal
Expression	6,1E-05*	0,318	0,837	0,318
	2,11	1,46	1,18	0,57
	52	14	2	7
Histonmethylierung	1,4E-04*	0,318	0,007*	0,001*
	2,35	1,61	5,70	2,82
	36	9	5	18

In Summe konnten 143 DMR-assoziierte Gene detektiert werden, welche während der Gehirnevolution human-spezifisch reguliert worden sind. Davon stammen 25 aus dem Kontrast nicht-neuronale Zellen (HSA vs. PTR), 7 aus dem Kontrast neuronale Zellen (HSA vs. PTR), 88 aus dem Kontrast Menschen (N+ vs. N-) und 23 aus dem Kontrast Schimpanse (N+ vs. N-) (Tabelle 23; Anhang-Tabelle 5). Im interspezifischen Vergleich wurden vier Gene detektiert, die sowohl in neuronalen als auch in nicht-neuronalen Zellen spezies-spezifische Methylierungsveränderungen zeigen (Tabelle 23). *OPCML* und *TNFRSF11A* waren im Menschen in beiden Zelltypen hypermethyliert. *ZNF717* dagegen zeigte eine spezies- und zelltyp-spezifische Methylierung. Die Promotorregion lag im Menschen in den neuronalen und nicht-neuronalen Zellen hypomethyliert im Vergleich zum Schimpansen vor. Die weiter downstream gelegene Region wies eine Hypermethylierung in den nicht-neuronalen Zellen auf; der Genkörper war im Menschen neuronal hypomethyliert und im Schimpansen hypermethyliert, im Vergleich zu den nicht-neuronalen Zellen (Tabelle 23; Anhang-Tabelle 5). In der nachfolgenden Tabelle wurden aufgrund des funktionellen Einflusses der DNA-

Methylierung in Promotorregionen auf die Transkription [91, 92, 94] zusätzlich die DMRs markiert, welche die Promotorregion der Gene abdecken (Tabelle 23).

Tabelle 23: Interspezifische DMRs assoziiert mit human-spezifischer Regulation im Gehirn. Angegeben sind die chromosomale Lage der DMRs, die Länge (bp) und Anzahl der abgedeckten CpG-Stellen (# CG), der Gennamen des DMR-überlappenden Gens und die durchschnittliche Methylierung dieser Region in Prozent.

Zelltyp	human-spezifische Regulation	Chr.	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
								Mensch	Schimpanse
Neuronal	H3K4me3	chr2	1817284	1818177	894	10	MYT1L	69	32
	H3K4me3	chr3	75834642	75834760	119	16	ZNF717*	1	51
	Expression, H3K4me3	chr10	106034918	106035147	230	6	GSTO2	8	85
	Expression, H3K4me3	chr11	132662857	132663105	249	7	OPCML*	84	10
	H3K4me3	chr18	60052067	60052345	279	16	TNFRSF11A*	53	1
Nicht-neuronal	H3K4me3	chr1	3376370	3376578	209	6	ARGF16	81	8
	H3K4me3	chr1	10811734	10813837	2104	10	CASZ1	77	32
	H3K4me3	chr1	17198730	17198847	118	14	CROCC	73	13
	H3K4me3	chr3	75827859	75829467	1609	4	ZNF717*	78	2
			75834604	75834754	151	18		7	33
	Expression	chr3	129317949	129318104	156	12	PLXND1	57	10
	H3K4me3	chr5	177732259	177733520	1262	5	COL23A1	91	46
	H3K4me3	chr6	656770	656921	152	14	EXOC2	83	31
	Expression	chr6	158409180	158411724	2545	3	SYNJ2	70	9
	Expression	chr7	44279728	44279853	126	3	CAMK2B	81	26
	Expression	chr8	1788181	1788341	161	3	ARHGEF10	95	18
			1877950	1879119	1170	9		72	18
			1882222	1885011	2790	3		89	11
	Expression	chr10	132953684	132953784	101	3	TCERG1L	79	39
	H3K4me3	chr11	859590	859702	113	4	TSPAN4	34	63
	Expression, H3K4me3	chr11	132662853	132663105	353	17	OPCML*	75	42
	H3K4me3	chr15	101523646	101523753	108	3	LRRK1	91	6
	H3K4me3	chr16	3187426	3188788	1363	5	ZNF213	88	23
	H3K4me3	chr17	7907136	7907487	352	25	GUCY2C	77	37
	H3K4me3	chr17	20799775	20799923	149	13	CCDC144NL	68	15
	H3K4me3	chr17	45439767	45439877	111	7	EFCAB13	46	11
Expression, H3K4me3	chr17	54989090	42989653	564	4	GFAP	76	32	
H3K4me3	chr18	60052067	60052297	231	28	TNFRSF11A*	78	3	
H3K4me3	chr19	36303414	36303574	161	3	PRODH2	86	28	
H3K4me3	chr20	25846963	25847117	155	5	FAM182B	22	78	
H3K4me3	chr22	44968061	44968163	103	5	LINC00207	79	24	
H3K4me3	chr22	46847905	46848027	123	6	CELSR1	89	45	

* Abdeckung der gleichen Region im selben Gen.

⁼ Abdeckung einer benachbarten Genregion im selben Gen.

^{||} Abdeckung der Promotorregion dieses Gens.

1.1.5 Repetitive Elemente assoziiert mit DMRs

Aufgrund der stetigen Aktivität und Akkumulation der Transposablen Elemente (TEs) im menschlichen Genom im Laufe mehrerer Jahrmillionen, haben diese einen starken Einfluss auf die Evolution des Primaten-Genoms genommen – sowohl in Struktur als auch Funktion [261]. Im Folgenden wurde überprüft, inwieweit eine differenzielle Methylierung in den repetitiven Elementen (REs) zwischen den Spezies bzw. Zelltypen vorliegt. Hierfür wurden alle annotierten REs mit den DMRs abgeglichen.

In Summe wurden 1015 REs durch den RRBS-Datensatz abgedeckt, davon waren 351 REs in DMRs lokalisiert. In den meisten Fällen war dasselbe RE in mehreren DMRs lokalisiert. Die meisten der REs konnten den beiden Hauptklassen, long terminal repeat (LTR) und non-long terminal repeat (non-LTR) Retroelemente zugeordnet werden. Innerhalb der LTRs waren vier Familien (LTR, HERV, MLT und MER) und zwei MER Subfamilien (Charlie, Tigger) vertreten. Bei den non-LTRs fanden sich überwiegend RE aus den Familien L1, ALU und MIR. In der nachfolgenden Analyse lag der Fokus auf den REs, die eine Hypomethylierung aufweisen, da bekannt ist, dass die Transkription und Transposition der REs mittels Methylierung supprimiert wird [136, 262].

In den RE-Familien MLT, MER, Charlie, Tigger, L1, ALU und MIR waren zwischen 70% und 90% aller DMR-assozierten REs neuronal hypomethyliert im Menschen. Rund 10-20% der nicht-neuronalen DMR-assozierten REs zeigten eine Hypomethylierung im Schimpansen. Alle weiteren Kontraste enthielten weniger als 10% an hypomethylierten REs. In der RE-Familie LTR waren jeweils rund 30% aller REs im Intraspezies-Vergleich neuronal im Menschen hypomethyliert und im interspezifischen Vergleich in den nicht-neuronalen Zellen im Schimpansen hypomethyliert. In der RE-Familie HERV waren dagegen 90% der hypomethylierten REs in Regionen lokalisiert, die im Menschen bzw. Schimpansen in den nicht-neuronalen Zellen hypomethyliert waren (Abbildung 33). Bei der weiteren Analyse wurde der Fokus ausschließlich auf die L1 REs gelegt, da angenommen wird, dass eine L1-Aktivität im Gehirn vorliegt, besonders in den neuronalen Zellen, wodurch es zu einer L1-induzierten genomischen Variabilität kommt. Dieser Mechanismus trägt möglicherweise zu neuronaler Diversität bei, welche einen Einfluss auf die neuronale Plastizität, Kognition und letztendlich das Verhalten hat [263]. In Tabelle 24 sind alle DMR-assozierten Gene dargestellt, welche ein L1 RE enthalten und in einem Zusammenhang mit Funktionen im

Gehirn stehen. Eine vollständige Liste aller DMR-assozierten Gene, die L1 REs enthalten, findet sich im Anhang (Anhang-Tabelle 6).

Vergleicht man die Gene, welche mehrfach in der Tabelle auftreten, findet sich mit *ARHGEF10* (Rho guanine nucleotide exchange factor 10) die einzige Übereinstimmung für die gleiche DMR. Des Weiteren war auffällig, dass 16% (n=25/156) der Regionen in den L1-assozierten DMRs mit MicroRNAs (miRNAs; MIR) übereinstimmten (Tabelle 24; Anhang-Tabelle 6).

Tabelle 24: Ausgewählte LINE1-assozierte DMRs. Aufgelistet sind die DMR-assozierten Gene, welche ein L1 RE enthalten und mit Funktionen im Gehirn in Zusammenhang gebracht werden. Angegeben sind die chromosomale Lage der DMRs, die Länge (bp) und Anzahl der abgedeckten CpG-Stellen (# CG), der Genname des DMR-überlappenden Gens und die durchschnittliche Methylierung dieser Region in Prozent.

	Name RE	Lage REs	Lage DMRs	up-stream	abgedeckt	down-stream	Methylierung (%)	
							Neuronal	Nicht-neuronal
Mensch N+→	L1M4	chr7:1530213-1530838	chr7:1528740-1533963	RP11-1246C19.1	INTS1	AC102953.6	18	68
	L1M5	chr1:1973799-1973912	chr1:1972875-1974300	GABRD		RP11-547D24.3	19	83
		chr10:134117562-134117711	chr10:134115790-134120496	DPYSL4	STK32C	LRR27	5	36
	L1MB3	chr17:74915816-74916009	chr17:74906329-74920018	LINC00868	MGAT5B	RP11-87G24.3	5	66
		chr3:66002407-66002828	chr3:66002247-66002949	AC106827.1	MAG11	SLC25A26	15	80
		chr7:154070706-154071073	chr7:154065566-154072194	AC006019.4	DPP6	RP11-476H24.1	16	72
	L1MB7	chr17:79033969-79034383; chr17:79034418-79034514	chr17:79032885-79038320	BAIAP2-AS1	BAIAP2	AATK	10	82
		chr17:79050785-79050991	chr17:79050932-79052330	BAIAP2-AS1	BAIAP2	AATK	17	75
	L1MB8	chr17:79062172-79062553	chr17:79060745-79066325	BAIAP2-AS1	BAIAP2	AATK	13	79
		chr7:1530893-1531246; chr7:1531556-1532242	chr7:1528740-1533963	RP11-1246C19.1	INTS1	AC102953.6	18	68
		chr9:132981261-132981346	chr9:132975862-132982367	AL360004.1	NCS1	HMCN2	19	78
	L1MC1	chr8:1850670-1850956	chr8:1849599-1853008	MIR596	ARHGEF10*	KBTBD11-OT1	4	60
	L1MC2	chr22:49062154-49062309; chr22:49062416-49062539	chr22:49060730-49062509	CTA-299D3.8	FAM19A5	MIR4535	11	73
	L1MC4	chr1:154546653-154547023	chr1:154544691-154548824	UBE2Q1	CHRN2	RP11-61L14.6	8	81
	chr2:99157465-99157508; chr2:99157806-99157902	chr2:99155916-99159945	CNGA3	INPP4A	COA5	14	77	
	chr7:674216-674280; chr7:674577-674763	chr7:672307-674963	AC147651.4	PRKAR1B	HEATR2	6	53	
	chr9:140724526-140725011; chr9:140725684-140725847	chr9:140723061-140727855	RP11-188C12.2	EHMT1	MIR602	2	56	

L1MC4a	chr10:133777999-133778557; chr10:133779451-133779690	chr10:133777895-133782644	PPP2R2D	BNIP3	JAKMIP3	7	72	
	chr7:672787-672982	chr7:672307-674963	AC147651.4	PRKAR1B	HEATR2	6	53	
	L1MC5	chr2:233757914-233758067	chr2:233757595-233759265	C2orf82	NGEF	AC106876.2	30	86
		chr7:44318660-44318802; chr7:44318825-44318915	chr7:44317254-44321187	YKT6	CAMK2B	NUDCD3	11	87
		chr9:140724419-140724469	chr9:140723061-140727855	RP11-188C12.2	EHMT1	MIR602	2	56
	L1MD	chr10:134055029-134055141	chr10:134053498-134056965	DPYSL4	STK32C	LRRC27	18	78
	L1MD2	chr4:1228774-1229113	chr4:1228936-1229135	SPON2	CTBP1	CTBP1-AS2	19	73
		chr9:140615055-140615585; chr9:140616142-140616545	chr9:140614916-140617481	SETP5	EHMT1	RP11-188C12.2	9	79
	L1ME1	chr17:79035129-79035500; chr17:79035513-79035746; chr17:79036049-79036149	chr17:79032885-79038320	BAIAP2-AS1	BAIAP2	AATK	10	82
		chr4:3204237-3204456; chr4:3204751-3204905	chr4:3201467-3205977	HTT-AS	HTT	MSANTD1	3	60
L1ME2z	chr20:44808197-44808282	chr20:44806684-44810899	RP5-998H6.2	CDH22	SLC35C2	11	80	
	chr9:140723775-140723886	chr9:140723061-140727855	RP11-188C12.2	EHMT1	MIR602	2	56	
L1ME3	hr22:49060427-49061116; chr22:49061135-49061462	chr22:49060730-49062509	CTA-299D3.8	FAM19A5	MIR4535	11	73	
L1MEa	chr9:140725959-140726208	chr9:140723061-140727855	RP11-188C12.2	EHMT1	MIR602	2	56	
	chr7:728576-728758; chr7:729057-729113	chr7:727590-732295	AC147651.4	PRKAR1B	HEATR2	8	90	
L1PA16	chr7:729949-730113	chr7:727590-732295	AC147651.4	PRKAR1B	HEATR2	8	90	
	chr8:1850154-1850381	chr8:1849599-1853008	MIR596	ARHGEF10*	KBTBD11-OT1	4	60	

						Mensch	Schimpanse			
Nicht-neuronal	HSA ↓	L1MB5	chr16:1160593-1160798; chr16:1160900-1160999	chr16:1158680-1162581	LA16c-349E11.1	LA16c-381G6.1	CACNA1H	27	84	
		L1ME2z	chr17:1032842-1032929; chr17:1033240-1033407	chr17:1029918-1033700	MIR3183		ABR	AC016292.1	27	62
		L1PA8	chr17:1032409-1032518; chr17:1032825-1032837	chr17:1029918-1033700	MIR3183		ABR	AC016292.1	27	62

HSA ↑	L1M1	chrX:153390703-153392264	chrX:153391129-153391259	MECP2		OPN1LW	45	8
	L1M5	chr8:1882726-1883281	chr8:1882222-1885011	MIR596	ARHGEF10*	KBTBD11-OT1	89	11
	L1MA10	chr20:60543018-60543194; chr20:60543305-60543411	chr20:60541116-60545298	MIR1257	TAF4	RP11-11M20.2	60	27
	L1MB3	chr17:77241511-77241625; chr17:77241626-77241710	chr17:77239436-77242699	RP11-398J5.1	RBFOX3	MIR4739	60	12
	L1MC1	chr8:1883277-1883653	chr8:1882222-1885011	MIR596	ARHGEF10*	KBTBD11-OT1	89	11
	L1MC3	chr17:77241497-77241510	chr17:77239436-77242699	RP11-398J5.1	RBFOX3	MIR4739	60	12
	L1ME3	chr20:60545138-60545654	chr20:60541116-60545298	MIR1257	TAF4	RP11-11M20.2	60	27

* Abdeckung der gleichen Region im selben Gen.

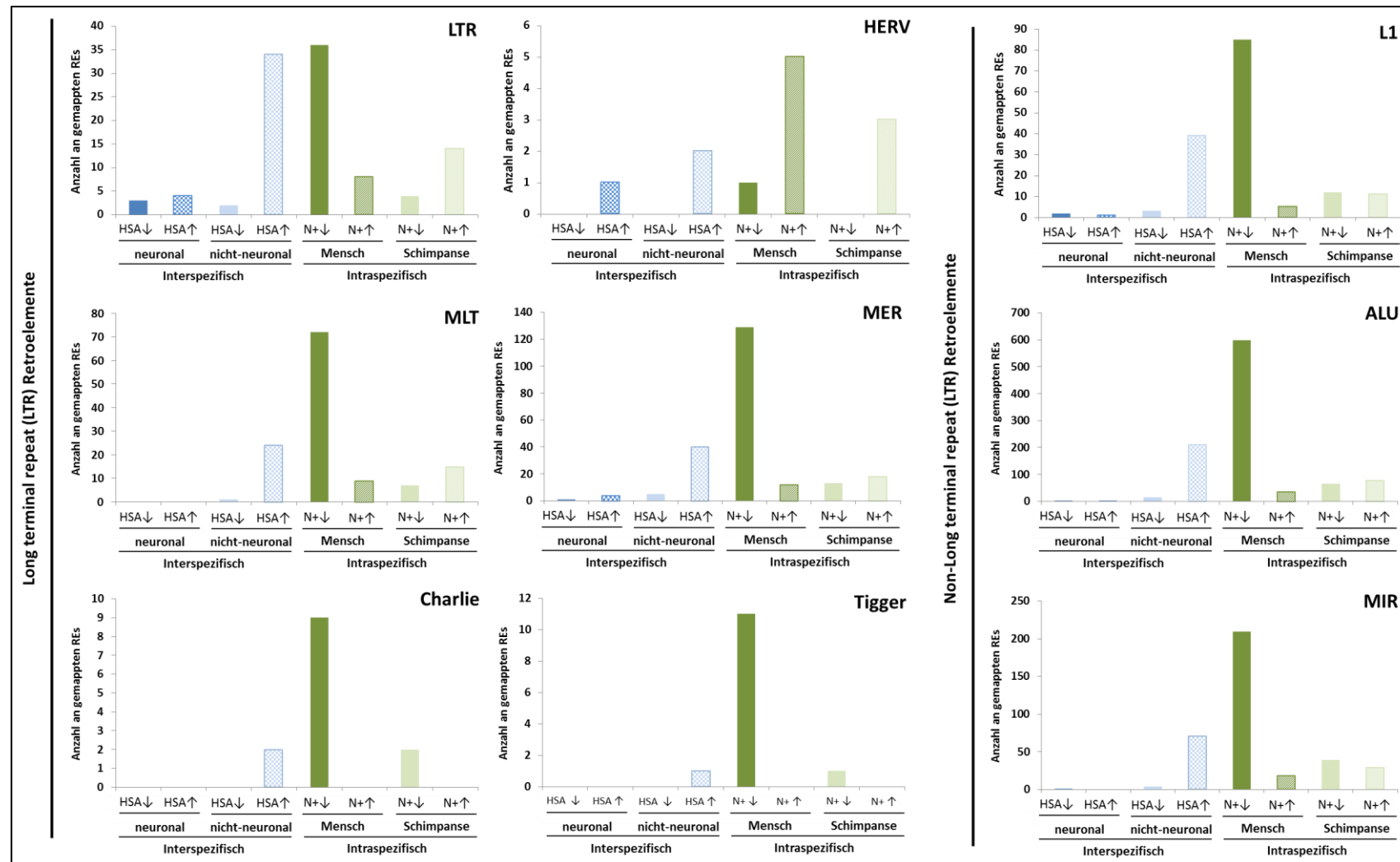


Abbildung 33: Assoziierung verschiedener Repeat-Klassen mit intra- und interspezifischen DMRs. Die Anzahl DMR-assoziiierter repetitiver Elemente ist auf der Y-Achse dargestellt und die jeweiligen DMR-Kontraste auf der X-Achse. Die blauen Balken repräsentieren die interspezifischen DMRs und die grünen Balken die intraspezifischen DMRs. Die unterschiedlichen Repeat-Klassen weisen eine verminderte Methylierung in den humanen neuronalen Zellen im Vergleich zu den

nicht-neuronalen Zellen auf. Zudem zeigen die humanen nicht-neuronalen Zellen eine höhere Methylierung in einigen Repeats als in nicht-neuronalen Zellen des Schimpansen. N+: neuronale Zellen; HSA: Mensch; ↑ hypermethyliert; ↓ hypomethyliert.

1.2 Validierung der DMRs via Bisulfit-Pyrosequenzierung

Für die Validierung der RRBS-Daten wurde je Kontrast ein Gen ausgewählt und mittels Bisulfit-Pyrosequenzierung in den neuronalen und nicht-neuronalen Zellen validiert. Bei der Wahl der Gene wurde der Fokus auf eine signifikant inverse differenzielle Methylierung zwischen den Zelltypen bzw. Spezies gelegt. Ein Gen wird als invers differenziell methyliert eingestuft, wenn ein Zelltyp/Spezies eine Hypomethylierung (<25%) und der andere Zelltyp/Spezies eine Hypermethylierung (>75%) aufweist. Des Weiteren wurde mittels Literaturrecherche überprüft, welches dieser invers methylierten Gene in einem funktionellen Zusammenhang mit der Gehirnaktivität steht.

Calmodulin binding transcription activator 1 (*CAMTA1*) wies einen Methylierungsunterschied in den nicht-neuronalen Zellen zwischen Mensch und Schimpanse auf. Mittels des Pyro-Assays konnten alle 5 CpG-Stellen (Differenz \emptyset Methylierung = +38%; T-Test $p = 0,000$) überprüft werden und die RRBS-Ergebnisse (Differenz \emptyset Methylierung = +56%; T-Test $p = 0,008$) erfolgreich bestätigt werden (Abbildung 34, A). Proline and serine rich 2 (*PROSER2*) zeigte einen Methylierungsunterschied in den neuronalen Zellen zwischen Mensch und Schimpanse. Über den Pyro-Assay konnten nur 3 CpG-Stellen der *PROSER2* DMR getestet werden, da die Sequenz des Schimpansen an dieser Stelle nicht exakt definiert werden konnte. Jedoch konnte die Übereinstimmung der Methylierungswerte zwischen den Pyro- (Differenz \emptyset Methylierung = -46%; T-Test $p = 0,000$) und RRBS-Daten (Differenz \emptyset Methylierung = -53%; T-Test $p = 0,000$) anhand dieser 3 CpGs gezeigt werden, wie in Abbildung 34, B dargestellt. RNA binding fox-1 homolog 3 (*RBFOX3*) war im Menschen zwischen den neuronalen und nicht-neuronalen Zellen invers differenziell methyliert. Durch den Pyro-Assay wurden 5 der 7 CpG-Stellen (Differenz \emptyset Methylierung = -39%; T-Test $p = 0,006$) innerhalb der DMR untersucht und die differenzielle Methylierung der RRBS-Daten (Differenz \emptyset Methylierung = -42%; T-Test $p = 0,000$) erfolgreich bestätigt. Die Werte sind grafisch in Abbildung 34, C dargestellt. Reticulon 4 receptor (*RTN4R*) wies eine Hypomethylierung in den neuronalen Zellen und eine Hypermethylierung in den nicht-neuronalen Zellen im Schimpansen auf.

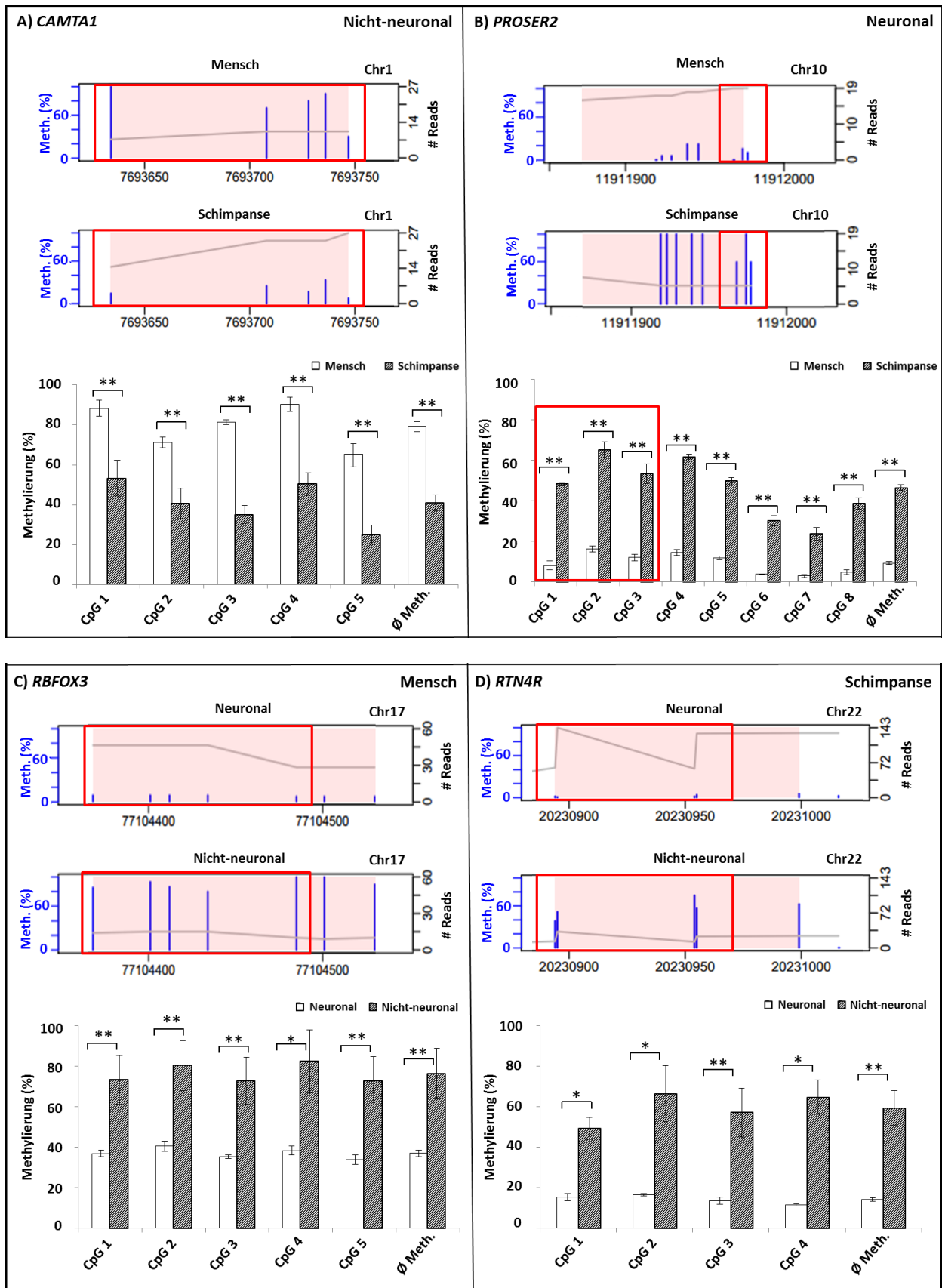


Abbildung 34: Validierung der RRBS-Daten mittels Pyrosequenzierung. Die obere Abbildung stellt die RRBS-Daten dar und die untere die Ergebnisse der Pyrosequenzierung für das jeweilige Gen (*CAMTA1*, *PROSER2*, *RBFOX3*, *RTN4R*). In der oberen Abbildung sind die dünnen blauen Balken repräsentativ für die Methylierungswerte (Y-Achse, links) der einzelnen CpG-Stellen und die graue

Linie für die Read-Coverage an den CpG-Stellen (Y-Achse, rechts). Auf der X-Achse ist die genomische Position der CpGs angegeben. In der unteren Abbildung sind die Methylierungswerte auf der Y-Achse und die einzelnen CpG-Stellen auf der X-Achse dargestellt. Signifikante intra- oder interspezifische Methylierungsunterschiede sind mit einem Stern (T-Test $p < 0,05$) oder mit zwei Sternen (T-Test $p < 0,01$) markiert. Der rote Rahmen hebt die zwischen RRBS und Pyrosequenzierung übereinstimmenden CpG-Stellen hervor.

Via Pyro-Assay wurden 4 von 6 CpG-Stellen (Differenz \emptyset Methylierung = -45%; T-Test $p = 0,000$) überprüft. Das invers-differenzielle Methylierungsmuster der RRBS-Daten (Differenz \emptyset Methylierung = -41%; T-Test $p = 0,000$) konnte ebenfalls validiert werden (Abbildung 34, D).

2. Methylierungsanalysen bei Brustkrebs via DBAS

Um das Methylierungsprofil von *BRCA1/BRCA2*-mutationsnegativen Brustkrebspatientinnen auf Einzel-Allel-Ebene erstellen zu können, wurden die Promotorbereiche ausgewählter TSG (*BRCA1*, *BRCA2*, *RAD51C*, *ATM* (Region 1 und 2), *PTEN*, *TP53* (Region 1 und 2), *MLH1*, und *RB1*) sowie die „promoter flanking region“ von *ESR1* mittels Deep Bisulfite Amplicon Sequencing (DBAS) untersucht. DNA aus peripherem Blut wurde zum einen von 43 Frauen mit early-onset Brustkrebs (EO BC) und zum anderen von 36 Brustkrebspatientinnen mit einem berechneten Heterozygotenrisiko von mehr als 95% (HR BC) analysiert. Zudem wurden 80 alters-gematchte Proben von krebsfreien Personen als Kontrollproben eingesetzt, bei denen auch innerhalb der Familie keine Brust- und Ovarienkrebserkrankungen aufgetreten bzw. bekannt sind. Ergänzend zu den Methylierungsanalysen wurde bei einigen ausgewählten Patientenproben eine Expressionsanalyse der TSG durchgeführt. Alle Ergebnisse wurden in der Publikation „Single CpG hypermethylation, allele methylation errors, and decreased expression of multiple tumor suppressor genes in normal body cells of mutation-negative early-onset and high-risk breast cancer patients“ im International Journal of Cancer veröffentlicht [264].

2.1 Analyse der DBAS-Daten

Das DBAS liefert Informationen über den Methylierungsgrad aller CpG-Stellen des gesamten Amplikons, aber auch auf Einzel-Allel-Ebene. In der nachfolgenden Grafik sind die untersuchten Amplikons und ihre durchschnittliche Read-Coverage inklusive Standardabweichung (engl. standard deviation, SD) abgebildet (Abbildung 35). Die Anzahl der Reads variiert je nach untersuchter Region stark. Die geringste Abdeckung findet sich bei *BRCA2* (5.307 ± 3.207), *BRCA1* (6.987 ± 4.827) und *TP53* Region 2 (7.147 ± 5.818), wohingegen *MLH1* (29.966 ± 28.175) die höchste durchschnittliche Coverage, aber auch eine starke Varianz aufweist.

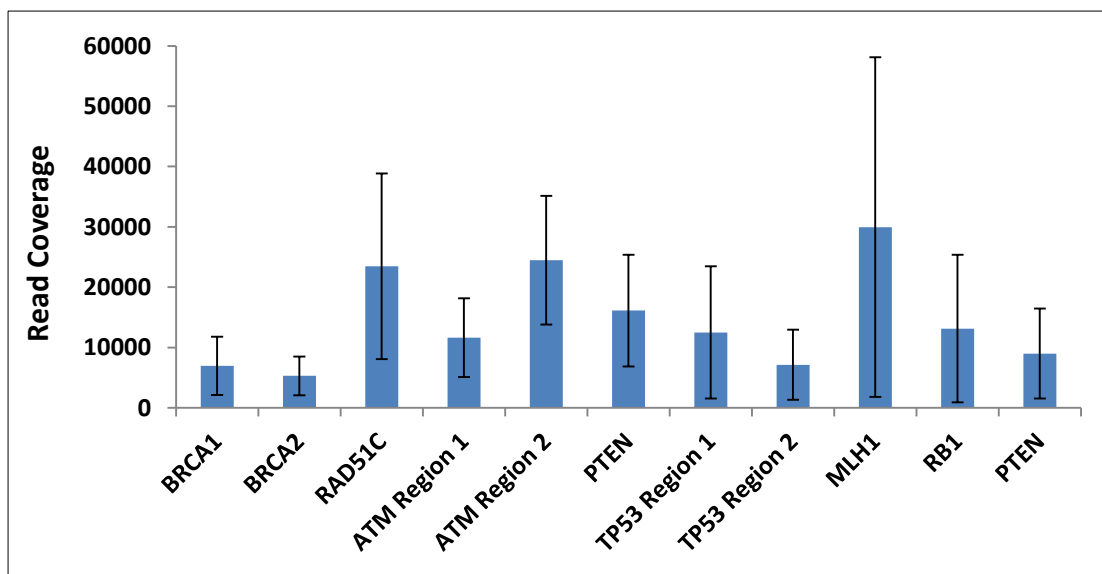


Abbildung 35: Durchschnittliche Coverage je Amplicon inklusive Standardabweichung (SD).

2.1.1 DNA-Methylierungsprofile von TSG

Die durchschnittliche Methylierung über alle im Amplicon abgedeckten CpG-Stellen liegt bei den Tumorsuppressorgenen unter 1% und für *ESR1* bei rund 3%. Es konnte kein signifikanter Methylierungsunterschied zwischen den Patienten- und Kontrollproben detektiert werden; ebensowenig zwischen den Patienten beider Brustkrebskohorten (Tabelle 25).

Tabelle 25: Durchschnittliche Methylierungslevel mit Angaben zu den Interquartilbereichen je Amplicon und Kohorte bzw. Gruppe.

	Durchschnittliche Methylierungslevel ± SD (%)			
	EO BC		Kontrolle	
	Interquartilbereich (%)			
	EO BC	Kontrolle	HR BC	Kontrolle
BRCA1	0,540 ± 0,036 (0,2)	0,662 ± 0,059 (0,2)	0,850 ± 0,352 (0,2)	0,369 ± 0,014 (0,1)
BRCA2	N/A	N/A	0,347±0,013 (0,1)	0,337±0,014 (0,1)
RAD51C	0,549±0,049 (0,2)	0,585±0,027 (0,1)	0,494±0,039 (0,1)	0,481±0,032 (0,1)
ATM Region 1	0,342±0,014 (0,1)	0,375±0,010 (0,1)	0,292±0,012 (0,1)	0,264±0,009 (0,1)
ATM Region 2	0,388±0,014 (0,1)	0,460±0,012 (0,1)	0,356±0,032 (0,0)	0,333±0,001 (0,1)
PTEN	0,403±0,018 (0,2)	0,434±0,016 (0,1)	0,308±0,015 (0,0)	0,289±0,010 (0,0)
TP53 Region 1	0,366±0,015	0,410±0,013	0,319±0,011	0,319±0,009

	(0,1)	(0,1)	(0,1)	(0,1)
TP53 Region 2	0,612±0,059 (0,2)	0,514±0,016 (0,2)	0,389±0,022 (0,1)	0,403±0,027 (0,1)
MLH1	0,381±0,017 (0,2)	0,456±0,014 (0,1)	0,386±0,039 (0,1)	0,339±0,008 (0,1)
RB1	0,448±0,015 (0,1)	0,507±0,012 (0,1)	0,356±0,008 (0,1)	0,367±0,011 (0,1)
ESR1	2,902±0,120 (1,4)	3,055±0,098 (0,8)	3,078±0,142 (1,3)	3,064±0,112 (0,8)

Beide Patientenkohorten kombiniert zeigen durchschnittliche Methylierungswerte von 0,683 ± 0,163% für *BRCA1*; 0,347 ± 0,014% für *BRCA2*; 0,524 ± 0,032% für *RAD51C*; 0,319 ± 0,010% für *ATM* region 1; 0,373 ± 0,016% für *ATM* region 2; 0,352 ± 0,012% für *PTEN*, 0,344 ± 0,010% für *TP53* region 1; 0,510 ± 0,035% für *TP53* region 2; 0,383 ± 0,020% für *MLH1* und 0,405 ± 0,010% für *RB1* (Abbildung 36). Die durchschnittliche Methylierung der „promoter flanking“ Region von *ESR1* war etwa 5 bis 10 Mal höher (3,123 ± 0,090%) im Vergleich zu den TSG. In den 159 analysierten Patienten- und Kontrollproben konnte bei 28 Proben (17,6%) eine erhöhte durchschnittliche Methylierung zwischen 1% und 3% gemessen werden. Daher kann angenommen werden, dass der unmethylierte Status von TSG im Promotorbereich sehr streng reguliert ist. In einer früheren Studie wurden via Bisulfit-Pyrosequenzierung mehr als 600 *BRCA1/BRCA2* mutations-negative BC und/oder Ovarienkrebspatientinnen untersucht [265]. Ein Schwellenwert von 6% durchschnittlicher Methylierung wurde hier eingesetzt, um Individuen mit einer konstitutiven Epimutation ausfindig machen zu können. In unserem Datenset konnte ebenfalls eine Probe mit einer durchschnittlichen Methylierung >6% in *BRCA1* (HR BC 30: 12,9%) detektiert werden. Dies entspricht, alle untersuchten EO und HR BC-Patienten einbezogen, einem Anteil von 1,3% und ist vergleichbar mit dem Epimutationsscreening von Hansmann *et al.* (2012). In dieser Studie wiesen 9 von 641 (1,4%) Patienten einen Methylierungswert von mehr als 6% auf.

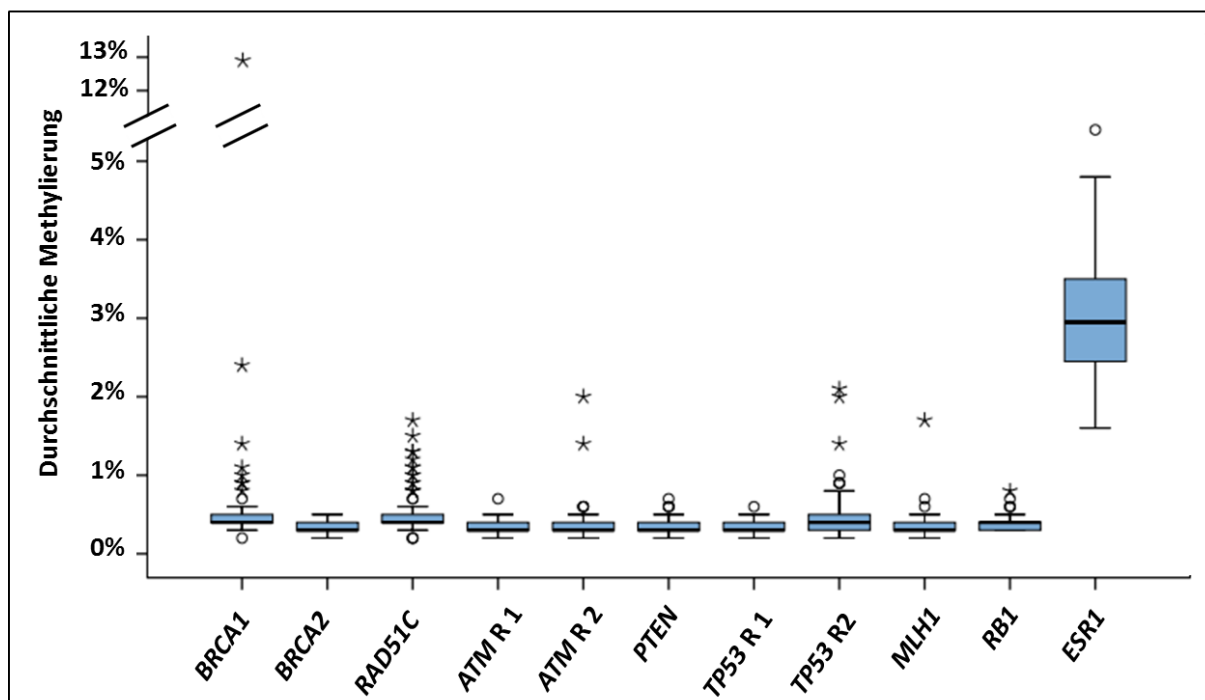


Abbildung 36: Durchschnittliche Methylierungslevel je Amplikon, zusammengefasst für alle Patienten und Kontrollproben, dargestellt als Boxplot Diagramm. Der Methylierungswert wurde über alle im Amplikon enthaltenen CpG-Stellen je Probe ermittelt. Je Amplikon wurden alle Patienten- und Kontrollproben kombiniert, da zwischen den Gruppen und Kohorten keine signifikanten Methylierungsunterschiede detektiert werden konnten. Die untere bzw. die obere Linie der Box repräsentiert die 25% Quartile bzw. 75% Quartile. Die dicke Linie in der Mitte der Box gibt den Median an. Die Whiskers zeigen den minimalen bzw. maximalen Wert. Der Interquartilbereich (IQB) berechnet sich aus der Differenz der 25% und 75% Quartile. Proben, die mit einem Kreis bzw. Stern gekennzeichnet sind, entsprechen den Ausreißern bzw. Extremwerten. Diese berechnen sich aus 1,5 Mal bzw. 3 Mal dem IQB plus der 75% Quartile.

2.1.2 Epimutationsraten in TSG

Angelehnt an die bereits erwähnte Studie von Hansmann *et al.* (2012) wurden die Proben einem Screening zur Detektion von Epimutationen auf Einzel-Allel-Ebene unterzogen. Das DBAS ist verglichen mit der Plasmid-Bisulfit-Sequenzierung wesentlich günstiger und effizienter und war somit die Technik der Wahl.

Übereinstimmend mit früheren Studien [266-268] wurde ein Allel als Epimutation mit funktioneller Relevanz definiert, wenn 50% oder mehr der CpG-Stellen des Amplikons eine aberrante Methylierung zeigten. Im Fall der TSG entsprechen Allele mit einer Methylierung von 50% und mehr einer Epimutation, da diese im Promotorbereich in unmethylierter Form vorliegen sollten. Dies konnte bereits über die niedrigen durchschnittlichen

Methylierungslevel bestätigt werden. Da es sich bei *ESR1* nicht um ein TSG handelt, wurde dieses Gen/Amplikon aus der Epimutationsanalyse ausgeschlossen.

In Summe wurden bei 49% (21/43) der EO BC, bei 44% (16/36) der HR BC und bei 43% (34/80) der Kontrollkohorte Proben mit Epimutationen detektiert. In Tabelle 26 ist eine Übersicht über die Anzahl der EO/HR BC-Patientinnen und Kontrollen je Gen dargestellt, die Allele mit veränderter Methylierung zeigen. Die größte Anzahl an aberranten Proben wurden in *RAD51C* (43/163; 26%), *BRCA1* (20/156; 13%) und *TP53* Region 2 (25/159; 16%) beobachtet. Am wenigsten betroffen waren die Promotorbereiche von *ATM* Region 1 (1/159; 1%) und 2 (3/159; 2%), *MLH1* (2/159; 1%) und *PTEN* (3/135; 2%). Die untersuchten Bereiche für *BRCA2* und *TP53* Region 1 waren gänzlich unauffällig. In *BRCA1* waren signifikant mehr BC Proben im Vergleich zu den Kontrollen (jeweils kombinierte Kohorten) betroffen (p=0,016; Fisher exakt).

Tabelle 26: Anzahl an Epimutationen (>50% methylierte CpGs) in BC- und Kontrollproben je Kohorte und Amplikon.

Anzahl der Proben mit Epimutationen (%)					
	EO BC	Kontrolle	HR BC	Kontrolle	p-Wert ^{a,b}
<i>BRCA1</i>	7/42 (17%)	4/42 (10%)	8/36 (22%)	1/36 (3%)	0,016
<i>BRCA2</i>	N/A	N/A	0/36 (0%)	0/36 (0%)	
<i>RAD51C</i>	14/43 (33%)	13/44 (30%)	9/36 (25%)	7/36 (19%)	
<i>ATM</i> Region 1	0/43 (0%)	0/44 (0%)	1/36 (3%)	0/36 (0%)	
<i>ATM</i> Region 2	1/43 (2%)	0/44 (0%)	1/36 (3%)	1/36 (3%)	
<i>PTEN</i>	0/31 (0%)	0/32 (0%)	3/36 (8%)	0/36 (0%)	
<i>TP53</i> Region 1	0/41 (0%)	0/42 (0%)	0/36 (0%)	0/36 (0%)	
<i>TP53</i> Region 2	2/43 (5%)	5/44 (11%)	9/36 (25%)	9/36 (25%)	
<i>MLH1</i>	0/43 (0%)	0/44 (0%)	2/36 (3%)	0/36 (0%)	
<i>RB1</i>	0/42 (0%)	0/42 (0%)	0/36 (0%)	1/36 (3%)	

^a Fisher-Test; Signifikanzniveau ist 0,05.

^b Gruppenvergleich BC und Kontrollen (kombinierte Kohorten).

Die Verteilung der Epimutations-Raten (EMR), berechnet aus der Gesamtzahl der Reads durch die Anzahl an aberrant methylierten Allelen je Probe und Amplikon, verhielt sich sehr unterschiedlich in den einzelnen Kohorten und Regionen (Abbildung 37). Mehr als 90% aller

Proben zeigten eine EMR unter 1% und rund 8% der Proben eine EMR von 1% - 2,5%. Es konnte lediglich eine Patientenprobe in *BRCA1* mit einer EMR von 14,7% (durchschnittliche Methylierung 12,9%) detektiert werden.

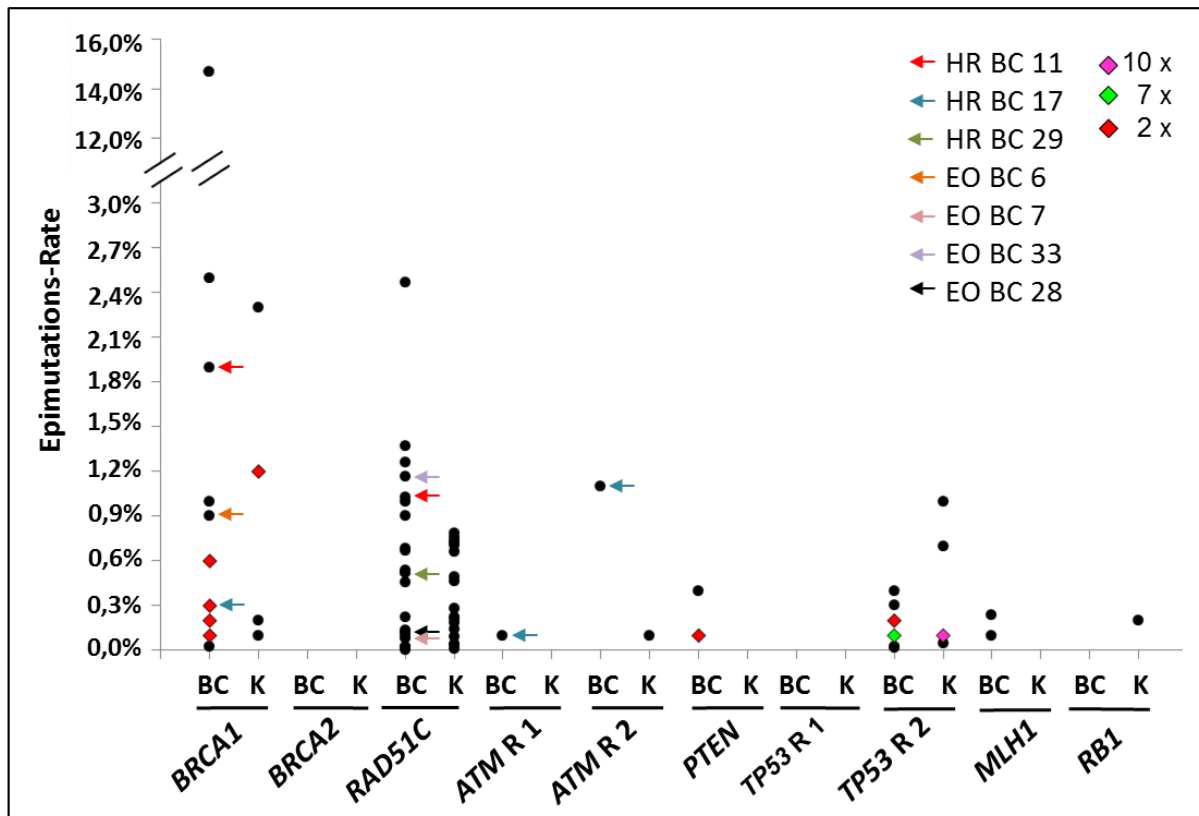


Abbildung 37: Verteilung der Epimutations-Raten (EMR) in Prozent je Amplicon und Gruppe. Die Kreise stellen die Patienten- und Kontrollproben (kombinierte Kohorten je Gruppe) dar, die Epimutationen aufgewiesen haben. Die farbigen Rauten stehen für EMR, die bei mehreren Proben den selben Wert gezeigt haben. Die EMR berechnet sich aus der gesamten Read-Anzahl geteilt durch die Anzahl der aberrant methylierten Reads. Die meisten Epimutationen wurden in *BRCA1* und *RAD51C* detektiert. Der Großteil der Proben zeigt eine EMR kleiner 1%. Eine HR BC Probe zeigte eine EMR >14% in *BRCA1*. Die Pfeile zeigen die Proben an, die zum Zeitpunkt der Blutabnahme nicht tumorfrei waren.

Abgesehen von dem Extremwert, konnten nicht nur die meisten Proben mit Epimutationen in *BRCA1* und *RAD51C* detektiert werden, sondern auch die Proben mit den höchsten EMR. Es konnte ein signifikanter Gruppenunterschied zwischen Patienten und Kontrollen (kombinierte Kohorten) für die durchschnittliche EMR in *BRCA1* ($p=0,03$; Mann-Whitney-U) und eine leichte Tendenz in *PTEN* ($p=0,08$; Mann-Whitney-U) gemessen werden. Der Gruppenunterschied in *BRCA1* bleibt auch nach Ausschluss des Extremwertes erhalten ($p=0,02$; Mann-Whitney-U). Tabelle 27 zeigt die durchschnittlichen EMR mit

Standartabweichung und den minimalen und maximalen EMR-Wert je Amplikon und Kohorte.

Tabelle 27: Durchschnittliche Epimutations-Raten (EMR) mit Standartabweichung (SD) und minimalem bzw. maximalem EMR Wert je Amplikon und Kohorte.

Durchschnittliche Epimutations-Rate ± SD in % (Min. – Max. in %)					
	EO BC	Kontrolle	HR BC ^c	Kontrolle	p-Wert ^{a,b}
BRCA1	0,086±0,037 (0-1,0)	0,114±0,067 (0-2,3)	0,567±0,413 (0-14,7)	0,006±0,006 (0-0,2)	0,03 (0,02) ^c
BRCA2	N/A	N/A	0	0	
RAD51C	0,162±0,053 (0-1,4)	0,088±0,028 (0-0,8)	0,128±0,051 (0-1,2)	0,065±0,035 (0-0,8)	
ATM Region 1	0	0	0,003±0,003 (0-0,1)	0	
ATM Region 2	0,001±0,001 (0-0,1)	0	0,031±0,031 (0-1,1)	0,003±0,003 (0-0,1)	
PTEN	0	0	0,017±0,012 (0-0,4)	0	0,08
TP53 Region 1	0	0	0	0	
TP53 Region 2	0,007±0,004 (0-0,1)	0,011±0,005 (0-0,1)	0,040±0,016 (0-0,4)	0,065±0,033 (0-1,0)	
MLH1	0	0	0,003±0,003 (0-0,3)	0	
RB1	0	0	0	0,006±0,006 (0-0,2)	

^a Mann-Whitney-U; Signifikanzniveau ist 0,05.

^b Gruppenvergleich BC und Kontrollen (kombinierte Kohorten).

^c durchschnittliche EMR für HR BC in *BRCA1* unter Ausschluss des Extremwertes: 0,163±0,089 (0-2,5).

Alle Proben, mit oder ohne Epimutation, zeigen das gleiche Verteilungsprofil an unterschiedlich stark methylierten Allelen, bei einer Einteilung der Methylierung einzelner Allele in 10% Intervallen. Proben ohne Epimutationen zeigen ausschließlich unmethylierte bzw. niedrig methylierte Allele von 0%- 29%. Individuen mit aberranten Allelen weisen ebenfalls hauptsächlich unmethylierte oder schwach methylierte Allele von 0% -29% auf, aber auch Allele mit einer Methylierung von 30% - 100% (Abbildung 38).

Die Amplikons von *BRCA1*, *RAD51C*, *ATM* Region 1 und 2, *MLH1* und *RB1* enthalten jeweils einen SNP mit einer hohen MAF (> 0,1), über den die Allele bei heterozygoten Proben nach ihrer elterlichen Herkunft unterschieden werden können. In 14 heterozygoten *BRCA1* und 10

RAD51C Proben konnten alle aberrant methylierten Allele nur einem Elternteil zugeordnet werden, da diese an der SNP Position alle die gleiche Base trugen. Im Gegensatz dazu befanden sich die Epimutationen in den Amplikons von *ATM* Region 1 und 2, *MLH1* und *RB1* über beide parentalen Allele verteilt.

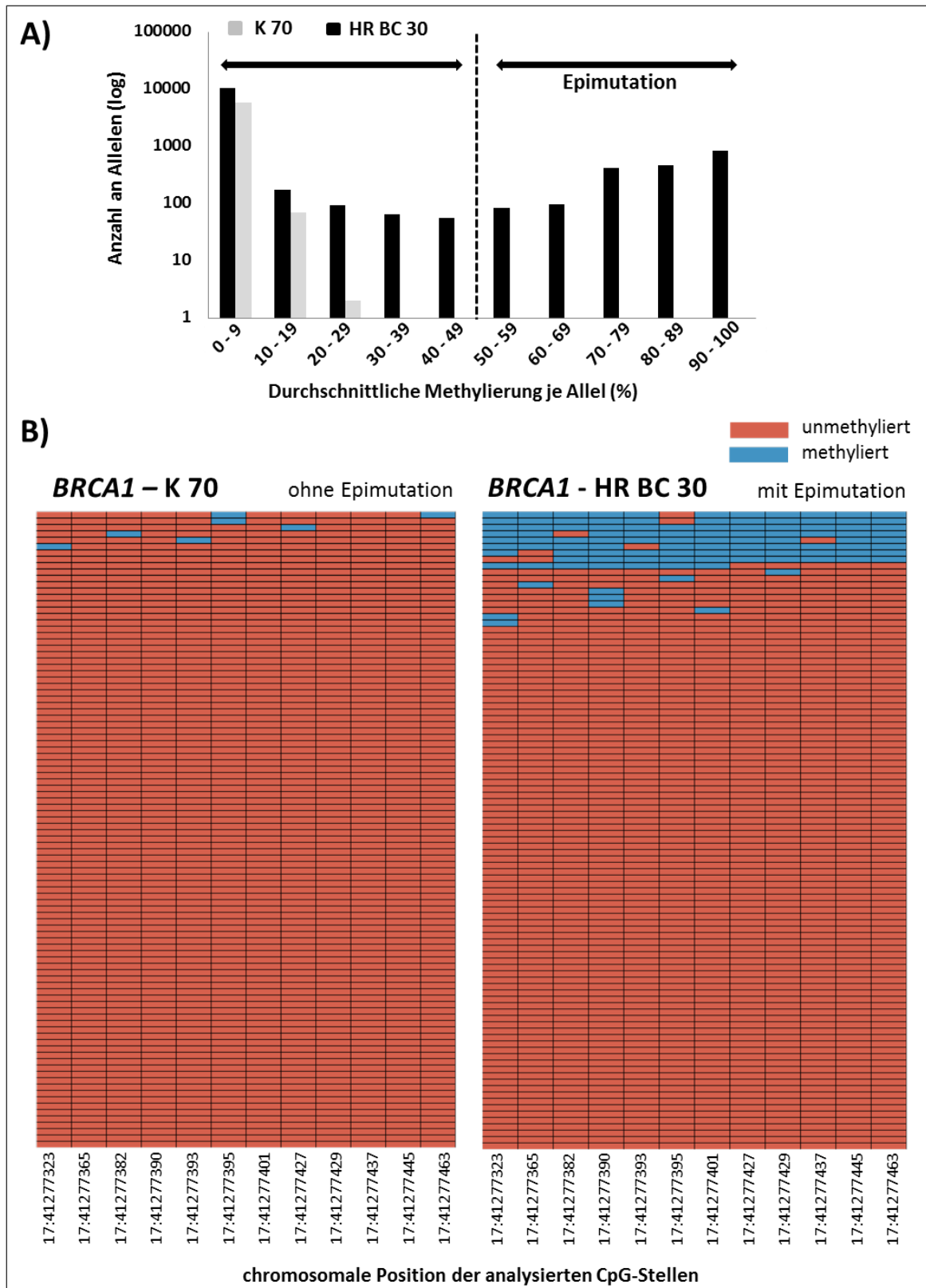


Abbildung 38: Methylierungsprofil auf Einzel-Allel-Ebene von Proben mit und ohne Epimutation.
A) Balkendiagramm zur Darstellung allelischer Methylierungsprofile einer Probe mit (HR BC 30) und ohne Epimutationen (K 70) in 10% Methylierungsintervallen. **B)** „Heatmaps“ für je eine exemplarische Probe ohne (links, K70) und mit (rechts, HR BC30) Epimutation. Dargestellt ist jeweils ein Ausschnitt aus dem analysierten Amplikon (17: 41277323 - 41277463) von *BRCA1*, sowie 100 Reads/Allele je Probe. Die Auswahl der Allele erfolgte nach dem Zufallsprinzip. Die einzelnen Zeilen

stehen jeweils für ein einzelnes Allel und die Spalten für eine CpG-Stelle. Die Kästchen sind rot bzw. blau eingefärbt wenn die CpG-Stelle unmethyliert bzw. methyliert vorliegt.

2.1.3 Hypermethylierung auf Einzel-CpG-Ebene

Zusätzlich zum Epimutations-Screening wurden die Proben auf Einzel-CpG-Methylierungsfehler hin untersucht. Zur Unterscheidung zwischen normal und abnormal methylierten CpG-Stellen wurde für jedes einzelne CpG ein individueller Schwellenwert berechnet. Extremwerte definieren sich in der Statistik über 3 Mal den Interquartilbereich (IQB) plus den Wert der 75% Quartile. In dieser Studie lässt sich diese Definition für Extremwerte nicht anwenden, da die IQBs sehr gering sind, wie in Tabelle 25 dargestellt. Um dennoch einen Schwellenwert zu definieren, der es möglich macht, stochastisches Rauschen von potentiell relevanten aberrant methylierten CpG-Stellen unterscheiden zu können, wurde der Schwellenwert auf 5 Mal den IQB plus den Wert der 75% Quartile erweitert. Für die Berechnung wurden ausschließlich Kontrollproben ohne Epimutationen verwendet.

Tabelle 28 gibt einen Überblick über die Anzahl an Proben, bei denen sogenannte „single CpG errors“ detektiert werden konnten. „Single CpG methylation errors“ beschreiben Allele, die ein oder zwei aberrant methylierte CpG-Stellen aufweisen, wohingegen alle umliegenden CpGs unmethyliert vorliegen (Abbildung 39, B).

Am häufigsten traten Einzel-CpG-Hypermethylierungen in der EO BC Kohorte in den Promotorbereichen von *PTEN* (39%), *MLH1* (23%), *TP53* Region 1 (22%) und 2 (17%), sowie *RB1* (17%) auf. Des Weiteren findet sich in diesen Bereichen ein signifikanter Gruppenunterschied zwischen BC-Patientinnen und den Kontrollproben (kombinierte Kohorten). Über alle untersuchten Proben und Amplikons gesehen, weisen signifikant ($p=0,001$; Fisher exakt) mehr EO BC (60/389; 15%) Patientinnen im Vergleich zu den HR BC (8/363; 2%) Patientinnen aberrant methylierte CpG-Stellen auf.

Tabelle 28: Anzahl an Einzel-CpG-Hypermethylierungen in BC- und Kontrollproben je Kohorte und Amplikon.

Anzahl der Proben [*] mit Einzel-CpG-Hypermethylierung (in Prozent)					
	EO BC	Kontrolle	HR BC	Kontrolle	p-Wert ^{a,b}
BRCA1	1/35 (3%)	0/38 (0%)	0/28 (0%)	0/35 (0%)	
BRCA2	N/A	N/A	2/36 (6%)	1/36 (3%)	
RAD51C	1/29 (4%)	0/31 (0%)	0/27 (0%)	0/29 (0%)	
ATM region 1	7/43 (16%)	2/44 (5%)	0/35 (0%)	1/36 (3%)	
ATM region 2	4/42 (10%)	0/44 (0%)	0/35 (0%)	0/35 (0%)	
PTEN	12/31 (39%)	1/32 (3%)	1/33 (3%)	0/36 (0%)	0,001
TP53 region 1	9/41 (22%)	1/42 (2%)	1/36 (3%)	0/36 (0%)	0,005
TP53 region 2	7/41 (17%)	0/39 (0%)	0/27 (0%)	0/27 (0%)	0,015
MLH1	10/43 (23%)	3/44 (7%)	1/34 (3%)	0/36 (0%)	0,025
RB1	7/42 (17%)	1/42 (2%)	1/36 (3%)	0/35 (0%)	0,017
ESR1	2/42 (5%)	0/42 (0%)	2/36 (6%)	1/36 (3%)	

^a Fisher-Test; Signifikanzniveau ist 0,05.

^b Gruppenvergleich BC und Kontrollen (kombinierte Kohorten).

* Anzahl unter Ausschluss aller Proben mit Epimutationen.

In Abbildung 39 ist exemplarisch das Methylierungsprofil für eine HR BC Probe mit „single CpG errors“ im Promoterbereich von *MLH1* aufgezeigt. Diese Probe weist an 10 von 33 analysierten CpG-Stellen (Position 9, 10, 14, 15, 18, 19, 23, 29, 31, 33), im Vergleich zum Schwellenwert, eine Einzel-CpG-Hypermethylierung auf (über alle Reads berechnete) (Abbildung 39, A). Auf Einzel-Allel-Ebene erkennt man deutlich, dass pro Allel/Read immer nur ein bis maximal zwei CpG-Stellen aberrant methyliert und zufällig über das Allel verteilt sind (Abbildung 39, B).

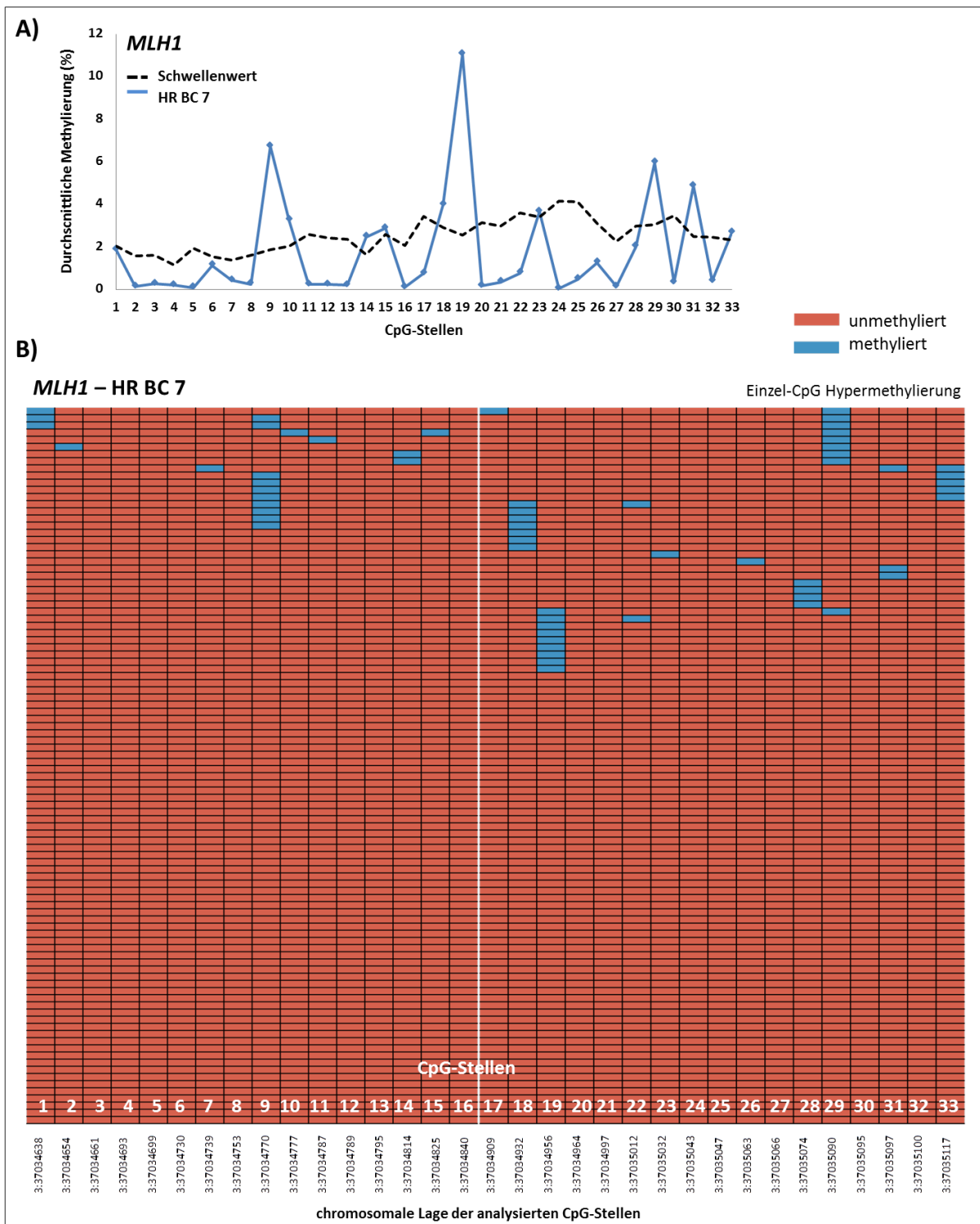
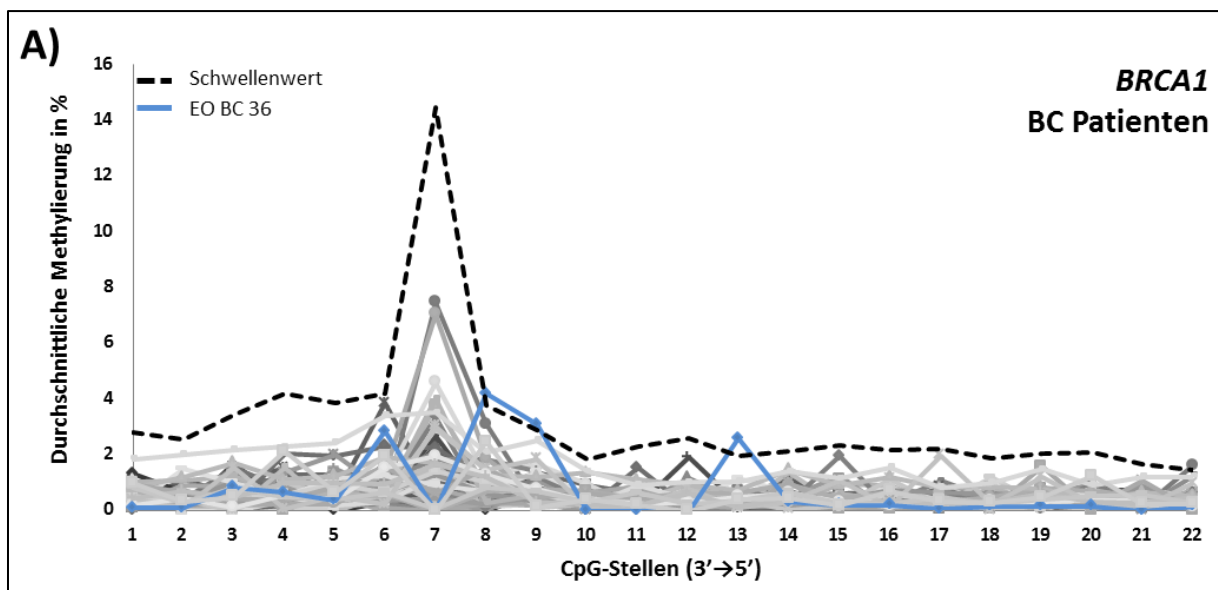
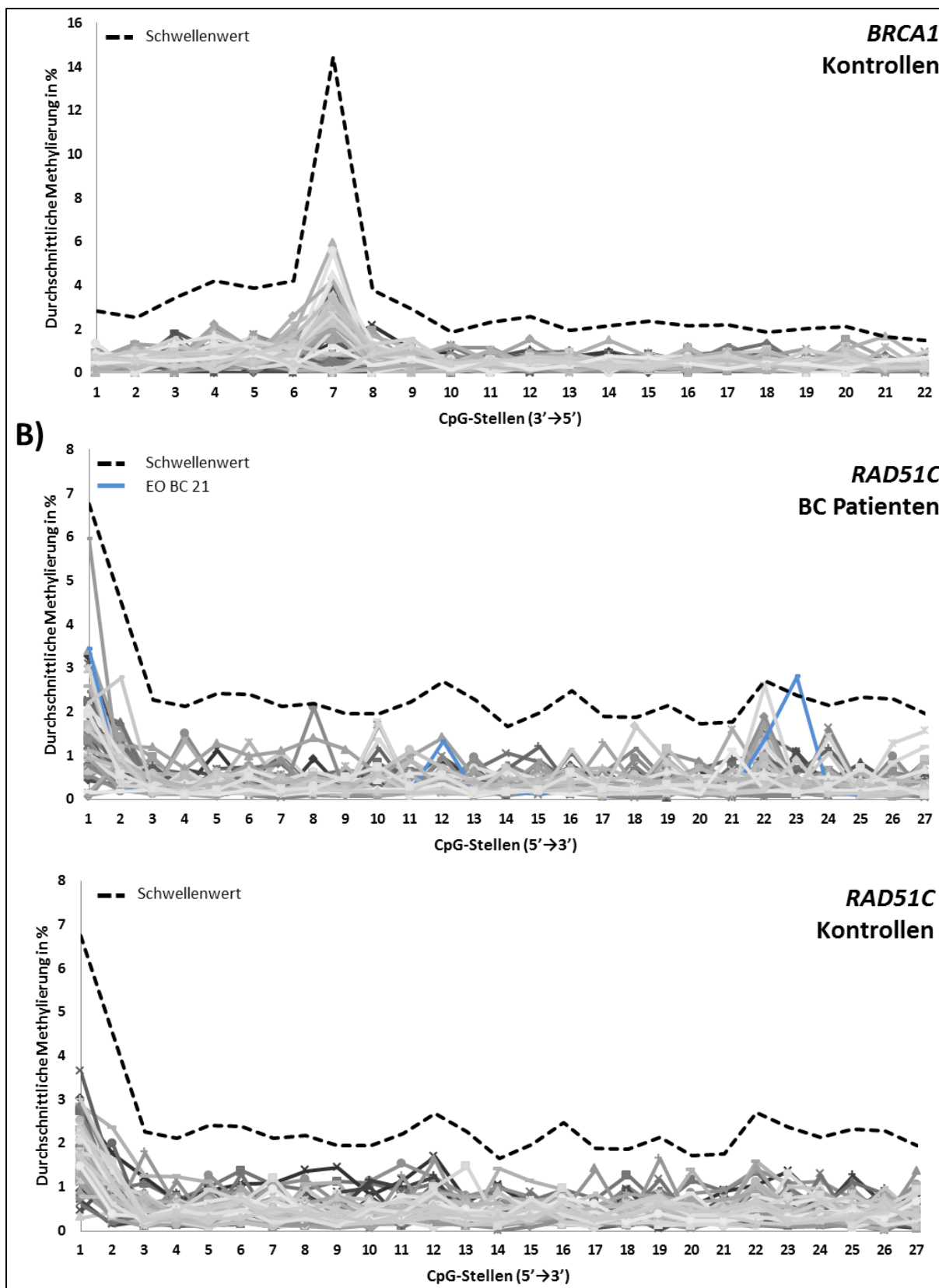


Abbildung 39: Methylierungsprofil, exemplarisch dargestellt für eine HR BC Probe mit „single CpG errors“ im Promoterbereich von *MLH1*. **A)** Die durchschnittliche Methylierung je CpG-Stelle des *MLH1* Amplikons für die Probe HR BC 7 ist als blaue Linie und der Schwellenwert (5*IQR + 75% Quartil Wert) als gestrichelte schwarze Linie gezeigt. Diese Patientenprobe weist an 10 von 33 analysierten CpG-Stellen (Pos. 9, 10, 14, 15, 18, 19, 23, 29, 31, 33), im Vergleich zum Schwellenwert, eine Einzel-CpG-Hypermethylierung auf. **B)** Die sogenannte „Heatmap“ zeigt die bereits unter **A)** dargestellte Probe HR BC 7 auf Einzel-Allel-Ebene. Dargestellt sind repräsentativ 100 Reads/Allele. Die Auswahl der Allele erfolgte nach dem Zufallsprinzip. Die einzelnen Zeilen stehen jeweils für ein

einzelnes Allel und die Spalten für eine CpG-Stelle. Die Kästchen sind rot bzw. blau eingefärbt, wenn die CpG-Stelle unmethyliert bzw. methyliert vorliegt.

Auffällig war des Weiteren, dass in den Bereichen von *BRCA1* (Abbildung 40, A) und *RAD51C* (Abbildung 40, B) jeweils nur eine EO Patienten Probe und keine Kontrollprobe mit Einzel-CpG-Hypermethylierung gefunden werden konnte; diese Gene/Amplikons zeigten neben *TP53* Region 2 die meisten Proben, sowohl BC als auch Kontrollen, mit Epimutationen (Tabelle 26). In *TP53* Region 2 dagegen wiesen zudem rund ein Fünftel aller analysierten EO BC Proben „single CpG errors“ auf (Abbildung 40, C). Abbildungen zu den Methylierungsprofilen der restlichen untersuchten Amplikons (*BRCA2*, *ATM* Region 1/2, *TP53* Region 1, *MLH1*, *RB1* und *ESR1*) befinden sich in einer separaten Abbildung im Anschluss (Abbildung 41). In Summe wurden in 362 analysierten CpG-Stellen 119 CpGs (33%) mit einer aberranten Hypermethylierung in den BC-Proben und 13 (4%) aberrante CpGs in den Kontrollen detektiert.





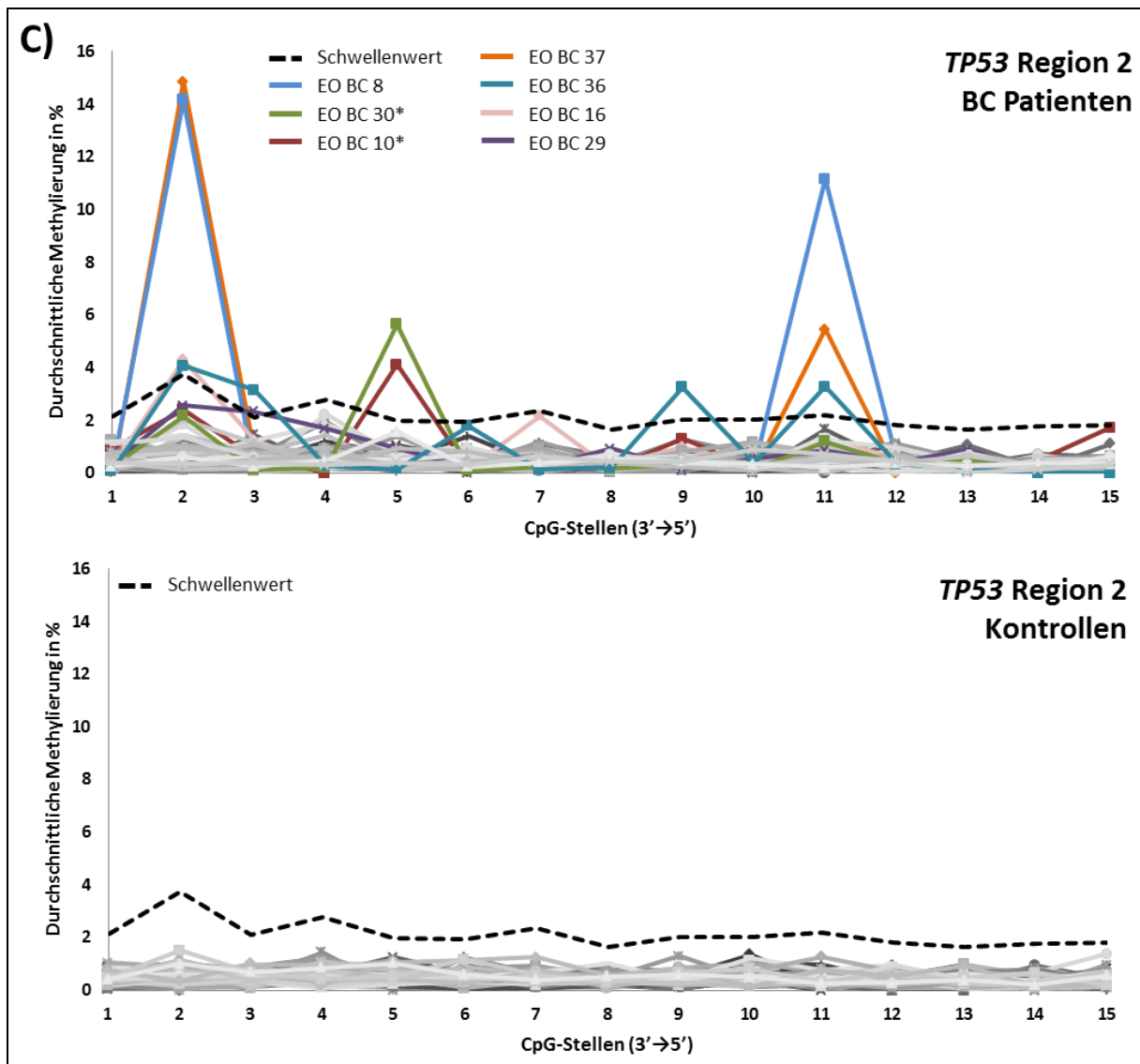
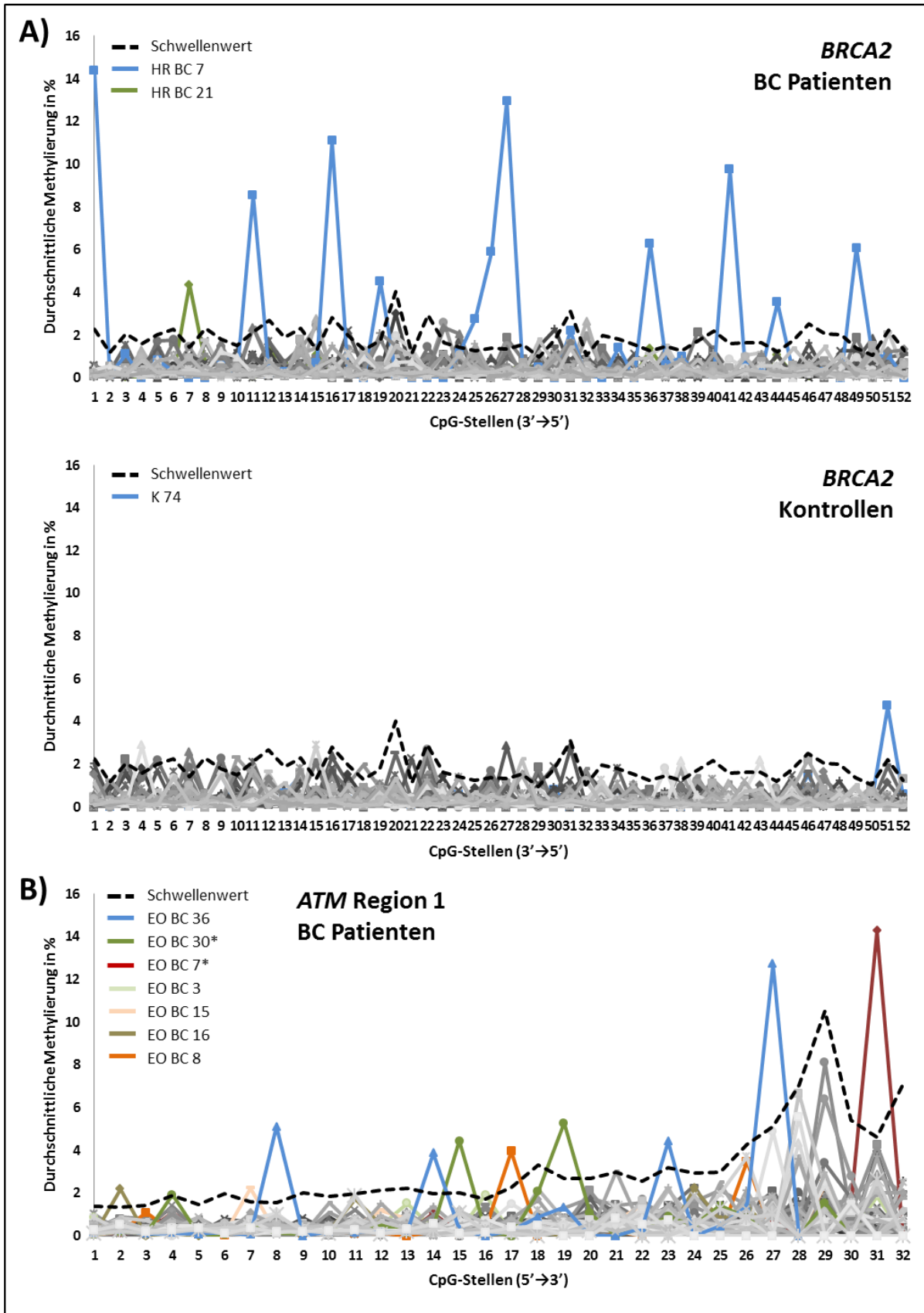
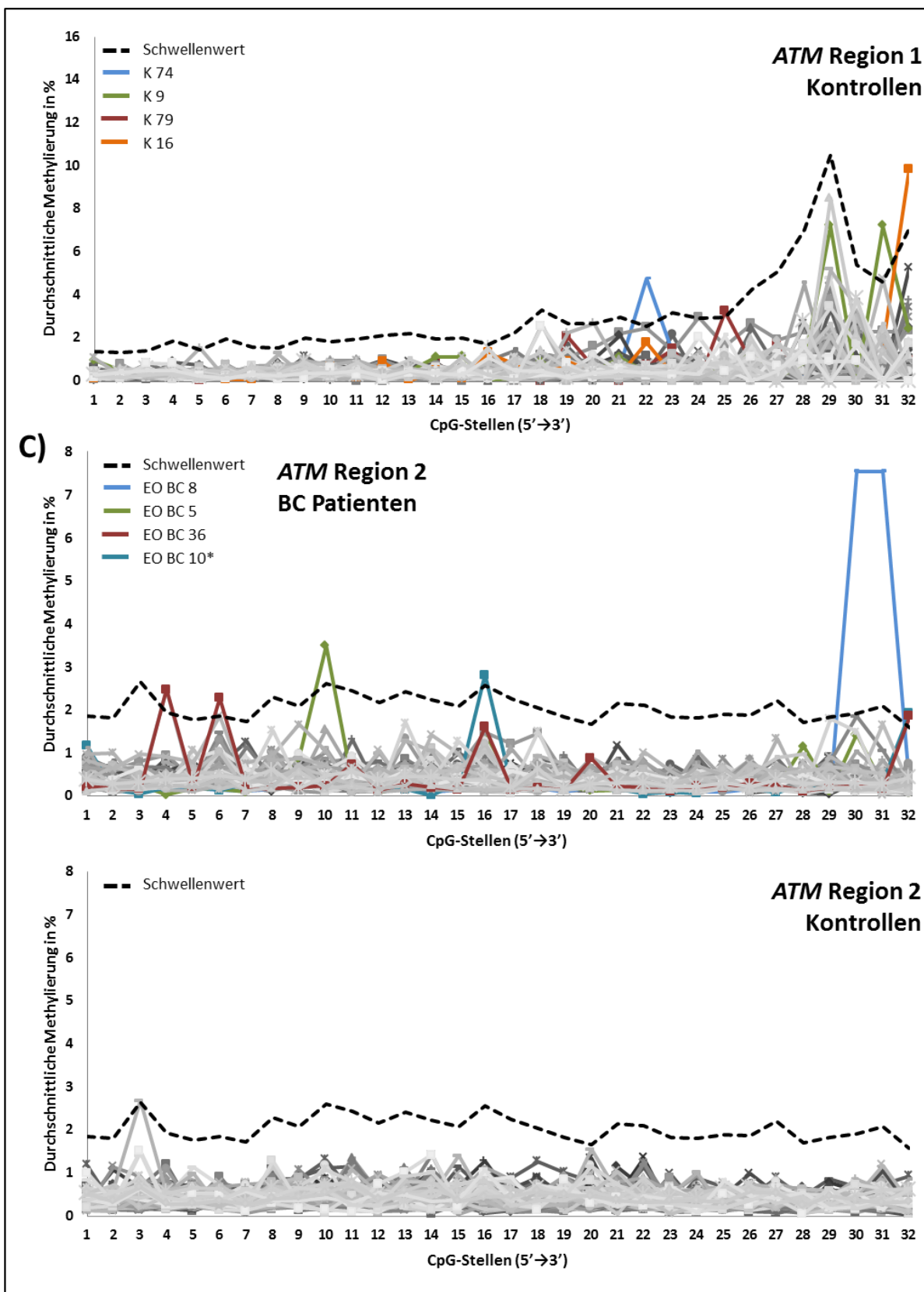
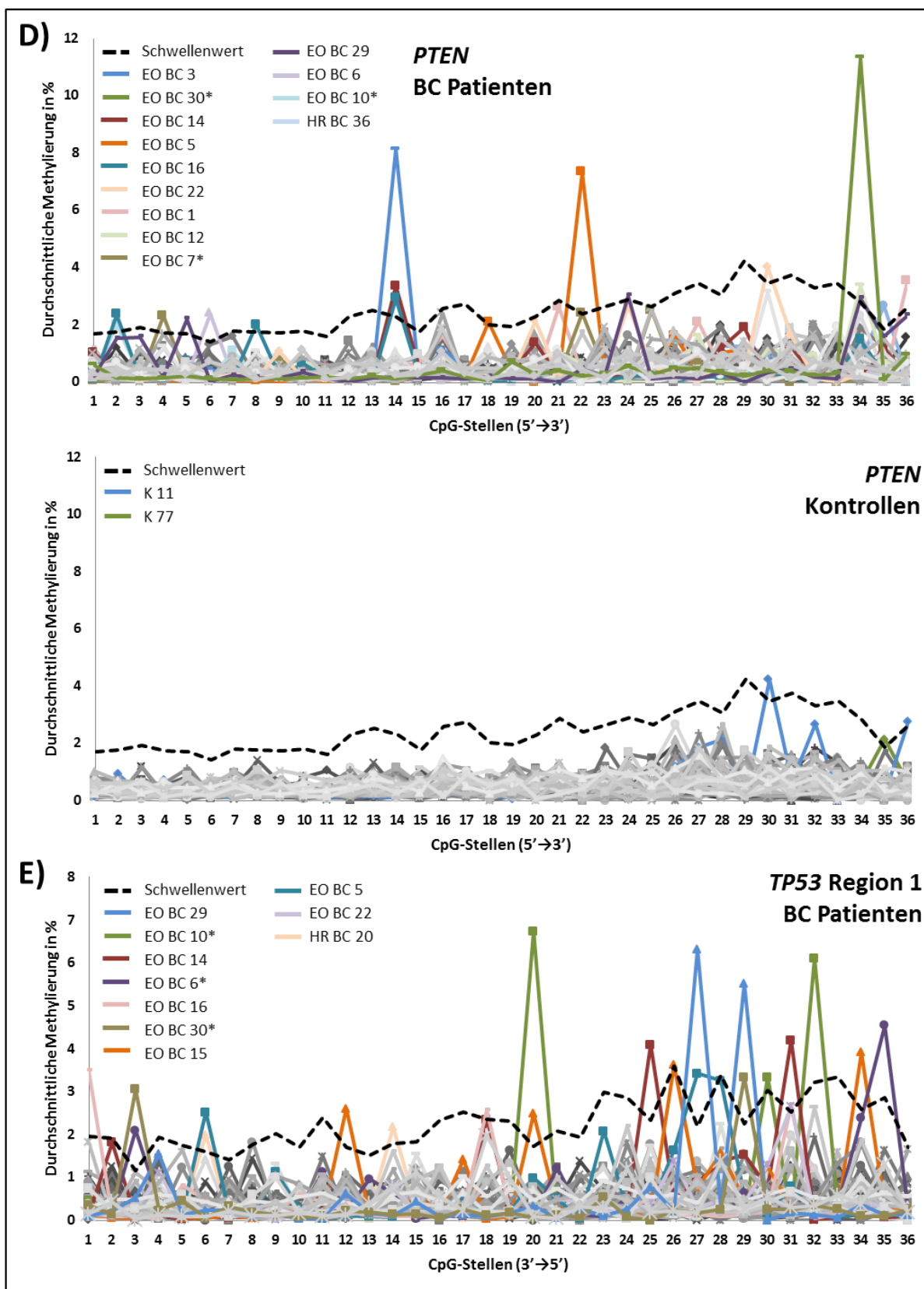
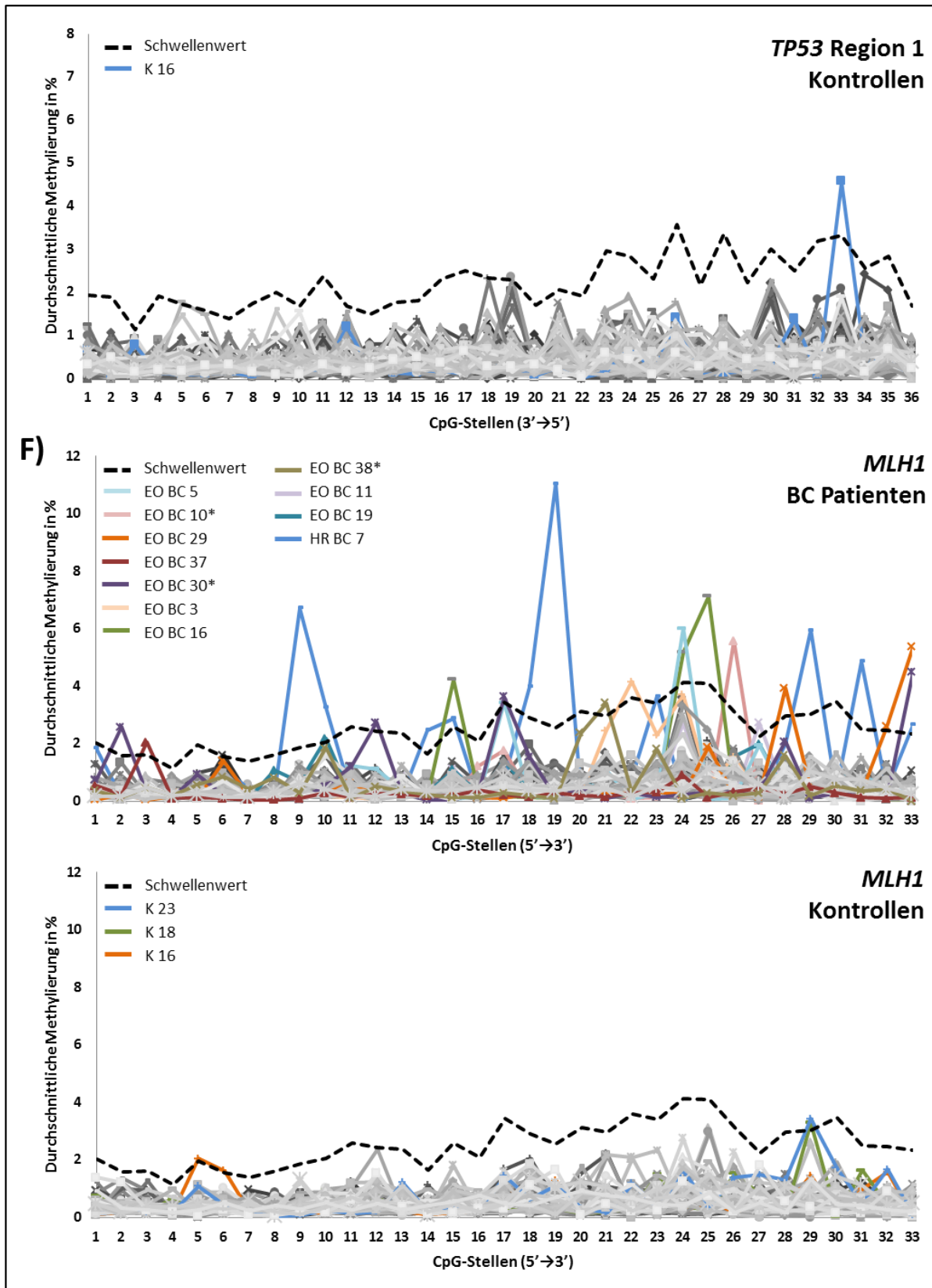


Abbildung 40: Durchschnittliche Methylierung je CpG-Stelle für *BRCA1*, *RAD51C* und *TP53* Region 2. In Grau sind alle Proben dargestellt, die unterhalb des Schwellenwertes liegen, und somit eine normale durchschnittliche Methylierung aufweisen. Die gestrichelte schwarze Linie stellt den Schwellenwert (5*IQR + 75% Quartil Wert) dar. Alle Proben, die mindestens eine aberrant hypermethylierte CpG-Stelle besitzen, sind unterschiedlich farbig dargestellt. Die mit Stern markierten Proben waren zum Zeitpunkt der Blutabnahme nicht tumorfrei.









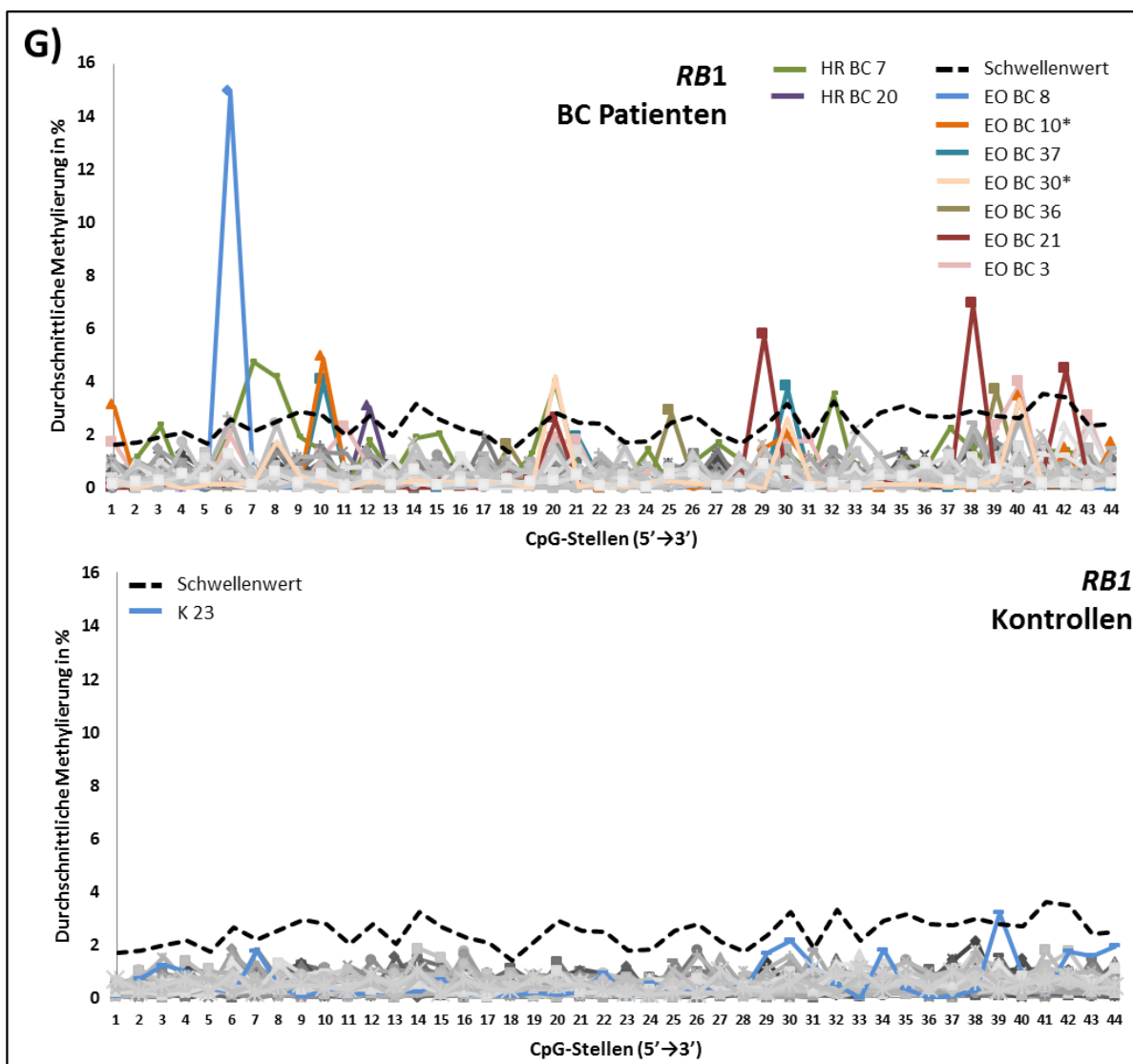


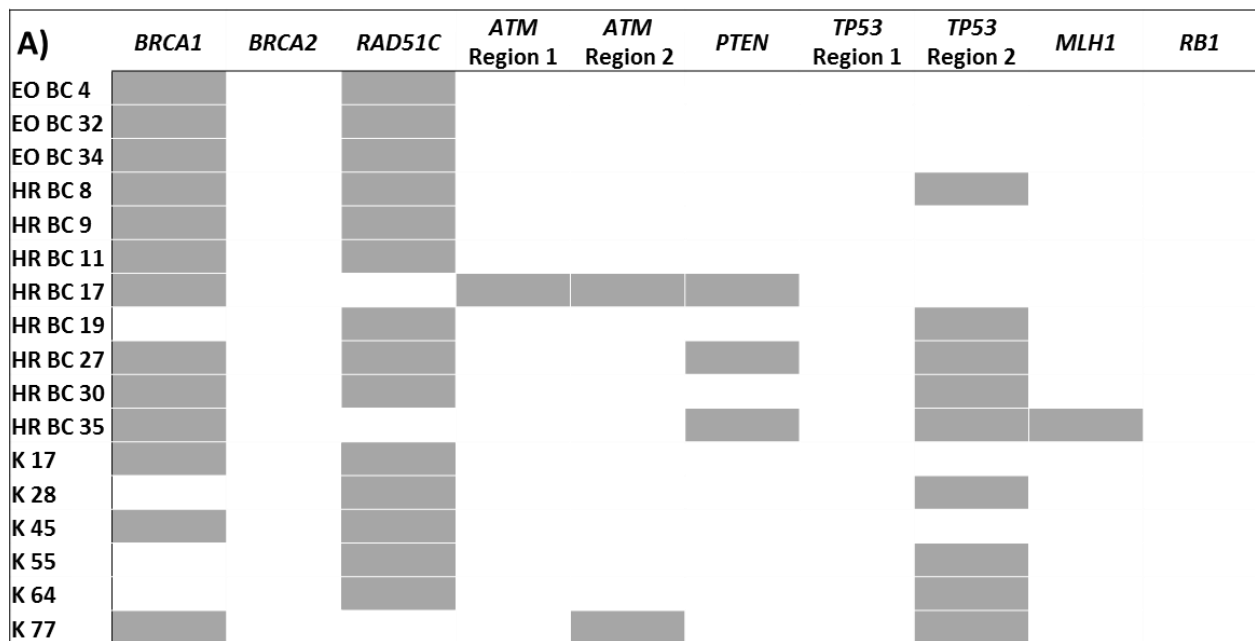
Abbildung 41: Durchschnittliche Methylierung je CpG-Stelle für *BRCA2*, *ATM* Region 1 und 2, *PTEN*, *TP53* Region 1, *MLH1* und *RB1*. In Grau sind alle Proben dargestellt, die unterhalb des Schwellenwertes liegen, und somit eine normale durchschnittliche Methylierung aufweisen. Die gestrichelte schwarze Linie stellt den Schwellenwert (5*IQR + 75% Quartil Wert) dar. Alle Proben, die mindestens eine aberrant hypermethylierte CpG-Stelle besitzen, sind unterschiedlich farbig dargestellt. Die mit Stern markierten Proben waren zum Zeitpunkt der Blutabnahme nicht tumorfrei. **A)** *BRCA2*. **B)** *ATM* Region 1. **C)** *ATM* Region 2. **D)** *PTEN*. **E)** *TP53* Region 1. **F)** *MLH1*. **G)** *RB1*.

2.1.4 Kummulative Promotor-Hypermethylierung

Vergleicht man alle Gene/Amplikons je Probe, so fällt auf, dass in einigen Proben mehrere Gene/Amplikons ein aberrantes Methylierungsmuster (Epimutationen oder „single CpG error“) aufweisen. In der HR BC Kohorte konnten aberrante Allele in mehreren Promotorbereichen für 50% (8/16), in der EO BC Kohorte bei 14% (3/21) und in den Kontrollen bei 18% (6/34) aller hypermethylierten Proben gemessen werden. Dabei war

auffällig, dass 82% (14/17) und 77% (13/17) aller analysierter Proben (BC und Kontrollen), die aberrant methylierte Allele in *RAD51C* bzw. *BRCA1* aufgezeigt haben, in mindestens einem weiteren Amplikon Epimutationen trugen. Einige Proben zeigten auch multiple Hits in *TP53* Region 2 (53%; 9/17), *PTEN* (18%; 3/17), *ATM* Region 2 (11.8%, 2/17), *ATM* Region 1 und *MLH1* (5.9%; 1/17) (Abbildung 42, A).

Eine kummulative Anhäufung von „single CpG errors“ in einzelnen Proben traten in 50% (2/4), 75% (15/20), und 38% (3/8) der HR BC, EO BC-Patientinnen und Kontrollen auf. Dabei war die Promotorregionen von *PTEN* und *TP53* Region 1 zu 55% (11/20), *MLH1* zu 50% (10/20), *ATM* Region 1 und *RB1* zu 45% (9/20), *TP53* Region 2 zu 35% (7/20), *ATM* Region 2 zu 20% (4/20), *ESR1* und *BRCA2* zu 10% (2/20) und *BRCA1* zu 5% (1/20) betroffen (Abbildung 42, B).



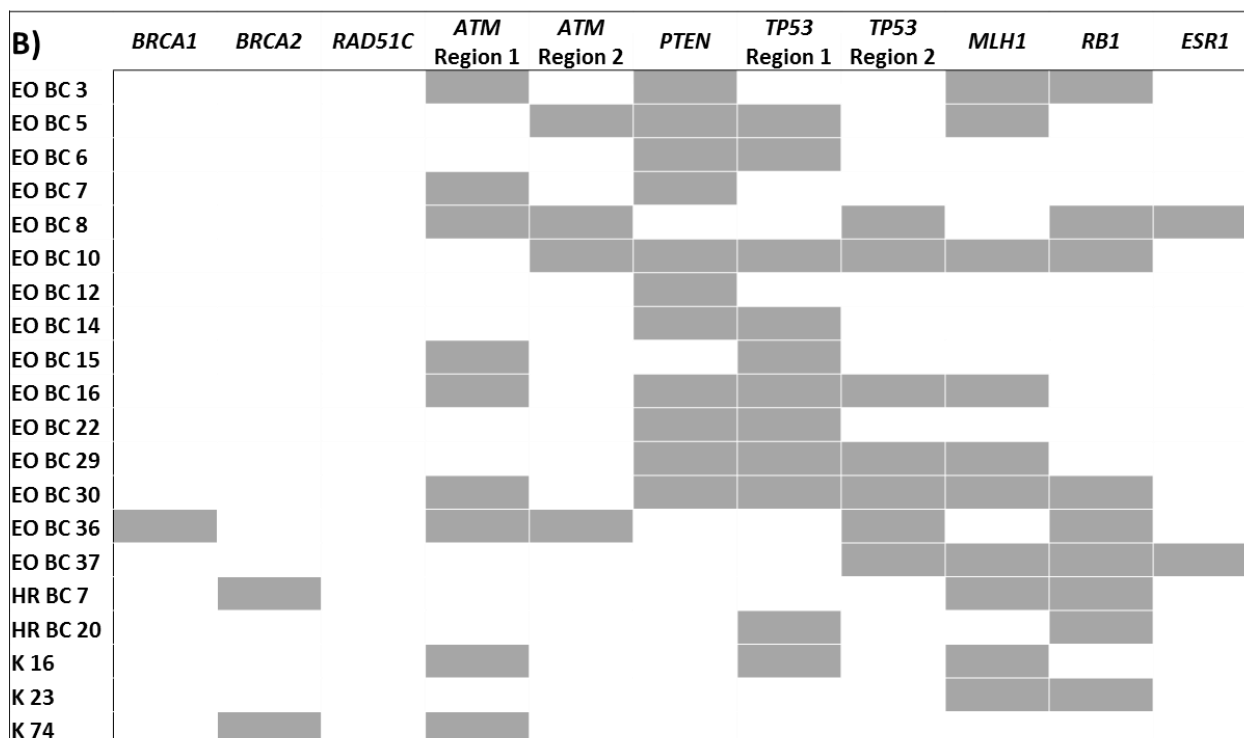


Abbildung 42: Kummulative Hypermethylierung in Proben mit Epimutationen bzw. Einzel-CpG-Fehlern. A) Multiple Hits in Proben mit aberrant methylierten Allelen. **B)** Multiple Hits in Proben mit Einzel-CpG-Hypermethylierung.

Führt man eine Korrelationsanalyse für die Anzahl an Hits je Probe und dem Alter der Patientinnen bei der Brustkrebsdiagnose (21 – 85 Jahre) durch, so findet sich eine signifikant negative Korrelation (Spearman rho = - 0,184; p = 0,018) zwischen der Anzahl an Genen pro Probe, die „single CpG error“ aufweisen, und dem Krebsbeginn (Abbildung 43). Es gab keine vergleichbare Korrelation für die Anzahl an Genen mit Epimutationen und dem Alter.

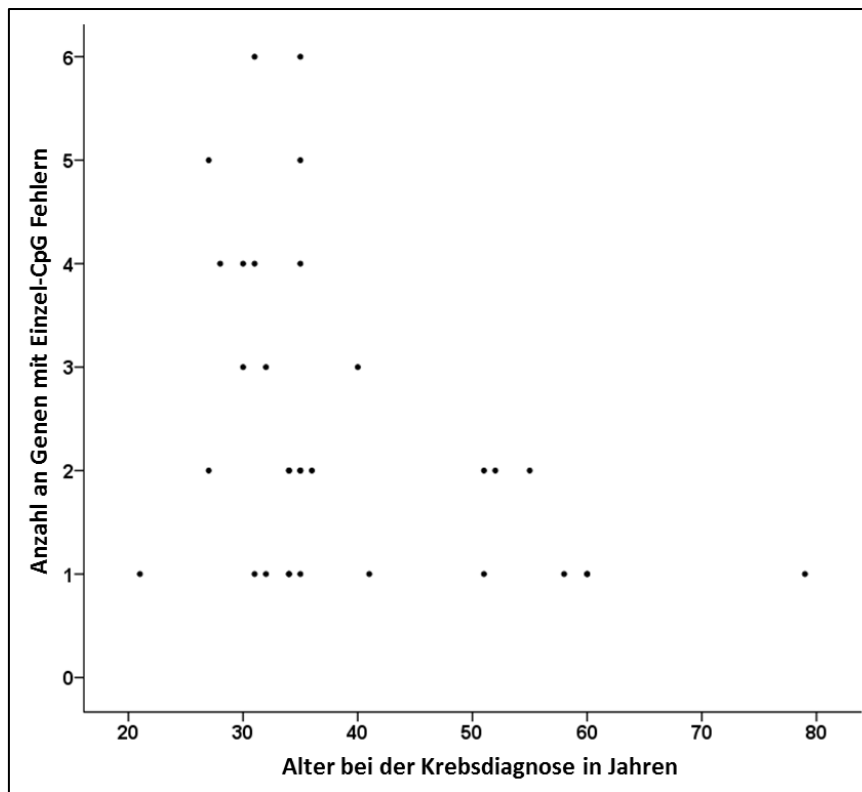


Abbildung 43: Beziehung zwischen der Anzahl an Genen mit single CpG errors einer Probe und ihrem Alter bei der Krebsdiagnose. Es besteht eine signifikant negative Korrelation mit einem Koeffizienten von Spearman $\rho = -0,184$ ($p = 0,018$) zwischen der Anzahl an Bereichen mit Einzel-CpG-Hypermethylierung einer Probe und deren Alter bei der Krebsdiagnose.

2.2 Expressionsanalysen für ausgewählte Proben und Gene

Um den Einfluss der aberranten Promotormethylierung auf die Expression des jeweiligen TSG zu prüfen, wurden Expressionsanalysen mittels TaqMan Technologie durchgeführt. Wie unter III 1.1.1 bereits beschrieben, handelte es sich, bei den untersuchten Blut DNA-Proben der EO und HR BC Proben für die Methylierungsanalysen, um Ausschussmaterial nach einer diagnostischen Untersuchung. Deshalb war es nicht möglich, für die RNA-Analyse alle Proben zu untersuchen. Für die Expressionsanalyse standen eine EO BC Probe ohne detektierte Methylierungsveränderungen (EO BC 40), eine EO BC und drei HR BC Proben mit niedrigen bis intermediären EMR (EO BC 34, HR BC 8, HR BC 11, EO BC 17), zwei EO BC Proben mit „single CpG errors“ (EO BC 36, EO BC 30) und drei gesunde Kontrollproben (K 90, K 91, K92 = Referenzprobe) zur Verfügung. Bezogen auf die relativen Expressionswerte der Gene innerhalb der Kontrollproben wurde der Schwellenwert zur Unterscheidung zwischen normaler und niedriger bzw. erhöhter Expression zwischen 0,5 und -0,5 festgesetzt. Ein \log_2

RQ Wert von 1 bzw. -1 entspricht einer Verdoppelung bzw. Halbierung der Expression bezogen auf die Referenzprobe (Abbildung 44).

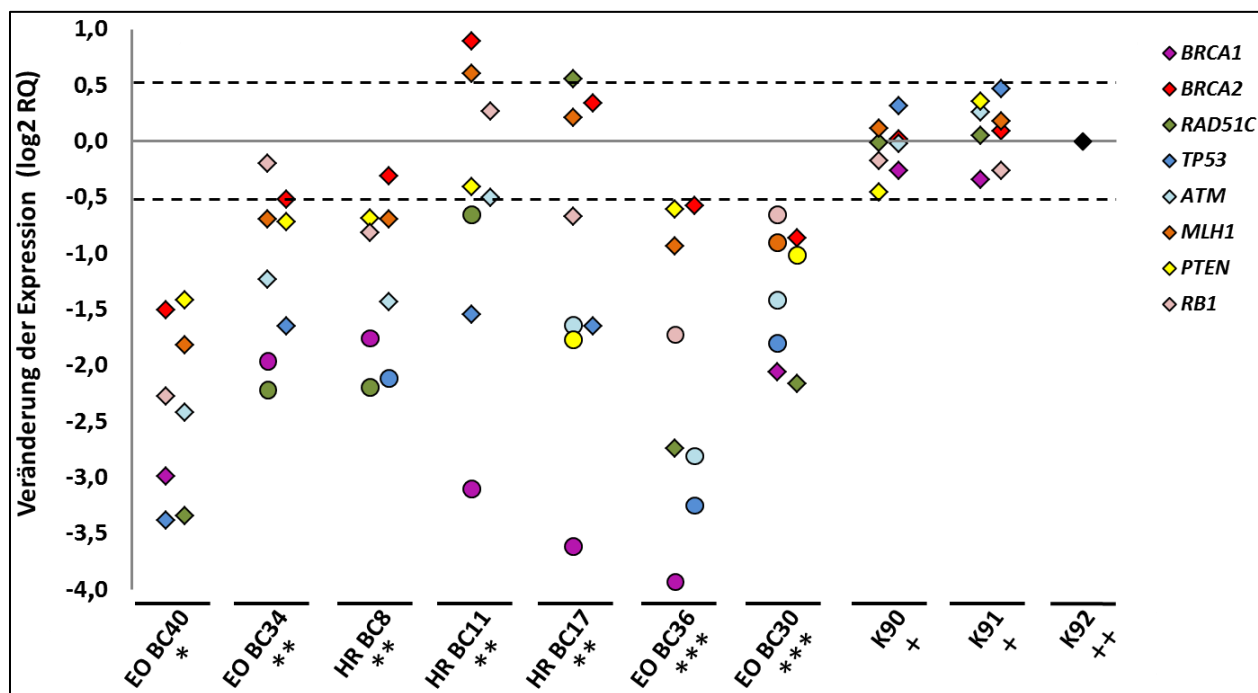


Abbildung 44: Ergebnisse der Expressionsanalyse in den Tumorsuppressorgenen *BRCA1*, *BRCA2*, *RAD51C*, *TP53*, *ATM*, *MLH1*, *PTEN* und *RB1* bei ausgewählten Brustkrebspatientenproben sowie gesunden Kontrollen. Die relativen Expressionswerte werden durch die farbigen Rauten bzw. Kreise dargestellt. Die Kreise zeigen zudem an, dass in diesem Gen für diese Probe eine aberrante Promotormethylierung via DBS detektiert wurde. Die Proben können in fünf Gruppen unterteilt werden: * Patient ohne Methylierungsveränderung, ** Patient mit Epimutationen, *** Patient mit „single CpG errors“, + gesunde Kontrolle, ++ gesunde Kontrolle, welche als Kalibrator fungiert. Die gestrichelten Linien geben den Schwellenwert des Normalbereichs bezogen auf die Kontrollproben an. Ein log2 RQ Wert von 1 bzw. -1 entspricht einer Verdoppelung bzw. Halbierung der Expression bezogen auf die Referenzprobe (K 92).

In allen untersuchten Patientenproben konnte eine verminderte Expression einiger oder aller TSG im Vergleich zu den gesunden Kontrollproben gemessen werden. Die Patientenproben (EO BC 40) ohne veränderte Methylierung und die Proben mit Einzel-CpG-Fehlern (EO BC 36, EO BC 30) zeigten eine herabgesetzte Expression aller TSG. Bei den Patienten mit Epimutationen (EO BC 34, HR BC 8, HR BC 11, EO BC 17) waren die meisten Gene ebenfalls herunterreguliert, jedoch zeigten einige Gene zudem eine normale oder sogar eine erhöhte Expression bezogen auf die Kontrollproben. Die stärkste Herabregulation konnte in allen Patienten in den TSG *BRCA1*, *RAD51C*, *TP53* und *ATM* gemessen werden.

Zudem waren die Gene mit Hypermethylierung am niedrigsten exprimiert bezogen auf einen Patienten.

V DISKUSSION

1. Identifizierung differenziell methylierter Regionen im präfrontalen Cortex

Die Unterschiede in der proteincodierenden Sequenz zwischen dem Menschen und seinen nächsten Verwandten, den Primaten, sind zu gering, als dass diese die verschiedenen Phänotypen bedingen könnten. Deshalb entstand die Hypothese, dass diese Unterschiede eher durch Veränderungen auf Ebene der Genregulation erklärt werden könnten als durch Veränderungen in der primären genomischen Sequenz. Die DNA-Methylierung ist eine wichtige epigenetische Modifikation und ist bekanntermaßen in viele biologische Prozesse involviert; ihr Einfluss auf die Evolution des menschlichen Gehirns ist noch nicht besonders gut verstanden. Jedoch haben verschiedene Studien gezeigt, dass die DNA-Methylierung eine wichtige Rolle bei der Gehirnentwicklung und Regulation der Transkription beim Menschen einnimmt, mit Auswirkungen auf die Funktion der Neuronen und die Entstehung von Krankheiten. Die Differenzierung neuronaler und nicht-neuronaler Zellen [269-271], die axonale Wegfindung [272] und auch die Remodellierung von Synapsen sind mit der DNA-Methylierung assoziiert [272-274]. Des Weiteren ist die DNA-Methylierung an der Gedächtnisbildung beteiligt [274]. Im Laufe der letzten Jahre wurden einige Studien durchgeführt, die sich mit der differentiellen Methylierung im Gehirn zwischen dem Menschen und seinen nächsten Verwandten beschäftigen [255, 275-278]. Jedoch zeigen all diese Studien eine entscheidende Schwachstelle, da die Analysen mit Gesamtgewebe durchgeführt wurden – während es mittlerweile bekannt ist, dass die DNA-Methylierung signifikant zwischen verschiedenen Gehirnregionen und auch zwischen der weißen und grauen Substanz innerhalb einer Region variiert [117, 279-281]. Deshalb wurde im Zuge dieses Projekts eine genomweite differenzielle Methylierungsanalyse neuronaler und nicht-neuronaler Zellen aus dem präfrontalen Cortex zwischen Mensch und Schimpanse durchgeführt. Da bisher, nach unserem Wissensstand, keine vergleichende Studie durchgeführt und veröffentlicht wurde, die auf separiertem Cortexgewebe basiert, handelt es sich hier um eine exploratorische Studie zur Erlangung erster Eindrücke und Hypothesenbildung.

1.1 RRBS zur Detektion aberranter Methylierungsmuster

Die Analyse der Methylierungsprofile birgt einige Schwierigkeiten. Es handelt sich zwar nur um ein Genom je Organismus, jedoch kann dieses mehr als 1.000 Epigenome je Individuum beinhalten, da die DNA-Methylierung je Zelltyp und auch während der Entwicklung oder bei Krankheiten sowie in Abhängigkeit von Umweltweinflüssen variiert [81]. Aktuell gibt es noch keine standardisierten Techniken, welche es ermöglichen, die Methylierungsinformationen über eine direkte Sequenzierung der DNA oder mittels Hybridisierungsmethoden auslesen zu können, da diese während der Amplifikation oder einer Klonierung gelöscht werden [230]. Deshalb behilft man sich zur Unterscheidung von methylierten und unmethylierten Cytosinen mit dem Einsatz von methylierungs-sensitiven und/oder -insensitiven Restriktionsenzymen, Affinitätsaufreinigungen oder der Bisulfitkonvertierung [111, 230, 282-287].

Die Techniken zur Untersuchung der globalen DNA-Methylierungsmuster haben sich in den letzten zehn Jahren rasant weiterentwickelt: Beginnend mit dem Restriction Landmark Genomic Scanning, über Microarrays bis hin zu sequenzierungsbasierten Analysen mittels NGS [111, 230, 282-285, 288, 289]. Der Genome Analyzer (GA) von Illumina war die erste Sequenzierplattform, mit der kurze Fragmente mit einer Länge von 35 bp oder 2x75 bp sequenziert werden konnten. Das Template wird *in situ* über Brückenbildung amplifiziert und anschließend, basierend auf der „Sequencing-by-Synthesis“ Technologie, sequenziert. Durch die kürzere Readlänge und die hohe Sequenzierleistung waren die Sequenzierkosten pro Base damals günstiger im Vergleich zur 454 Sequenzierung von Roche Applied Science und wurde somit zur meist genutzten Technik für genomweite DNA-Methylierungsanalysen [290]. Trotz der schnellen Weiterentwicklung der NGS-basierten Techniken und der damit einhergehenden Preisreduktion je sequenzierter Base, sind die Kosten für eine genomweite Sequenzierung via Whole-Genome Sequencing nach wie vor hoch. Da in dieser Arbeit das Methylom untersucht werden sollte, bot sich das Reduced Representation Bisulfite Sequencing (RRBS) als Technik der Wahl an. Durch die gezielte Anreicherung CpG-haltiger Sequenzen wird mittels RRBS nur ein für das Epigenom repräsentativer Teil eines Individuums analysiert; dies ist somit im Vergleich zum Whole-Genome Bisulfite Sequencing wesentlich effizienter und kostengünstiger. Ein weiterer positiver Aspekt zeigt sich in der geringen DNA-Menge (30-200 ng), die je Ansatz eingesetzt werden muss, sowie in der

genomischen Abdeckung von CpG-Inseln, Promotorbereichen, Exons, 3' untranslatierten Bereichen und repetitiven Elementen [291].

1.2 Intraspezifisch differenziell methylierte Regionen

Das Verständnis der epigenetischen Unterschiede, welche zur Differenzierung der Vorläuferzellen in neuronale oder nicht-neuronale Zellen führen, nimmt eine wichtige Rolle für neuroepigenetische Studien ein. Es konnte gezeigt werden, dass die DNA-Methylierung signifikant zwischen verschiedenen Gehirnregionen und auch zwischen der weißen und grauen Substanz innerhalb einer Region variiert [117, 279-281]. Das Gehirn ist durch eine hohe Komplexität charakterisiert, welche u.a. durch die Heterogenität der Zelltypen und Subpopulationen innerhalb dieser Zelltypen hervorgerufen wird. Neurone und Gliazellen sind zudem unterschiedlich stark in verschiedenen Gehirnregionen vertreten. Die Gehirnregionen wiederum weisen ebenfalls eine heterogene Zytoarchitektur, Konnektivität und Funktion auf [279, 281]. Deshalb ist es essentiell für das Erhalten von Einblicken in die epigenetische Landschaft des Gehirns, dass die epigenetischen Signaturen, wie DNA-Methylierung, individuell je Zelltyp und separat für jede Region untersucht werden.

Für unseren Datensatz wurde vor der Suche nach differenziell methylierten Regionen (DMRs) eine Korrespondenzanalyse basierend auf den 10.000 variabelsten Methylierungsloci angefertigt. Hierbei zeigte sich der höchste Methylierungsunterschied zwischen den Zelltypen, gefolgt von dem Unterschied zwischen den Spezies. Dieser Fund bestätigt die oben angesprochene Bedeutung der getrennten Analyse von neuronalen und nicht-neuronalen Zellen.

Im Menschen waren mehr als 90% der intraspezifischen DMRs neuronal hypomethyliert. Im Schimpansen dagegen war die Anzahl an hyper- und hypomethylierten neuronalen DMRs nahezu ausgeglichen. Der Vergleich der intraspezifischen DMRs mit anderen Spezies zeigt, dass ein Teil der DMRs im Menschen [292], Schimpansen und der Maus [293] konserviert vorliegen. Dies wurde ebenfalls in der Studie von Kessler *et al.* (2015) zur Auffindung differenziell methylierter CpG-Stellen beim Menschen und der Maus postuliert [293].

Für eine bessere Einordnung der funktionellen und biologischen Rolle der in den intraspezifischen DMRs gelegenen Gene wurde eine GO-Analyse durchgeführt. Diese ergab eine Anreicherung der Gene für zahlreiche biologische Prozesse, welche eine wichtige Rolle bei der Entwicklung und Funktion des Gehirns einnehmen. Dieses Ergebnis bestätigt die

Meinung, dass die Gene der intraspezifischen DMRs eine wichtige Aufgabe bei der Regulation der hoch komplexen Zusammenarbeit zwischen neuronalen und nicht-neuronalen Zellen übernehmen.

1.2.1 Kandidatengenanalysen in intraspezifischen DMRs

1.2.1.1 Neuropsychiatrische Erkrankungen

Des Weiteren wurde eine Kandidatengenanalyse mit Genen, die zur Entstehung von neuropsychiatrischen Erkrankung beitragen, und den DMRs durchgeführt, um weitere Gene identifizieren zu können, die einen Einfluss auf die Gehirnevolution genommen haben könnten. Einige Studien haben bereits gezeigt, dass verschiedene neurodegenerative und neuropsychiatrische Erkrankung sowie Krankheiten, die zu einer gestörten Entwicklung des Nervensystems führen, teilweise durch aberrante epigenetische Modifikationen herbeigeführt werden können [132, 294, 295]. Die Kandidatengenanalyse zeigte im Menschen eine Anreicherung verschiedener Gene für Mentale Retardierung, Autismus, Bipolare Störungen und Alzheimer. Im Schimpansen ergaben sich Anreicherungen für Autismus und Bipolare Störungen. Beide Spezies zeigten einige Gene für Autismus und Bipolare Störungen, die sowohl in den humanen als auch in den DMRs des Schimpansen übereinstimmten und in beiden Spezies neuronal hypomethyliert vorliegen. Dabei handelt es sich um die mit Autismus assoziierten Gene *SND1* (staphylococcal nuclease and tudor domain containing 1), *BAIAP2* (BAI1 associated protein 2) und *ADORA2A* (adenosine A2a receptor). *SND1* codiert einen transkriptionellen Co-Aktivatoren und ist essentiell für ein normales Zellwachstum (OMIM *602181). *BAIAP2* codiert ein brain-specific angiogenesis inhibitor (BAI1)-Bindepotein, das als Substrat der Insulinrezeptor-Tyrosinkinase fungiert. Deshalb wird diesem Protein eine wichtige Rolle bezüglich der Insulinversorgung im Zentralnervensystem zugesprochen. Des Weiteren ist *BAIAP2* in die Bildung von Lamellipodien und Filopodien bei freibeweglichen Zellen involviert und beeinflusst das neuronale Wachstum (OMIM * 605475). *ADORA2A* codiert für ein Mitglied der guanine nucleotide-binding protein (G-protein)-coupled receptor (GPCR) Überfamilie und sorgt für den Anstieg intrazellulärer cAMP Level. Dieses Protein hat eine Vielzahl verschiedener biologischer Funktionen, wie die Aufrechterhaltung des Herzrhythmus und die Durchblutung der Herzgefäße, sowie die Kontrolle des cerebralen und renalen Blutflusses, aber auch eine Beteiligung bei Immunfunktion, Schmerzregulation und Schlaf (OMIM * 102776). Beide

Spezies zeigten auch Gene für Bipolare Störungen, bei denen die DMRs beider Spezies überlappten. Dabei handelte es sich um *SDK1* (sidekick cell adhesion molecule 1), *PRKAG2* (protein kinase AMP-activated non-catalytic subunit gamma 2), *ST3GAL1* (ST3 beta-galactoside alpha-2,3-sialyltransferase 1), *CACNA1B* (calcium voltage-gated channel subunit alpha1 B), *OLFM1* (olfactomedin 1) und *KACNQ2* (potassium voltage-gated channel subfamily Q member 2). *SDK1* codiert für ein Protein der Überfamilie der Immunglobuline. Basierend auf Studien in Hühnern wird angenommen, dass *SDK1* ein bestimmender Faktor bei der lamina-spezifischen synaptischen Konnektivität ist (OMIM * 607216). *PRKAG2* codiert für eine nicht-katalytische Untereinheit der AMP-activated protein kinase (AMPK), die ein wichtiger Sensor für die Überprüfung des zellulären Energielevels ist und dieses reguliert. Mutationen im *PRKAG2* Gen sind bisher nur mit Herz-Kreislauf-Erkrankungen, wie beispielsweise dem autosomal dominanten Wolff-Parkinson-White Syndrom (WPW; 194200), assoziiert (OMIM * 602743). *ST3GAL1* codiert für ein Typ II Membranprotein, das den Transfer von Sialinsäure katalysiert (OMIM * 607187). Das Genprodukt von *OLFM1* zeigt weitgehende Sequenzähnlichkeit in der Abfolge der Aminosäuren mit einem Protein in Ratten, welches in dem neuronalen olfactomedin-related Endoplasmatischen Retikulum lokalisiert ist. Über die Funktion dieses Proteins ist noch nichts bekannt, jedoch wird durch die reichliche Expression im Gehirn angenommen, dass es eine essentielle Rolle für das Nervengewebe spielt (OMIM * 605366). *CACNA1B* codiert für eine porenbildende Untereinheit eines N-Typ spannungsabhängigen Calciumkanals, der die Kontrolle der Neurotransmitterfreisetzung in Neuronen kontrolliert (OMIM * 601012). *KCNQ2* codiert für einen spannungsabhängigen Kaliumkanal, der im Gehirn exprimiert wird und eine wichtige Funktion bei der Regulation der neuronalen Erregbarkeit einnimmt. Defekte in diesem Gen sind mit Epilepsie assoziiert (OMIM * 602235).

1.2.2 Repetitive Elemente in intraspezifischen DMRs

Da mittels RRBS auch teilweise repetitive Elemente (REs) detektiert werden können, wurde im nächsten Schritt überprüft, inwieweit eine differenzielle Methylierung in den REs zwischen den Zelltypen bzw. Spezies vorliegt. Es konnte eine neuronale Hypomethylierung in einigen humanen intraspezifischen DMRs detektiert werden, welche RE enthalten. Es ist bekannt, dass Transposable Elemente (TE) durch ihre stetige Aktivität und Akkumulation im menschlichen Genom über mehrere Jahrtausende hinweg einen starken Einfluss auf die

Evolution des Primaten-Genoms genommen haben [261]. Die Expression und die Mobilisation der TE ist normalerweise in den meisten Zelltypen reprimiert, da die Retrotranspositionsaktivität von LINE1 (Long INterspersed Element-1) und anderen TE mit einer genomischen Instabilität assoziiert ist. Die DNA-Methylierung und Veränderungen in der Chromatinstruktur nehmen eine wichtige Rolle bei der Regulation der Aktivität von TE ein [296]. Die Insertion von TE kann zur Entstehung von Krankheiten führen. Im Menschen ist LINE1 das einzige aktive autonome TE, welches nicht nur die Fähigkeit besitzt, die eigene RNA über einen „copy-and-paste“ Mechanismus einzubauen, sondern auch fremde RNA wie beispielsweise von Alu, SVA (SINE/VNTR/Alu) und zellulären RNAs. Bis heute sind 124 LINE1-vermittelte Insertionen bekannt, die zur Entwicklung genetischer Erkrankungen führen [297]. Während der Differenzierung der neuronalen Vorläuferzellen in neuronale und nicht-neuronale Zellen zum Zeitpunkt der Gehirnentwicklung, kommt es zu einer transkriptionellen Aktivierung des LINE1 Promotors [298, 299]. Die Retrotranspositionsereignisse finden zudem wesentlich häufiger in Neuronen im Vergleich zu anderen neuronalen Zelltypen statt, was zur Entstehung von neuronalen somatischen Mosaiken führt [263, 300, 301].

Vergleicht man jene Gene, welche mehrfach in LINE1 auftreten und mit Funktionen im Gehirn assoziiert sind, findet sich mit *ARHGEF10* (Rho guanine nucleotide exchange factor 10) die einzige Übereinstimmung für die gleiche DMR. *ARHGEF10* liegt beim Menschen in den neuronalen Zellen und in den nicht-neuronalen Zellen im Schimpansen hypomethyliert vor. Dieses Gen gehört zur Familie der Rho guanin nucleotide exchange factors (GEFs), welche in die neuronale Morphogenese und Konnektivität involviert sind. Verhoeven *et al.* konnten über Expressionsstudien in der Maus zeigen, dass *ARHGEF10* im zentralen und peripheren Nervensystem stark exprimiert wird und möglicherweise eine Rolle bei der Myelinisierung peripherer Nerven während der Entwicklung einnimmt [302]. Mutationen in diesem Gen konnten bereits mit einer verlangsamten Nervenleitungsgeschwindigkeit in motorischen und sensorischen Neuropathien in Verbindung gebracht werden (OMIM *608136).

Des Weiteren war auffällig, dass 16% (n=25/156) der Regionen in den L1-assoziierten DMRs mit MicroRNAs (miRNAs; MIR) übereinstimmten. In der aktuellen Literatur wird angenommen, dass miRNAs die Genexpression in vielen biologischen Prozessen regulieren.

Viele miRNAs werden im adulten Nervensystem stark exprimiert und sind in verschiedene Aspekte der Dendriten-Remodellierung und synaptischen Plastizität involviert [303].

1.3 Interspezifisch differenziell methylierte Regionen

Im Vergleich zu anderen Säugern ist das menschliche Gehirn außergewöhnlich groß – relativ gemessen an der Körpergröße [55]. Einige Studien haben gezeigt, dass die Evolution das menschliche Gehirn auf Ebene der Organisation, wie beispielsweise durch Veränderungen in der weitreichenden corticalen Konnektivität, Histologie und lokaler Organisation [304], sowie bei der Gen-, Protein- und Metabolitenexpression modifiziert hat [251, 305, 306]. Das menschliche Gehirn ist rund dreimal so groß wie das Gehirn unseres nächsten lebenden Verwandten, des Schimpansen. Insbesondere enthält der humane cerebrale Cortex doppelt so viele Zellen wie der des Schimpansen, zudem weisen die Zellen der Netzwerke ein unterschiedliches Verhalten zwischen den beiden Spezies auf. Es wird angenommen, dass die Expansion des Neocortex während der Evolution der Primaten zur verstärkten kognitiven Leistungsfähigkeit im Menschen im Vergleich zu unseren nächsten lebenden Verwandten beigetragen hat [307-309]. Durch die Vergrößerung des Neocortex kam es auch zu einer Zunahme der corticalen Neuronen, die während der fetalen Entwicklung gebildet werden [20, 64, 310, 311]. Dies spiegelt primär eine vergrößerte und verlängerte Wachstumskapazität der humanen neuralen Stamm- und Vorläuferzellen innerhalb der germinalen Matrixzone bei der Entwicklung des Neocortex wieder [312].

Innerhalb des interspezifischen Vergleichs konnte nur eine sehr geringe Anzahl an DMRs in den neuronalen Zellen (N=96) im Vergleich zu den nicht-neuronalen Zellen (N=666) detektiert werden. Dieser große Unterschied zwischen der Anzahl an DMRs in den neuronalen und nicht-neuronalen Zellen könnte auf adaptive Veränderungen in der Homöostase des menschlichen Gehirns zur Erhöhung des Metabolismus und der synaptischen Transmission hindeuten. Es ist bekannt, dass das menschliche Gehirn aufgrund seiner Größe einen enormen Energiebedarf hat. Das sich entwickelnde Gehirn eines Neugeborenen benötigt rund 60-80% und das Gehirn eines Erwachsenen rund 20-25% an RMR (*resting metabolic rate*). Im Vergleich zu nicht-menschlichen Primaten, die durchschnittlich weniger als 10% im Ruhezustand benötigen, ist der Energieverbrauch beim Menschen ungewöhnlich hoch [313, 314]. Die Evolution der höheren kognitiven Funktionen beim Menschen steht in Verbindung mit einem erhöhten Verbrauch des Gehirns an Glukose

und Sauerstoff. Zur Aufrechterhaltung der elektrischen und synaptischen Aktivität der Neuronen ist eine ausreichende Verfügbarkeit von ATP notwendig. Die Gliazellen, welche mit trophischen Funktionen assoziiert sind, scheinen eine zentrale Rolle bei der metabolischen Versorgung des Nervensystems und der Deckung des zunehmenden Energiebedarfs der aktiven Neuronen einzunehmen [315]. Zudem wird angenommen, dass die Interaktion zwischen Neuronen und Gliazellen zu einer Modulation der neuronalen Transmission und synaptischer Plastizität führt [316, 317].

Die GO-Analyse ergab eine zweifache Anreicherung in den interspezifischen DMRs für Gene, die mit der Neurogenese assoziiert sind, wie beispielsweise *CELSR1* und *SHANK3*. Das *CELSR1* (cadherin EGF LAG seven-pass G-type receptor 1) Gen codiert für ein Transmembranprotein der Flamingo-Unterfamilie, welche zur Überfamilie der Cadherine gezählt wird. Mutationen in diesem Gen wurden bereits mit dem Neuralrohrdefekt Spina bifida in Verbindung gebracht (OMIM * 604523). *SHANK3* (SH3 and multiple ankyrin repeat domains 3) gehört zur Familie der Shank Genfamilie. Die Shank Proteine spielen eine Rolle bei der Synapsenbildung und der Reifung der dendritischen Dornfortsätze. Mutationen in diesem Gen sind mit Autismus-Spektrum-Störungen (ASD) assoziiert, zu denen frühkindlicher Autismus, das Asperger-Syndrom und atypischer Autismus gezählt werden. Die ASD ist charakterisiert durch ein gestörtes soziales Interaktions- und Kommunikationsverhalten, sowie beschränkte Verhaltensmuster und Interessen. Des Weiteren können Mutationen in *SHANK3* ursächlich für Schizophrenie und Phelan-McDermid Syndrom sein (OMIM * 606230). Bei diesen beiden Genen könnte es sich um epigenetische Regulatoren handeln, welche in die menschliche Gehirnevolution involviert waren. Sie stellen somit gute Kandidatengene für nachfolgende Analysen dar.

1.3.1 Kandidatengenanalysen in interspezifischen DMRs

1.3.1.1 Neuropsychiatrische Erkrankungen

Die Kandidatengenanalyse mit Genen, die zur Entstehung von neuropsychiatrischen Erkrankungen beitragen sowie die Analyse der interspezifischen DMRs zeigten keine signifikanten Anreicherungen, lediglich eine Tendenz ($p=0,081$) für Gene, die mit Mentaler Retardierung (MR) assoziiert sind. Interessanterweise liegen 3 von 9 MR Genen auf dem X-Chromosom. Man nimmt an, dass die X-chromosomal lokalisierten Gene einen Einfluss auf die Entwicklung der kognitiven Fähigkeiten beim Menschen haben [318]. Diese Gene wiesen

alle eine Hypermethylierung des Genkörpers beim Menschen im Vergleich zum Schimpansen auf, was normalerweise mit einer genaktivierenden Wirkung in Verbindung gebracht wird [260]. Die einzige Ausnahme stellte die Region von *GDI1* dar. In diesem Gen wurde die Hypermethylierung im Promotorbereich detektiert. *GDI1* (GDP dissociation inhibitor 1) codiert für ein Protein, das den GDP-GTP Austausch von Mitgliedern der Rab-Proteinfamilie reguliert. *GDI1* wird überwiegend in neuralem und sensorischem Gewebe exprimiert (OMIM * 300104). Das *OPHN1* Gen, ebenfalls auf dem X-Chromosom gelegen, zeigte sowohl in den neuronalen als auch in den nicht-neuronalen Zellen im Menschen eine Hypermethylierung im Vergleich zum Schimpansen. *OPHN1* (oligophrenin 1) codiert für ein Rho-GTPase-aktivierendes Protein, welches die GTP Hydrolyse einiger Mitglieder der Rho Subfamilie fördert. Das Protein spielt eine Rolle in X-linked Mentaler Retardierung, welche einhergeht mit zerebellarer Hypoplasie und charakteristischem fazialen Erscheinungsbild. Die Inaktivierung des RhoGAP Proteins führt möglicherweise zu einer konstitutiven Aktivierung des GTPase Targets, welches dafür bekannt ist, die Zellmigration und den Auswuchs der Axone und Dendriten zu beeinflussen. Des Weiteren wird angenommen, dass es einen Zusammenhang zwischen kognitiven Störungen und einem Defekt in einem der Rho GTPase abhängigen Signalwege gibt [319]. Kasri *et al.* (2009) konnten an kultivierten Neuronen aus dem Hippocampus einer Ratte zeigen, dass *Ophn1* eine Schlüsselrolle in der aktivitätsabhängigen Reifung und Plastizität exzitatorischer Synapsen innehat [320] (OMIM *300127). Bei *IL1RAPL1* (interleukin 1 receptor accessory protein like 1) handelt es sich ebenfalls um ein X-chromosomales Gen, welches hypermethyliert in humanen nicht-neuronalen Zellen vorliegt. Dieses Gen codiert für ein Mitglied der Interleukin 1-zugehörigen Proteine und befindet sich in einer Region auf dem X-Chromosom, welche mit einer nicht-syndromalen Form der geistigen Behinderung assoziiert ist. *IL1RAPL1* wird stark in postnatalen Gehirnstrukturen, die in die hippocampale Gedächtnisbildung involviert sind, exprimiert. Dies deutet auf eine spezialisierte Rolle bei den physiologischen Prozessen, die dem Lern- und Erinnerungsvermögen zugrunde liegen, und bei der Synapsenbildung und -stabilisierung hin (OMIM * 300206).

1.3.1.2 Human-spezifisch regulierte Gene

Neben der Anreicherung von MR Genen in den interspezifischen DMRs wurde auch eine Anreicherung von Genen detektiert, welche eine human-spezifische Regulation im Gehirn

aufweisen. In Summe tauchten 25 DMR-assoziierte Gene in nicht-neuronalen Zellen, 7 in neuronalen Zellen, 88 im Menschen und 23 im Schimpansen auf. Im interspezifischen Vergleich wurden vier Gene beobachtet, die sowohl in neuronalen als auch nicht-neuronalen Zellen spezie-spezifische Methylierungsveränderungen zeigen. Hierbei handelte es sich u.a. um *OPCML* und *TNFRSF11A*. Diese Gene waren im Menschen in beiden Zelltypen hypermethyliert. Das *OPCML* (opioid binding protein/cell adhesion molecule like) Protein ist ein Mitglied der IgLON Unterfamilie und ist in der Plasmamembran lokalisiert. Da dieses Protein hoch konserviert in verschiedenen Spezies während der Evolution vorliegt, könnte es eine fundamentale Rolle bei Säugern innehaben (OMIM * 600632). *TNFRSF11A* (Tumor Necrosis Factor receptor superfamily member 11a; auch bekannt als RANK) codiert für einen Rezeptor der TNF-Rezeptor Superfamilie und interagiert mit verschiedenen Proteinen der TRAF Familie. Dieser Rezeptor und sein Ligand TNFSF11 (Tumor Necrosis Factor superfamily member 11; auch bekannt als RANKL) sind wichtige Regulatoren bei der Interaktion zwischen T-Zellen und dendritischen Zellen. Hanada *et. al* (2009) fanden heraus, dass RANK und RANKL eine essentielle Rolle im Gehirn einnehmen. Sie untersuchten, inwieweit RANK bei der Auslösung des Fiebermechanismus, der über astrozytische Zytokine stimuliert wird, beteiligt ist. Es konnte gezeigt werden, dass RANKL zur Aktivierung der Gehirnregionen beiträgt, welche in die Thermoregulation und Fieberinduktion via COX2-PGF2/EP3R Signalweg involviert sind. Daraus kann geschlossen werden, dass die Schlüsselenzyme der Osteoklasten-Differenzierung auch eine zentrale Rolle bei der Auslösung von Fieber einnehmen (OMIM * 603499) [321]. *ZNF717* war in mehreren Kontrasten vorhanden und wies im Menschen sowohl in den neuronalen als auch in den nicht-neuronalen Zellen eine Hypomethylierung der Promotorregion auf, wohingegen der Promotor im Schimpansen in beiden Zelltypen hypermethyliert vorliegt. Die weiter downstream gelegene Region wies eine Hypermethylierung in den nicht-neuronalen Zellen auf und der Genkörper war im Menschen neuronal hypomethyliert und im Schimpansen hypermethyliert, im Vergleich zu den nicht-neuronalen Zellen. *ZNF717* (zinc finger protein 717) codiert für ein Kruppel-associated box (*KRAB*) zinc-finger Protein, welches zu einer großen Gruppe von Transkriptionsregulatoren in Säugern gehört. Diese Proteine binden Nukleinsäuren und nehmen eine wichtige Rolle in verschiedenen zellulären Funktionen, wie Zellproliferation, Differenzierung und Apoptose ein. In einer Studie von Nowick wurden zwischen Mensch und Schimpanse im Gehirn 90 differenziell exprimierte Gene identifiziert, die für

Transkriptionsfaktoren codieren. Diese werden über ein Koexpressionsnetzwerk organisiert, welches aus zwei Modulen mit unterschiedlichen Funktionen besteht. *ZNF717* übernimmt in diesem Netzwerk die Funktion eines Bindeglieds zwischen diesen beiden Modulen [254]. Durch diese zentrale Rolle bei der Koordination aufeinander abgestimmter Veränderungen innerhalb dieses Expressionsnetzwerks ist es vorstellbar, dass epigenetische Veränderungen in der Regulation von *ZNF717* einen möglichen Auslöser für die menschliche Gehirnevolution darstellten.

1.4 Schlussfolgerung und Ausblick

Zusammenfassend kann gesagt werden, dass epigenetische Unterschiede zwischen den humanen neuronalen und nicht-neuronalen Zellen wesentlich häufiger detektiert werden konnten als interspezifische Unterschiede. Evolutionäre Veränderungen der DNA-Methylierung in neuronalen und nicht-neuronalen Zellen könnten zur Vergrößerung des menschlichen Gehirns und der Zunahme der kognitiven Fähigkeiten beigetragen haben. Die Veränderungen im humanen Gehirn führten wahrscheinlich zur Hypermethylierung vieler Gene in den nicht-neuronalen Zellen. Dies führt zu dem Schluss, dass der Einfluss der verschiedenen nicht-neuronalen Zelltypen auf die humane Gehirnevolution bis dato unterschätzt worden ist. Es scheint, dass die nicht-neuronalen Zellen nicht allein eine unterstützende physikalische Funktion innehaben oder für die Versorgung der Neuronen mit Nährstoffen verantwortlich sind, sondern auch einen erheblichen Einfluss auf die Funktionen der Neuronen haben. Da die Interaktion zwischen neuronalen und nicht-neuronalen Zellen eine wichtige Rolle bei gedanklicher Aktivität und der Entstehung von neuropsychiatrischen Erkrankungen einnimmt, entsprechen die vielen intra- und interspezifischen DMRs, welche mit neuropsychiatrischen Erkrankungen assoziiert werden konnten, den Erwartungen. Durch diese Studie konnten erstmals Kenntnisse darüber gewonnen werden, inwiefern sich die DNA-Methylierungsmuster zwischen Mensch und Schimpanse auf Ebene der neuronalen bzw. nicht-neuronalen Zellen des präfrontalen Cortex unterscheiden. Dies führt zu einigen interessanten Erkenntnissen, aber auch zu einer Sammlung von Kandidatengenen, die in weiteren Studien genauer in diesem Kontext analysiert werden können. Es wäre notwendig, die DNA-Methylierung in einen Zusammenhang mit der Expression dieser Gene zu bringen. Zudem wäre eine zusätzliche Spezies, wie beispielsweise der Rhesus-Affe von Vorteil, um klären zu können, ob die Veränderungen durch Adaption beim Mensch entstanden sind; dies

kann nicht allein durch den Vergleich zwischen Mensch und Schimpanse geklärt werden. Um logistische und ethische Hürden bei der Probenbeschaffung zu umgehen, würde sich auch die Arbeit an neuronalen oder nicht-neuronalen Zelllinien, sowie induzierten pluripotenten Stammzellen (iPSCs) in Zellkultur anbieten. Ein entscheidender Vorteil wäre hier, dass keine Zellseparation durchgeführt werden müsste und auch ein einzelner neuronaler oder nicht-neuronaler Zelltyp untersucht werden könnte. Zudem können über die iPSCs auch interspezifische Unterschiede in der frühen Phase der Gehirnentwicklung studiert werden. Es besteht weiterhin die Möglichkeit, einer verbesserten Standardisierung und damit der Vorteil, einige Störfaktoren zu beseitigen, die sich negativ auf die Ergebnisse auswirken, wie beispielsweise die ungleiche Anzahl an neuronalen und nicht-neuronalen Zellen des Gesamtgewebes, das Alter der Zellen bzw. Individuen, sowie die post mortem Zeit vor der Gehirnsektion. Auch besteht in Zellkultur die Möglichkeit, Expressionstudien und funktionelle Studien durchführen zu können. In weiterführenden Studien sollte zudem die Analyse der nicht-CpG Methylierung und Hydroxymethylierung einfließen. Die Methylierung an nicht-CpG-Stellen ist auf bestimmte Zelltypen, wie pluripotente Stammzellen, Oozyten, Neurone und Gliazellen beschränkt [322]. In differenzierten somatischen Zellen kommt diese bei etwa 0,02% aller methylierten Cytosine vor [80, 81]. Es wird angenommen, dass die nicht-CpG Methylierung mit einer transkriptionellen Silllegung in Zellen des Gehirns korreliert [323] und verschiedene Gehirnzelltypen unterschiedliche nicht-CpG Methylierungslevel aufweisen. Beispielsweise zeigen Neurone deutlich höhere Methylierungslevel an nicht-CpG-Stellen als Gliazellen [324]. Die Hydroxymethylierung (5hmC) ist eine kovalente Modifikation der Cytosine und ein oxidatives Derivat von 5mC, welches bei der aktiven Demethylierung gebildet wird [325, 326]. 5hmC ist im vollentwickelten Zentralnervensystem stark vertreten, insbesondere in der Nachbarschaft von Genen, die mit synaptischen Funktionen assoziiert sind [327, 328]. Diese epigenetische Modifikation spielt vermutlich eine wichtige Rolle bei der Funktion der Synapsen, dem Lernen und der Gedächtnisbildung [329]. Deshalb könnten sowohl die nicht-CpG Methylierung als auch die Hydroxymethylierung einen wichtigen Faktor bei der Evolution des menschlichen Gehirns ausgemacht haben und interessant für nachfolgende Studien sein.

2. Methylierungsanalysen von TSG bei Brustkrebspatientinnen

Bei rund 25% aller diagnostizierten Krebserkrankungen bei Frauen handelt es sich um Brustkrebs; er stellt somit weltweit die häufigste Krebserkrankung bei Frauen dar [Int. 30]. Die Überlebensrate variiert sehr stark je nach Herkunftsland, so beläuft sich diese auf 80% in Nordamerika, Schweden und Japan; auf rund 60% in Ländern mittleren Einkommens und fällt gar unter 40% in Ländern niedrigen Einkommens [228]. In Ländern mittleren und niedrigen Einkommens ist für viele Frauen keine Möglichkeit gegeben, ein routinemäßiges Mammographie-Screening durchführen zu lassen. Dadurch wird die Krebserkrankung meist in einem sehr späten Stadium diagnostiziert. Zudem ist keine adäquate Behandlung oder palliative Betreuung möglich [330].

Die Diagnoseverfahren in Ländern hohen Einkommens dahingegen beinhalten standardmäßig Techniken wie Mammographie, Ultraschalluntersuchungen, Magnetresonanztomographie und Biopsien [Int. 31]. Zudem werden bei Verdacht auf familiären Brustkrebs genetische Untersuchungen veranlasst. Bei rund 25% aller Frauen mit Brustkrebs kann eine familiäre Häufung beobachtet werden. Jedoch können nur bei 5-10% aller Patientinnen pathogene genetische Mutationen identifiziert werden [Int. 6]. Da neben den genetischen Ursachen auch die Kausalität von epigenetischen Veränderungen auf die Karzinogenese bekannt sind, war die Zielsetzung dieser Arbeit, die Analyse von Methylierungsveränderungen in Promotorregionen von TSG und diese als mögliche Ursache für die Brustkrebskrankung bei *BRCA1/BRCA2* mutations-negativen Patientinnen ausfindig machen zu können.

2.1 DBAS zur Detektion aberranter Methylierungsmuster

Da es sich in dieser Studie um eine Kandidatengenanalyse handelte, eignete sich für dieses Projekt nicht die oben erwähnte RRBS Methode. Für die gerichtete Methylierungsanalyse spezifischer Zielregionen oder CpG-Stellen gilt die Bisulfit-Pyrosequenzierung nach wie vor als Goldstandard. Jedoch können mit dieser Technik nur kleine Regionen (bis zu 115 bp) und somit auch wenige CpG-Stellen in einem Lauf untersucht werden [331]. Es besteht auch keine Möglichkeit, einzelne Allele zu analysieren, da nur durchschnittliche Methylierungswerte je CpG-Stelle gemessen werden können. In dieser Studie lag der Fokus jedoch auf der Detektion von aberrant methylierten Allelen. Die Plasmid-Bisulfit-Sequenzierung dagegen ermöglicht die Analyse einzelner Allele. Jedoch ist diese Methode

sehr arbeitsintensiv und ebenfalls über die niedrige Länge der Amplikons limitiert [332]. Das Deep Bisulfite Sequencing (DBS) stellt demnach eine attraktive Lösung dar. Diese Technik wurde erstmals auf dem GS Junior von Roche etabliert und ermöglicht eine hochauflösende Sequenzierung spezifischer Amplikons auf Einzel-Allel-Ebene [333-335]. Da mit dem DBS-Protokoll von Roche nur 24 Proben und maximal zwei Gene pro Lauf analysiert werden können, wurde im Rahmen dieses Projekts eine neue, an das oben erwähnte RRBS Protokoll angelehnte Deep Bisulfite Amplicon Sequencing (DBAS) Methode entwickelt und etabliert, welche mit dem MiSeq kompatibel ist. Das DBAS ermöglicht die Analyse von 48 Proben und 12 verschiedene Amplikons je Probe in einem einzigen MiSeq-Lauf. Die Limitierung hierbei liegt ausschließlich in der Anzahl an Indexkombinationen und der Kapazität des Sequenziergeräts. Die Coverage liegt durchschnittlich zwischen 5.000 und 25.000 Reads je Probe. Deshalb eignet sich diese Anwendung besonders, um seltene Epimutationen bzw. geringe Methylierungsunterschiede detektieren zu können und stellt somit eine kostengünstige und effiziente Alternative zum ursprünglichen DBS auf dem Roch GS Junior dar.

2.2 Methylierungslevel der Promotorbereiche ausgewählter TSG

Zahlreiche Studien, v.a. im Bereich der Krebsforschung, haben gezeigt, dass der Promotorbereich von TSG in gesunden Individuen in unmethylierter Form vorliegt [265, 336-340]. Dies ist entscheidend für die aktive Transkription dieser Gene, da eine Hypermethylierung in dieser Genregion zu einer Stilllegung der TSG führen kann [341-343]. Im Gegensatz zu Protoonkogenen und onkogenischen Proteinen, inhibieren Proteine, die über TSG codiert werden, das Zellwachstum oder das Überleben einer Zelle. Die Inaktivierung von TSG führt deshalb durch die Eliminierung negativ-regulatorischer Proteine zur Entwicklung von Tumoren. In einigen Fällen inhibieren die tumorsuppressorgenen Proteine dieselben zellregulatorischen Signalwege, die über onkogenische Produkte stimuliert werden [344]. Deshalb wurde in dieser Studie die Promotormethylierung von TSG, bei *BRCA1/BRCA2* mutations-negativen Brustkrebspatientinnen auf Einzel-Allel-Ebene untersucht.

BRCA1 (BRCA1, DNA repair associated) und *BRCA2* (BRCA2, DNA repair associated) sind zwei der wichtigsten Brustkrebs-Suszeptibilitätsgene und werden zu den Tumorsuppressorgenen gezählt. Beide Proteine kommen an unterschiedlichen Stellen der DNA-Schadensantwort

(engl. *DNA damage response*, DDR) und DNA-Reparatur zum Einsatz und dienen somit dem Schutz des Genoms vor doppelsträngigen DNA-Schäden während der Replikation. BRCA1 ist ein pleiotropes DDR Protein und ist sowohl bei der Checkpoint-Aktivierung als auch der DNA-Reparatur beteiligt [345, 346]. Das BRCA1 Protein formt in Verbindung mit anderen TSG sensorischen Einheiten, die den DNA-Schaden erkennen und Signalgebern einen großen Proteinkomplex, welcher aus mehreren Untereinheiten besteht. Dieser Komplex wird auch als *BRCA1-associated genome surveillance complex* (BASC) bezeichnet [345]. BRCA2 spielt eine essentielle Rolle in verschiedenen DNA-Reparatursignalwegen und nimmt eine Schlüsselrolle bei der Reparatur von DNA-Schäden ein, wie DNA-Doppelstrangbrüchen und Querverbindungen innerhalb eines DNA-Stranges. Darüberhinaus verhindert BRCA2, dass es zu einer nukleolytischen Degradierung an der Replikationsgabel kommt. Beide Funktionen sind direkt oder indirekt in die Aufrechterhaltung der Telomere involviert [347].

Keimbahnmutationen in einer Kopie dieser beiden Gene führen zum erblichen Brust- und Eierstockkrebsyndrom (engl. *hereditary breast and ovarian cancer*, HBOC), welches einer autosomal-dominanten Vererbung unterliegt. Dieses Syndrom ist neben early-onset Brustkrebs auch mit einem erhöhten Risiko für Eierstock-, Eileiter-, Kehlkopf-, Bauspeicheldrüsen-, Magen- und Prostatakrebs assoziiert. Das HBOC Syndrom erklärt rund 5-7% aller Brustkrebsfälle. Davon betroffene Individuen haben ein Lebenszeitrisiko von 50-80% für die Entwicklung von Brustkrebs und rund 30-50% für die Entwicklung von Eierstockkrebs [346, 348].

Neben *BRCA1* und *BRCA2* werden unter anderem *TP53* (tumor protein p53) [337] und *PTEN* (phosphatase and tensin homolog) [338] zu den Suszeptibilitätsgenen mit hoher Penetranz und *ATM* (ATM serine/threonine kinase) [339] zu den Suszeptibilitätsgenen mit einer mäßigen Penetranz gezählt und als Kandidatengene mit in diese Studie aufgenommen. Das Genset wurde durch weitere TSG ergänzt, die bereits mit einer aberranten Promotormethylierung bei Brustkrebspatienten in Verbindung gebracht werden konnten. Hierzu gehören *MLH1* (mutL homolog 1) [340], *RAD51C* (RAD51 paralog C) [265] und *RB1* (RB transcriptional corepressor 1) [336]. Zudem wurde *ESR1* (estrogen receptor 1) in diese Studie mitaufgenommen. *ESR1* ist kein TSG, aber in die Entwicklung von Brusttumoren involviert und dient deshalb als internes Kontrollgen [349].

Die in dieser Studie analysierten TSG wiesen eine sehr niedrige durchschnittliche Methylierung des Promotors zwischen 0,3% und 0,9% in allen untersuchten Gruppen auf.

Diese Werte sind konkordant mit bereits veröffentlichten Studien, die ebenfalls eine Hypomethylierung im Promotorbereich dieser TSG beschreiben [265, 337-340]. Zudem konnte kein signifikanter Unterschied zwischen den Patienten- und Kontrollgruppen gefunden werden. Die Promotorregion von *ESR1* zeigte, in Übereinstimmung mit der Literatur [350], 5-10 Mal höhere durchschnittliche Methylierungswerte von rund 3%, verglichen mit den TSG.

Aufgrund der niedrigen durchschnittlichen Methylierungswerte (<1%) wurden rund 18% der analysierten Patienten- und Kontrollproben, basierend auf statistischen Berechnungen, als hypermethyliert (\emptyset Methylierung: 1%-3%) eingestuft. Dies deutet auf eine sehr strenge Regulation zur Aufrechterhaltung des unmethylierten Status des Promotorbereichs bei TSG hin. Die Wichtigkeit der TSG wird alleine schon deutlich durch die essentiellen Aufgaben, die diese bei der DNA-Reparatur, Zellzykluskontrolle und anderen Mechanismen zur Erhaltung der genomischen Integrität einnehmen [351]. Die Hypermethylierung von CpG-Inseln spezifischer Promotorregionen kann über die transkriptionelle Stilllegung von TSG zu Karzinogenese und die damit verbundene Initiierung und Entwicklung einer Krebserkrankung führen [352, 353].

In einer Studie von Hansmann *et al.* (2012) wurden via Bisulfit-Pyrosequenzierung die Promotorbereiche von *ATM*, *BRCA1*, *BRCA2*, *RAD51C*, *PTEN* und *TP53* in Blutzellen von mehr als 600 *BRCA1/BRCA2* mutations-negative BC und/oder Eierstockkrebspatientinnen untersucht. Es wurde ein Schwellenwert von 6% durchschnittlicher Methylierung eingesetzt, um Individuen mit einer konstitutiven Epimutation auffindig machen zu können. Unter Verwendung dieses Schwellenwerts konnte in unseren Daten eine Patientenprobe mit einer durchschnittlichen Methylierung von 12,9% in *BRCA1* detektiert werden. Dies entspricht, über alle untersuchten EO und HR BC-Patienten berechnet, einem Anteil von 1,3% und ist konkordant mit dem Epimutationsscreening von Hansmann *et al.* (2012). In dieser Studie wiesen rund 1,4% der untersuchten Individuen einen Methylierungswert von mehr als 6% in *BRCA1* auf [265].

2.3 Analyse der Epimutationen in TSG

Angelehnt an die oben erwähnte Studie von Hansmann *et al.* (2012) wurde bei allen Patienten- und Kontrollproben ein Screening zur Detektion von Epimutationen auf Einzel-Allel-Ebene durchgeführt. In Summe wurden bei 49% der EO BC, 44% der HR BC und 43% der

Kontrollgruppe Proben mit Epimutationen detektiert. Am häufigsten zeigten die untersuchten Individuen aberrant methylierte Allele in *RAD51C* (26%), *BRCA1* (13%) und *TP53* Region 2 (16%). Am wenigsten betroffen waren die Promotorbereiche von *ATM* Region 1 (1%) und 2 (2%), *MLH1* (1%) und *PTEN* (2%). Bei *BRCA2* und *TP53* Region 1 konnten keine Proben mit Epimutationen beobachtet werden. In *BRCA1* waren signifikant mehr BC Proben im Vergleich zu den Kontrollproben betroffen. Im Vergleich zu Hansmann *et al.* (2012) konnten deutlich mehr Proben mit aberrant methylierten Allelen detektiert werden, jedoch kommt dieser Unterschied schon alleine durch die Anzahl an Proben zustande, die für das Epimutationsscreening verwendet wurden. In unserer Studie flossen alle Proben in die Analyse ein, unabhängig von ihrer durchschnittlichen Methylierung. Würde man ebenfalls nur Proben mit einer durchschnittlichen Methylierung von 6% betrachten, würde der Anteil der Proben mit aberrant methylierten Allelen (1,3%) für *BRCA1* dem Wert (1,4%) von Hansmann *et al.* (2012) entsprechen.

Die Verteilung der Epimutationsraten (EMR) war sehr variabel in den einzelnen Kohorten und Promotorregionen. Mehr als 90% aller Proben zeigten EMR <1% und rund 8% wiesen EMR zwischen 1% und 2,5% auf. Es wurde lediglich eine Patientenprobe in *BRCA1* mit einer EMR von 14,7% detektiert. Diese Probe entspricht der oben genannten HR BC Probe mit einer durchschnittlichen Methylierung von 12,9%. Auch unter Ausschluss dieses Extremwertes konnte für *BRCA1* ein signifikanter Gruppenunterschied zwischen Patienten und Kontrollproben gemessen werden. Die analysierten Regionen von *BRCA1*, *RAD51C*, *ATM* Region 1 und 2, *MLH1* und *RB1* enthalten einen SNP mit einer hohen MAF, über den die Allele nach ihrer elterlichen Herkunft unterschieden werden konnten. Jedoch konnte keine Aussage darüber getroffen werden, welche Allele von der Mutter bzw. vom Vater stammen, da für diese Analyse kein elterliches Material zur Verfügung stand.

2.3.1 Biologische Aspekte für die Entstehung der Epimutationen

Bei den Epimutationen in *ATM*, *MLH1*, *RB1*, *PTEN* und *TP53* Region 1 kann von spontanen stochastischen und durch Umwelteinflüsse hervorgerufenen somatischen Epimutationen innerhalb des untersuchten Gewebes ausgegangen werden. Hierfür sprechen zum einen die extrem niedrigen Epimutationsraten (0,001%-0,031%) und zum anderen, dass die aberrant methylierten Allele keinem parentalen Ursprung zugewiesen werden konnten. Gegenteilig dazu konnten in *BRCA1* und *RAD51C* aufgrund des heterozygoten SNP die Epimutationen in

allen Proben nur einem Elternteil zugeordnet werden. Dies lässt grundsätzlich auf die Entstehung dieser Epimutationen während einer der beiden Reprogrammierungsphasen schließen. Während der Gametogenese und der frühen Embryogenese unterliegen die parental Genome einer Demethylierung, gefolgt von einer zell- und gewebsspezifischen Remethylierung. In dieser Phase kann es zu stochastischen und/oder zu durch Umwelteinflüsse bedingten Defekten in der Methylierung während der Reprogrammierungsphase kommen. Zu Beginn der Reprogrammierung werden die geschlechtsspezifischen Methylierungsmuster angelegt und nach erfolgter Befruchtung die somatischen Methylierungsmuster zur Aktivierung oder Stilllegung bestimmter Gene etabliert, welche für eine normale Entwicklung unabdingbar sind [354]. Es ist wichtig, dass während dieser Reprogrammierungsphase der unmethylierte Status in TSG aufrechterhalten wird. Es kann jedoch nicht sicher gesagt werden, ob es sich bei den Epimutationen von *BRCA1* und *RAD51C* um eine Mutation während einer der beiden Reprogrammierungsphasen handelt oder ebenfalls um spontane stochastische und durch Umwelteinflüsse hervorgerufene somatische Epimutationen innerhalb des untersuchten Gewebes. Bei einer Keimbahnmutation müssten alle Allele des betroffenen Elternteils diese Epimutation tragen und die Probe somit eine EMR von 50% aufweisen. Aufgrund der großen interindividuellen Schwankungen in den EMR bei *BRCA1* kann auch nicht mit Sicherheit von einem Mutationsereignis während der Reprogrammierungsphase im frühen Embryo ausgegangen werden. Jedoch steht fest, dass die Epimutationen in *BRCA1* und *RAD51C* sehr wahrscheinlich aus einer einzelnen Vorläuferzelle hervorgegangen sein müssen.

In allen Proben und Amplikons wurde überprüft, ob genetische Varianten vorliegen, um zum einen ausschließen zu können, dass diese ursächlich für die Krebserkrankung sind; zum anderen, dass diese einen Einfluss auf die aberrante Methylierung hatten, da bekannt ist, dass Sequenzvariationen zu Veränderungen in der Methylierung führen können [113].

2.3.2 Störfaktoren — Datenanalyse und technische Artefakte

An dieser Stelle muss zudem auch angemerkt werden, inwiefern es sich bei den niedrigen Epimutationsraten möglicherweise um technische und/oder bioinformatische Artefakte und nicht um biologisch relevante Epimutationen handeln könnte. Da mittels DBAS eine hohe Sequenziertiefe erreicht werden kann und es sich zudem um bisulfit-konvertierte DNA

handelt, müssen bei der Analyse der Daten strenge Filterkriterien angewandt werden, um diese Möglichkeit weitgehend ausschließen zu können.

Das Alignment bisulfit-konvertierter DNA ist im Vergleich zu unbehandelter DNA wesentlich schwieriger und aufwändiger, da die beiden DNA Stränge nach der Behandlung mit Natriumbisulfit nicht mehr komplementär zueinander sind und sich dadurch der Suchbereich signifikant erhöht. Des Weiteren ist die Komplexität der Sequenz herabgesetzt, da im codierenden und nicht-codierenden Strang weitgehend alle Cytosine zu Thymin umgewandelt werden und somit C-arm sind. Die jeweils revers komplementären Stränge hingegen sind G-arm und weisen einen normalen C-Gehalt auf. Zudem ist das C zu T Mapping assymetrisch, was bedeutet, dass die T's der Reads entweder als C oder als T in der Referenz gesehen werden können, aber nicht umgekehrt. Dies erhöht zwar nicht den Suchbereich, aber ein eindeutiges und genaues Mapping wird dadurch zusätzlich erschwert und kann zu falsch-positiven Alignments führen [355]. Um diese Problematik so klein wie möglich zu halten, wurden für das Alignment der Daten, während der Etablierung des Alignments und der weiteren Datenanalyse, unterschiedliche Mappingprogramme ausprobiert und auch verschiedene Strategien angewandt und miteinander verglichen. Diese Daten sind nicht in dieser Arbeit dargestellt. Hierbei hat sich herausgestellt, dass am wenigsten falsch-positive Alignments generiert werden, wenn das Programm Bismark (und nicht BSmap) verwendet wird und die Reads nicht nur je Amplikon gegen die Referenzsequenz des Amplikons, sondern gegen das gesamte Genom aligniert werden. Im Gegensatz zum Alignment gegen die Referenzsequenzen konnten somit die falsch-positiven Alignments zwischen den unterschiedlichen Amplikons stark vermindert werden.

Um weiter ausschließen zu können, dass es sich bei den Epimutationen um technische Artefakte entweder aus der Bisulfitkonvertierung, PCR oder Sequenzierung handelt, wurden alle unvollständig konvertierten Reads nach dem Alignment entfernt. Zudem wurden nur Reads als aberrant methyliert gewertet, die eine Länge von mindestens 100 bp aufwiesen, mindestens 5 CpG-Stellen beinhalteten und in denen keine Sequenzierfehler detektiert werden konnten. Durch diese Punkte kann weitgehend ausgeschlossen werden, dass es sich bei den Epimutationen um technische bzw. bioinformatische Artefakte handelt.

Was jedoch aufgrund der niedrigen EMR an dieser Stelle ebenfalls mit in die Überlegungen einbezogen werden muss, ist die aus der Bisulfitkonvertierung resultierende Sequenzdiversität zwischen methylierten und unmethylierten Allelen. Diese kann zu einem

Amplifikationsbias zum Vorteil des unmethylierten und somit T-reicheren Allels führen [356]. Dies würde wiederum eine Auswirkung auf die Amplifikation der methylierten Allele haben und könnte eine mögliche Erklärung für die niedrigen EMR sein.

Deshalb ist es für Nachfolgeprojekte unabdingbar, die Technik weiterzuentwickeln oder auf eine andere Technik umzusteigen, so dass es möglich ist, PCR-Duplikate bioninformatisch zu erkennen und entfernen zu können. Hierfür bietet sich das sogenannte Single Molecule Tagging an. Über dieses Verfahren ist es möglich, die Ausgangssequenzen vor der Amplifikation mittels Random-Tags zu markieren, sodass nach der Amplifikation alle Sequenzen, die denselben Barcode, tragen auf ein Ausgangsmolekül zurückgeführt werden können. Da es sich bei dieser Methode um eine noch nicht gängige und für bisulfit-konvertierte DNA etablierte Technik handelt, könnte diese in den Nachfolgeprojekten basierend auf den single molecule Molecular Inversion Probes (smMIP) von Hiatt *et al.* (2013) etabliert werden. Die Schwierigkeit bei zielgerichteten Anwendungen besteht darin, die spezifischen Regionen mit den Random Tags zu markieren, ohne gleichzeitig diesen Bereich zu amplifizieren. Um diese Problematik zu umgehen, verwenden Hiatt *et al.* (2013) MIPs, bei denen die zielgerichteten Primer über einen 28 nt-langen Strang verbunden sind, der den Random Tag enthält. Diese Sonden werden an die DNA hybridisiert und die Lücke zwischen den beiden Primern mittels Ligase aufgefüllt. Der dadurch entstandene einzelsträngige Kreis wird anschließend mit einer Exonuklease gespalten und kann dann mittels PCR amplifiziert werden [357].

2.4 Einzel-CpG-Fehler

Basierend auf den niedrigen durchschnittlichen Methylierungswerten von unter 1%, wurden bereits Proben mit einem durchschnittlichen Methylierungslevel zwischen 1% und 3% durch die statistischen Berechnungen als hypermethyliert eingestuft. Diese erhöhten Werte konnten nicht allein über die aberrant methylierten Allele begründet werden. Deshalb wurde zusätzlich ein Epimutationsscreening auf Einzel-CpG-Ebene durchgeführt.

Zur Unterscheidung zwischen normal und abnormal methylierten CpG-Stellen wurde für jedes einzelne CpG ein individueller Schwellenwert berechnet. In der Statistik werden Extremwerte als dreimal den Interquartilbereich (IQB) plus den Wert der 75% Quartile berechnet. Jedoch lässt sich diese Definition in dieser Studie nicht anwenden, um zwischen normal und abnormal klar unterscheiden zu können, da die IQBs sehr geringe Werte

aufweisen. Dies spricht zum einen für eine geringe technische Varianz der einzelnen Assays und zum anderen für eine niedrige biologische Varianz der Promotormethylierung von TSG und weist einmal mehr auf die Wichtigkeit des unmethylierten Status der Promotorregion von TSG hin.

Um dennoch einen Schwellenwert zu definieren, der eine Unterscheidung zwischen stochastischem Rauschen und potentiell relevanten aberrant methylierten CpG-Stellen ermöglicht, wurde der Schwellenwert auf fünfmal den IQB plus den Wert der 75% Quartile erweitert. Des Weiteren flossen in die Berechnung ausschließlich Kontrollproben ohne Epimutation ein.

Am häufigsten traten Einzel-CpG-Hypermethylierungen in der EO BC Kohorte in den Promotorbereichen von *PTEN* (39%), *MLH1* (23%), *TP53* Region 1 (22%) und 2 (17%), sowie *RB1* (17%) auf. In der HR BC Kohorte konnten lediglich acht Patientenproben mit aberrant methylierten CpG-Stellen detektiert werden. Dadurch ergab sich über alle Amplikons hinweg ein signifikanter Unterschied zwischen der Anzahl an betroffenen EO BC-Patientinnen (15%) im Vergleich zu den HR BC-Patientinnen (2%). Zudem waren in Summe mehr BC Proben im Vergleich zu den Kontrollproben betroffen.

Auf Einzel-Allel-Ebene erkennt man deutlich, dass pro Allel/Read immer nur ein bis zwei CpG-Stellen eine aberrante Methylierung zeigen und zufällig über das Allel verteilt sind; deshalb kann man hier von sogenannten „single CpG errors“ sprechen. Interessanterweise waren in *BRCA1* und *RAD51C* jeweils nur eine Patientenprobe und keine der Kontrollproben betroffen. Diese beide Regionen wiesen neben der *TP53* Region 2 die meisten BC und Kontrollproben mit Epimutationen auf.

Die Besonderheit der „single CpG errors“ ist, dass diese meist aus den Analysen ausgeschlossen oder nicht als relevant erachtet werden, da hier von stochastischen Fehlern aufgrund der biologischen Varianz oder Bisulfitkonvertierungsfehlern aufgrund technischer Varianzen ausgegangen wird [265, 358, 359]. Wie die aberrante Methylierung von CpG-Inseln in Promotorbereichen zustandekommt, ist nach wie vor nahezu ungeklärt [360]. Jedoch wird davon ausgegangen, dass aufgrund der Methylierungsmaschinerie nur benachbarte CpG-Stellen eine veränderte Methylierung aufweisen können. Umfangreiche epigenetische Studien haben gezeigt, dass CpG-reiche Regionen wesentlich weniger interindividuelle Variationen in der DNA-Methylierung zeigen als CpG-arme Regionen. Darüber hinaus zeigten die meisten CpG-Inseln relativ ähnliche, niedrige Methylierungswerte

in allen Individuen – mit einigen, zufällig verteilten, methylierten CpGs unter vielen unmethylierten CpG-Stellen [361]. Dieser Fund ist übereinstimmend mit der Ansicht, dass die Dichte der methylierten CpGs und weniger individuelle CpG-Stellen, in *cis*-regulatorischen Bereichen zur An- oder Abschaltung eines Gens führt [283, 362]. Diese Ergebnisse würden darauf schließen lassen, dass es sich bei den „single CpG errors“ ebenfalls rein um stochastische Fehler aufgrund der biologischen Varianz oder Bisulfitkonvertierungsfehlern aufgrund technischer Varianzen handelt; diese können somit keinen Einfluss auf die Brusttumorgenese haben. Jedoch spricht gegen diese These, dass in Summe 33% (119/362) aller analysierten CpG-Stellen in den BC-Patientinnen und nur 4% (13/362) in den Kontrollproben aberrant methylierte CpG-Stellen gezeigt haben. Zudem wurde akribisch bei der Datenanalyse vorgegangen und alle Reads mit unvollständiger Bisulfitkonvertierung konnten entfernt werden. Die BC- und Kontrollproben wurden je Kohorte (EO oder HR, und die zugehörige Kontrollgruppe) einer zeitgleichen Bisulfitkonvertierung auf einer 96-well Platte unterzogen; auch die restlichen Schritte der Library-Präparation wurden für alle Proben einer Kohorte in einem parallelen Ansatz durchgeführt. Pro Kohorte wurden die Daten in zwei Läufen generiert. Durch diesen Versuchsaufbau konnten Batcheffekte so weit wie möglich vermieden werden. Deshalb lassen diese Ergebnisse vermuten, dass bei den EO BC-Patientinnen möglicherweise ein Mechanismus beeinträchtigt ist, der für das Vermeiden oder das Entfernung von Einzel-CpG-Fehlern zuständig ist. Dies könnte zur Folge haben, dass der unmethylierte Status nicht aufrechterhalten werden kann und es zu einer verstärkten Methylierung des Promotorbereichs während der Tumorgenese kommt.

2.5 Kummulative Promotor-Hypermethylierung

Zieht man einen Vergleich zwischen allen untersuchten Genen/Amplikons pro Probe, so fällt auf, dass in einigen Proben mehrere Gene/Amplikons ein aberrantes Methylierungsprofil (Epimutationen oder single CpG error) aufweisen. 50% der HR BC, 14% der EO BC und 18% der Kontrollen zeigten Epimutationen in mehr als einer untersuchten Promotorregion. Interessanterweise konnten bei 82% der Proben, die aberrant methylierte Allele in *RAD51C* trugen und bei 77% in *BRCA1*, in mindestens einem weiteren Amplikon Epimutationen detektiert werden. Eine kummulative Anhäufung von Einzel-CpG-Fehlern traten in 50%, 75%,

und 38% der HR BC-, EO BC-Patientinnen und Kontrollen auf. Dabei waren die Regionen von *PTEN* und *TP53* Region 1 am häufigsten involviert.

Korreliert man nun die Anzahl an Hits je Probe und dem Alter der Patientinnen während der Brustkrebsdiagnose, so findet sich ein signifikant negatives Korrelationsmuster (Spearman $\rho = -0,184$; $p = 0,018$) zwischen der Anzahl an Gene mit „single CpG errors“ pro Probe und der Krebsdiagnose.

Multigenische Promotor-Hypermethylierung tritt bereits früh in entarteten Zellen auf. In einigen Studien wurde eine kummulative Hypermethylierung in TSG bei der Analyse von Brustkrebs- [363, 364] und Gebärmutterhalskarzinomen [365] detektiert, aber nicht in nicht-entartetem Gewebe oder in Blut. Aufgrund dieser Erkenntnisse muss in Betracht gezogen werden, dass die aberranten Methylierungsmuster – sowohl Epimutationen als auch Einzel-CpG-Fehler – aus der Analyse von DNA-Rückständen zirkulierender Tumorzellen (CTCs) und/oder zellfreier Tumor DNA (cfDNA) stammen könnten. 26% (11/43) der EO BC-Patientinnen und 19% (7/36) der HR BC-Patientinnen hatten zum Zeitpunkt der Blutabnahme einen Tumor, jedoch kann hier keine signifikante Korrelation (Spearman $\rho = -0,118$; $p = 0,302$) zwischen dem Tumorstatus und dem Auftreten von Epimutationen und single CpG Fehlern beobachtet werden. Des Weiteren spricht dagegen, dass unter normalen physiologischen Gegebenheiten der Anteil an cfDNA im Blutkreislauf sehr gering ist, da die apoptotischen und nekrotischen Tumorzellen über infiltrierende Phagozyten beseitigt werden. Zudem werden die meisten cfDNA Fragmente vom Tumor via Apoptose entlassen und weisen eine Länge von 180-200 bp auf [366-368]. Die in dieser Studie verwendeten Assays decken jedoch ausschließlich Regionen mit einer Länge von 379-597 bp ab. Solide Tumore können aber auch über Nekrose, Autophagie und andere physiologische Ereignisse cfDNA freisetzen [369, 370]. Diese Fragmente sind länger und könnten über unsere Assays abgedeckt werden. Jedoch wird cfDNA schnell aus dem Blutkreislauf beseitigt und besitzt eine sehr kurze Halbwertszeit von rund 16 min [371]. Eine neuere Studie hat gezeigt, dass die Halbwertszeit von zirkulierender fetaler DNA bei rund 14 Stunden liegt [372]. Aufgrund dieser begrenzten Stabilität der cfDNA muss die Präparation direkt nach der Blutabnahme erfolgen. In unserer Studie lag zwischen der Blutabnahme und der DNA Isolation mindestens ein Tag bis hin zu fünf Tagen. Des Weiteren ist die Analyse von cfDNA sehr schwierig und bedarf hochsensitiver Techniken, um den geringen Anteil an tumorspezifischer DNA unter der wildtypischen cfDNA ausfindig machen zu können [373].

Zirkulierende Tumorzellen (CTCs) treten ebenfalls mit einer sehr geringen Konzentration (1 Tumorzelle:1 Mio. Blutzellen) auf. Deshalb erweist sich die Detektion und Isolation von CTCs als schwierig und erfordert, wie bei der cfDNA, aufwändige Methoden und arbeitsintensive Prozesse, wie beispielsweise hochsensitive und spezifische analytische Methoden und komplexe Anreicherungsprozesse [374, 375]. Besonders Patienten mit einer frühen Form der Krebserkrankung weisen extrem niedrige CTC-Konzentrationen im Blut auf [375]. Ähnlich wie cfDNA zeigen isolierte CTCs eine geringe Stabilität und haben eine durchschnittliche Halbwertszeit von 1-2,4 Stunden [376]. Mit der von uns angewandten DBAS Technik wird zwar eine hohe Sequenziertiefe erreicht, jedoch ist die Probenpräparation nicht ausreichend sensitiv, um CTCs oder cfDNA detektieren zu können. Basierend auf diesen Punkten kann nahezu sicher ausgeschlossen werden, dass CTCs oder cfDNA in unserer Studie analysiert wurden und einen Einfluss auf die gemessenen ER, single CpG Fehler oder/und multigenischen Methylierungsveränderungen hatte.

2.6 Expressionsanalysen

Eine Vielzahl an Studien haben die Hypermethylierung des Promotorbereichs von TSG mit einer verminderten Gentranskription in Verbindung gebracht [377-380]. Deshalb wurde auch der Einfluss der Epimutationen und Einzel-CpG-Fehler auf die Expression der TSG untersucht. Da es sich hier um sehr geringe Effekte handelt, wurde für die Analyse der Expression auf die TaqMan Technologie zurückgegriffen. Es konnten nicht alle Patientenproben, die eine aberrante Methylierung gezeigt haben, untersucht werden, da es sich bei der DNA für die Methylierungsanalysen um Ausschussmaterial nach einer diagnostischen Untersuchung handelte und dabei routinemäßig keine RNA isoliert wird. Um dennoch Expressionsanalysen durchführen zu können, wurden einige Patientinnen ausgewählt und zu einer erneuten Blutabnahme gebeten. Dadurch konnten eine EO BC Probe ohne detektierte Methylierungsveränderungen, eine EO BC und drei HR BC Proben mit niedrigen bis indermediären ER, zwei EO BC Proben mit Einzel-CpG-Fehler und drei gesunde Kontrollproben rekrutiert werden. Interessanterweise konnte in allen untersuchten Patientenproben eine verminderte Expression einiger oder aller analysierten TSG im Vergleich zu den gesunden Kontrollproben gemessen werden. Entgegen den Erwartungen zeigte auch die Patientenprobe ohne veränderte Methylierungsmuster eine verminderte Genexpression um rund 50% bei allen TSG im Vergleich zu den Kontrollproben. Da von allen

Proben die gleiche RNA-Konzentration für die cDNA Synthese eingesetzt wurde und auch die gleiche cDNA Menge für die Real-Time PCR verwendet wurde, kann dies als negativer Faktor ausgeschlossen werden. Zudem wurden die Expressionslevel der Proben in Triplikaten und in einem einzigen Lauf gemessen. Deshalb können hier auch Batcheffekte ausgeschlossen werden. Des Weiteren wurde basierend auf den verwendeten Haushaltsgenen eine Datennormalisierung durchgeführt. Es kann sich auch nicht um Kreuzkontaminationen zwischen den Proben handeln, da für jeden Assay eine Negativkontrolle, in Triplikaten gemessen, in den Lauf inkludiert war. Selbst wenn man die Wahl der Haushaltsgene, *GAPDH* und *ACTB*, kritisch hinterfragt, kann auch hier weitgehend ausgeschlossen werden, dass die Verwendung dieser Haushaltsgene zu einer Verfälschung der Expressionslevel beigetragen hat. Laut den RNAseq Expressionsdaten des GTEx Portals [Int. 32, Int. 33], variieren die Expressionslevel von *ACTB* und *GAPDH* stark zwischen einzelnen Individuen in Gesamtblut. Generell ist es nicht besonders einfach, die Expression von Genen in Blut zu analysieren, da dieses Gewebe sich aus vielen verschiedenen Zelltypen, die sich in unterschiedlichen Stadien ihres Lebenszyklus befinden, zusammensetzt und jeder Zelltyp eine andere Menge an RNA beisteuert. Es gibt viele Studien, die sich mit der Wahl des richtigen Haushaltsgens für Expressionsstudien im Blut beschäftigt haben. Unsere Wahl basiert auf einer Veröffentlichung von Joehanes *et al.* (2013), die *GAPDH* und *ACTB* für die Expressionsstudie von Patienten mit Herz-Kreislauf-Erkrankungen mittels Real-Time TaqMan Technologie verwendet haben [381]. Unter dem Gesichtspunkt der interindividuellen Schwankungen der Expressionslevel dieser Haushaltsgene, hätte diese Varianz auch zwischen den in dieser Studie analysierten Kontrollproben sichtbar sein müssen. Zudem ist auch eine klare Trennung zwischen den Kontrollproben und Patientenproben detektierbar. Dies führt zu dem Schluss, dass *GAPDH* und *ACTB* sich als Haushaltsgene für die Normalisierung dieser Daten eignen. Um diese Frage jedoch abschließend von technischer Seite klären zu können, müssen in Nachfolgeprojekten zum einen mehr Proben und zum anderen, neben *GAPDH* und *ACTB*, weitere Haushaltsgene analysiert werden.

Nichtsdestotrotz weisen die Expressionsdaten, in Kombination mit den Methylierungsdaten, auf die Störung eines übergeordneten Mechanismus hin, der zu einer Prädisposition für mutations-negativen EO und HR BC führt. Es hält sich schon längere Zeit die Theorie, dass bevorzugt DNA-Bereiche mit Methylgruppen versehen werden, die fest auf Nukleosomen aufgewickelt sind. Eine deutsch-amerikanische Studie konnte zeigen, dass bei *Arabidopsis*

der Grad der Methylierung besonders hoch ist, wenn die DNA auf Nukleosomen aufgewickelt vorliegt. Dieser Fund konnte bereits in menschlichen Zellen bestätigt werden. Dies führt zu der Annahme, dass die nukleosomalen Positionen eine wichtige Rolle in der Gestaltung des Methyloms einnehmen. Deshalb ist es wichtig für nachfolgende Experimente, den Fokus nicht nur auf die DNA-Methylierung, sondern auch auf die Chromatinstruktur und insbesondere die Nukleosomen zu legen [382].

2.7 Schlussfolgerung und Ausblick

Zusammenfassend kann gesagt werden, dass die Aufrechterhaltung des unmethylierten Status in TSG einer starken Kontrolle unterliegt. Die durchschnittlichen Methylierungswerte lagen in allen Amplikons bei <1% und auch die ER waren sehr gering (0,001%-0,163%), mit einer Ausnahme in *BRCA1* (ER 14,7%). Ein stärkerer Effekt konnte dagegen bei den Einzel-CpG-Fehlern beobachtet werden, die bei rund einem Drittel der EO BC-Patientinnen in multiplen Genen vorlagen. Hierbei zeigte sich auch eine negative Korrelation zwischen der Häufigkeit an TSG mit Einzel-CpG-Fehlern und dem Alter der Patientinnen bei der Diagnose. Die nachfolgenden Expressionsanalysen ergaben eine erniedrigte Expression multipler TSG je Patient. Die Ergebnisse in ihrer Gesamtheit betrachtet, lassen den Schluss zu, dass es sich hier um die Störung eines übergeordneten Mechanismus handelt, der zu einer Prädisposition für mutations-negativen EO und HR BC führt. Deshalb könnte die Analyse der Methylierungsmuster und Expression der TSG als ein potentieller Biomarker für die Vorhersage des Tumorrisikos angewandt werden.

Da abschließend nicht ausgeschlossen werden kann, dass die Methylierungsveränderungen sowie die Herabregulation der TSG aus der Behandlung der Krebserkrankung resultieren, muss für nachfolgende Studien streng auf die Probenauswahl und das Material geachtet werden.

Deshalb wäre es von Vorteil, in nachfolgenden Studien allem voran eine Zellseparation vorzunehmen, um periphere mononukleäre Blutzellen (engl. *peripheral blood mononuclear cells*, PBMCs) und zirkulierende Tumorzellen (CTCs) vom Gesamtblut zu isolieren. Da zirkulierende Tumorzellen nur eine sehr geringe Halbwertszeit besitzen, könnte hier auch alternativ Brusttumorgewebe verwendet werden. Des Weiteren wäre es wichtig, Proben/Blut vor und nach der Brusttumorbehandlung zu analysieren. Dieses Material kann dann für Methylierungs- und Expressionsstudien eingesetzt werden, um das Ergebnis dieser

Studie zu replizieren und um zu überprüfen, ob der Einfluss der Behandlung des Tumors und der cfDNA bzw. CTCs auf die Ergebnisse ausgeschlossen werden kann. Für die Methylierungsanalysen wäre es wichtig, die DBAS Methode hinsichtlich des Single Molecule Taggings zu erweitern, um oben aufgeführte mögliche Schwankungen zu eliminieren. Es würde sich ebenfalls anbieten, ergänzend zu den Expressionsanalysen Western Blots durchzuführen, um den Einfluss der verminderten Expression auf die Proteinlevel der TSG zu messen. Zudem wäre es interessant, die Chromatinstruktur im Bereich der TSG zu analysieren, um zu sehen, inwieweit hier eine Stilllegung des Chromatins vorliegt. Es wird angenommen, dass Histonmodifikationen und die DNA-Methylierung miteinander agieren und es so zu transkriptioneller Reprimierung der Gene kommen kann [383-385]. Jedoch ist an dieser Stelle nicht klar, welche der beiden Modifikationen die Initiierung übernimmt. Zum einen könnte die DNA-Methylierung als Auslöser fungieren, was zur Modifikation der terminalen Enden der Histone führt und in einer Stilllegung des Chromatins resultiert [383, 386-388]; zum anderen könnte die DNA-Methylierung über das bereits stillgelegte Chromatin induziert werden [130]. Über Chromatin-Immunpräzipitations-(ChIP) Experimente konnte gezeigt werden, dass trimethyliertes H3 K9 (Lysin 9), nicht aber dimethyliertes H3 K9, Chromatinregionen für die Cytosinmethylierung in *N. crassa* markiert [389]. Ähnliche Ergebnisse wurden in *Arabidopsis* gefunden [390]. Dies führt zu der Annahme, dass bei diesen beiden Organismen die DNA-Methylierung downstream der H3 K9 Methylierung wirkt. Basierend auf diesen Ergebnissen wären die Durchführung eines ChIP-Experiments und die Untersuchung der Histonmethylierung bei Brustkrebspatienten von hinreichendem Interesse.

REFERENZEN

1. Literaturangaben

1. *Die Erschaffung der Welt, Genesis 1,1.* Einheitsübersetzung der Heiligen Schrift, Katholische Bibelanstalt Stuttgart, 2016.
2. Schneider, H. and G. Pohl-Apel, *Evolution der Lebewesen.* in Taschenlehrbuch Biologie - Ökologie und Evolution. Thieme Verlag, 2009. **Auflage 1:** p. 273-276.
3. Kutschera, U., *Entdeckungsgeschichte des Abstammungsprinzips und klassische Evolutionsbeweise.* in Lehrbuch Evolutionsbiologie, Eugen Ulmer Verlag, 2015. **Auflage 4:** p. 27-70.
4. Robson, S.L. and B. Wood, *Hominin life history: reconstruction and evolution.* J Anat, 2008. **212**(4): p. 394-425.
5. Clauss, W. and C. Clauss, *Stammesgeschichte des Menschen und Paläogenetik.* in Lehrbuch Humanbiologie kompakt, Spektrum Akademischer Verlag, 2013. **Auflage 5:** p. 2-34.
6. Qiu, J., *How China is rewriting the book on human origins.* Nature, 2016. **535**(7611): p. 218-20.
7. Pohl-Apel, G., *Evolution des Homo sapiens.* in Taschenlehrbuch Biologie - Ökologie und Evolution. Thieme Verlag, 2009. **Auflage 1:** p. 352-367.
8. Bräuer, G., *Der Mensch tritt auf.* in Raumschiff Erde, Spektrum der Wissenschaft Spezial, 2007. **2.**
9. *Die Evolution des Menschen.* Spektrum der Wissenschaft, 2004: p. 45.
10. Brunet, M., et al., *A new hominid from the Upper Miocene of Chad, Central Africa.* Nature, 2002. **418**(6894): p. 145-51.
11. Su, D.F., *The Earliest Hominins: Sahelanthropus, Orrorin, and Ardipithecus.* Nature Education Knowledge, 2013. **4**(4): p. 11.
12. Senut, B., *First hominin from the Miocene (Lukeino Formation, Kenya).* Comptes Rendus Académie de la Terres et des Planètes, 2001(332): p. 137-144.
13. White, T.D., et al., *Ardipithecus ramidus and the paleobiology of early hominids.* Science, 2009. **326**(5949): p. 75-86.
14. Alexeev, V., *The Origin of the Human Race.* Moscow: Progress Publishers. Translation by Campbell H., 1986.
15. Leakey, L.S., P.V. Tobias, and J.R. Napier, *A New Species of the Genus Homo from Olduvai Gorge.* Nature, 1964. **202:** p. 7-9.
16. Johanson, D.C., T.D. White, and Y. Coppens, *A new species of the genus Australopithecus (Primates: Hominidae) from the Pliocene of Eastern Afrika.* Kirtlandia, 1978(28): p. 2-14.
17. Dubois, E., *Pithecanthropus erectus: eine menschenaehnliche Uebergangsform aus Java.* Batavia: Landsdrukkerij, 1894.
18. King, W., *The reputed fossil man of the Neandertal.* Quartely Review of Science, 1864(1): p. 88-97.
19. Rughani, A., *Brain Anatomy.* Medscape, 2015.
20. Florio, M. and W.B. Huttner, *Neural progenitors, neurogenesis and the evolution of the neocortex.* Development, 2014. **141**(11): p. 2182-94.
21. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain.* J Comp Neurol, 2009. **513**(5): p. 532-41.

22. Kindersley, D., *Medizin: Die visuelle Geschichte der Heilkunst*. Dorling Kindersley Deutschland GmbH, 2017.
23. Damasio, A., *The frontal lobes*. Clinical Neuropsychology, 1985: p. 339-375.
24. Semendeferi, K., et al., *Prefrontal cortex in humans and apes: a comparative study of area 10*. Am J Phys Anthropol, 2001. **114**(3): p. 224-41.
25. Christoff, K., et al., *Rostrolateral prefrontal cortex involvement in relational integration during reasoning*. Neuroimage, 2001. **14**(5): p. 1136-49.
26. Holloway, R.L., *Brief communication: how much larger is the relative volume of area 10 of the prefrontal cortex in humans?* Am J Phys Anthropol, 2002. **118**(4): p. 399-401.
27. Jacobs, B., *Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study*. Cerebral Cortex, 2001. **11**(6): p. 558-571.
28. Bonin, G., *Essay on the cerebral cortex*. Thomas, Springfield Ill, 1950.
29. Sowell, E.R., et al., *Longitudinal mapping of cortical thickness and brain growth in normal children*. J Neurosci, 2004. **24**(38): p. 8223-31.
30. Peng, K., et al., *Brodman area 10: Collating, integrating and high level processing of nociception and pain*. Prog Neurobiol, 2017.
31. Lodish, H., A. Berk, and S. Zipursky, *Overview of neuron structure and function*. Molecular Cell Biology, 2000.
32. Campbell, N.A. and J.B. Reece, *Neurone, Synapsen und Signalgebung*. in Campbell Biologie, Pearson Verlag, Auflage 10, 2015: p. 1400-1418.
33. Stiles, J. and T.L. Jernigan, *The basics of brain development*. Neuropsychol Rev, 2010. **20**(4): p. 327-48.
34. Argente-Arizon, P., et al., *Role of non-neuronal cells in body weight and appetite control*. Front Endocrinol (Lausanne), 2015. **6**: p. 42.
35. Kettenmann, H., et al., *Physiology of microglia*. Physiol Rev, 2011. **91**(2): p. 461-553.
36. Aloisi, F., *Immune function of microglia*. Glia, 2001. **36**(2): p. 165-79.
37. Elkabes, S., E.M. DiCicco-Bloom, and I.B. Black, *Brain microglia/macrophages express neurotrophins that selectively regulate microglial proliferation and function*. J Neurosci, 1996. **16**(8): p. 2508-21.
38. Ashwell, K.W., et al., *The appearance and distribution of microglia in the developing retina of the rat*. Vis Neurosci, 1989. **2**(5): p. 437-48.
39. Dalmau, I., et al., *Development of microglia in the postnatal rat hippocampus*. Hippocampus, 1998. **8**(5): p. 458-74.
40. Marin-Teva, J.L., et al., *Microglia promote the death of developing Purkinje cells*. Neuron, 2004. **41**(4): p. 535-47.
41. Peri, F. and C. Nusslein-Volhard, *Live imaging of neuronal degradation by microglia reveals a role for v0-ATPase a1 in phagosomal fusion in vivo*. Cell, 2008. **133**(5): p. 916-27.
42. Svahn, A.J., et al., *Development of ramified microglia from early macrophages in the zebrafish optic tectum*. Dev Neurobiol, 2013. **73**(1): p. 60-71.
43. Thomas, W.E., *Brain macrophages: on the role of pericytes and perivascular cells*. Brain Res Brain Res Rev, 1999. **31**(1): p. 42-57.
44. Grossmann, R., et al., *Juxtavascular microglia migrate along brain microvessels following activation during early postnatal development*. Glia, 2002. **37**(3): p. 229-40.
45. Nimmerjahn, A., F. Kirchhoff, and F. Helmchen, *Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo*. Science, 2005. **308**(5726): p. 1314-8.

46. Batchelor, P.E., et al., *Activated macrophages and microglia induce dopaminergic sprouting in the injured striatum and express brain-derived neurotrophic factor and glial cell line-derived neurotrophic factor*. J Neurosci, 1999. **19**(5): p. 1708-16.
47. Batchelor, P.E., et al., *Macrophages and Microglia Produce Local Trophic Gradients That Stimulate Axonal Sprouting Toward but Not beyond the Wound Edge*. Mol Cell Neurosci, 2002. **21**(3): p. 436-53.
48. Zhong, Y., et al., *The direction of synaptic plasticity mediated by C-fibers in spinal dorsal horn is decided by Src-family kinases in microglia: the role of tumor necrosis factor-alpha*. Brain Behav Immun, 2010. **24**(6): p. 874-80.
49. Allen, N.J. and B.A. Barres, *Neuroscience: Glia - more than just brain glue*. Nature, 2009. **457**(7230): p. 675-7.
50. Kandell, E.R., J.H. Schwartz, and T.M. Jessel, *Principles of neural science*. 4th Edn New York, McGraw-Hill, 2000: p. 19-20.
51. Ullian, E.M., et al., *Control of synapse number by glia*. Science, 2001. **291**(5504): p. 657-61.
52. Doetsch, F., *The glial identity of neural stem cells*. Nat Neurosci, 2003. **6**(11): p. 1127-34.
53. Nishiyama, A., Z. Yang, and A. Butt, *Astrocytes and NG2-glia: what's in a name?* J Anat, 2005. **207**(6): p. 687-93.
54. Noctor, S.C., V. Martinez-Cerdeno, and A.R. Kriegstein, *Contribution of intermediate progenitor cells to cortical histogenesis*. Arch Neurol, 2007. **64**(5): p. 639-42.
55. Jerison, H., *Evolution of the brain and intelligence*. New York: Academic Press, 1973. **43**(1): p. 156-157.
56. Marino, L., *A comparison of encephalization between odontocete cetaceans and anthropoid primates*. Brain Behav Evol, 1998. **51**(4): p. 230-8.
57. Williams, R.W. and K. Herrup, *The control of neuron number*. Annu Rev Neurosci, 1988. **11**: p. 423-53.
58. Jerison, H.J., *Animal intelligence as encephalization*. Philos Trans R Soc Lond B Biol Sci, 1985. **308**(1135): p. 21-35.
59. Marino, L., *Convergence of complex cognitive abilities in cetaceans and primates*. Brain Behav Evol, 2002. **59**(1-2): p. 21-32.
60. MacDonald, D.W., *Dwindling resources and the social behavior of capybaras (Hydrochaeris hydrochaeris) (Mammalia)*. J. Zool. Lond. , 1981. **194**: p. 371-391.
61. Deaner, R.O., et al., *Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates*. Brain Behav Evol, 2007. **70**(2): p. 115-24.
62. Roth, G. and U. Dicke, *Evolution of the brain and intelligence in primates*. Prog Brain Res, 2012. **195**: p. 413-30.
63. Dicke, U. and G. Roth, *Neuronal factors determining high intelligence*. Philos Trans R Soc Lond B Biol Sci, 2016. **371**(1685): p. 20150180.
64. Herculano-Houzel, S., *The human brain in numbers: a linearly scaled-up primate brain*. Front Hum Neurosci, 2009. **3**: p. 31.
65. Roth, G. and U. Dicke, *Evolution of the brain and intelligence*. Trends Cogn Sci, 2005. **9**(5): p. 250-7.
66. Corballis, M.C., *Mirror neurons and the evolution of language*. Brain Lang, 2010. **112**(1): p. 25-35.
67. Friederici, A.D., *The brain basis of language processing: from structure to function*. Physiol Rev, 2011. **91**(4): p. 1357-92.

68. Atkinson, Q.D., *Phonemic diversity supports a serial founder effect model of language expansion from Africa*. *Science*, 2011. **332**(6027): p. 346-9.
69. Waddington, C.H., *The epigenotype. 1942*. *Int J Epidemiol*, 2012. **41**(1): p. 10-3.
70. Waterland, R.A. and R.L. Jirtle, *Transposable elements: targets for early nutritional effects on epigenetic gene regulation*. *Mol Cell Biol*, 2003. **23**(15): p. 5293-300.
71. Chong, S. and E. Whitelaw, *Epigenetic germline inheritance*. *Curr Opin Genet Dev*, 2004. **14**(6): p. 692-6.
72. Brooks, J., P. Cairns, and A. Zeleniuch-Jacquotte, *Promoter methylation and the detection of breast cancer*. *Cancer Causes Control*, 2009. **20**(9): p. 1539-50.
73. Hagood, J.S., *Beyond the genome: epigenetic mechanisms in lung remodeling*. *Physiology (Bethesda)*, 2014. **29**(3): p. 177-85.
74. Tollefsbol, T., *Epigenetics: The new science of genetics*. in *Handbook of Epigenetics. The New Molecular and Medical Genetics*, hg. Academic Press, 1, 1-6, 2010.
75. Lai, S.R., et al., *Epigenetic control of telomerase and modes of telomere maintenance in aging and abnormal systems*. *Front Biosci*, 2005. **10**: p. 1779-96.
76. Fu, Y., et al., *N6-methyldeoxyadenosine marks active transcription start sites in Chlamydomonas*. *Cell*, 2015. **161**(4): p. 879-892.
77. Zhang, G., et al., *N6-methyladenine DNA modification in Drosophila*. *Cell*, 2015. **161**(4): p. 893-906.
78. Greer, E.L., et al., *DNA Methylation on N6-Adenine in C. elegans*. *Cell*, 2015. **161**(4): p. 868-78.
79. Lindroth, A.M., et al., *Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation*. *Science*, 2001. **292**(5524): p. 2077-80.
80. Stansbury, M.S. and A.P. Moczek, *The function of Hox and appendage-patterning genes in the development of an evolutionary novelty, the Photuris firefly lantern*. *Proc Biol Sci*, 2014. **281**(1782): p. 20133333.
81. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences*. *Nature*, 2009. **462**(7271): p. 315-22.
82. Esteller, M., *Cancer epigenomics: DNA methylomes and histone-modification maps*. *Nat Rev Genet*, 2007. **8**(4): p. 286-98.
83. Goll, M.G. and T.H. Bestor, *Eukaryotic cytosine methyltransferases*. *Annu Rev Biochem*, 2005. **74**: p. 481-514.
84. Roje, S., *S-Adenosyl-L-methionine: beyond the universal methyl group donor*. *Phytochemistry*, 2006. **67**(15): p. 1686-98.
85. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development*. *Cell*, 1999. **99**(3): p. 247-57.
86. Wu, H. and Y. Zhang, *Reversing DNA methylation: mechanisms, genomics, and biological functions*. *Cell*, 2014. **156**(1-2): p. 45-68.
87. Ito, S., et al., *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine*. *Science*, 2011. **333**(6047): p. 1300-3.
88. Ambrosi, C., M. Manzo, and T. Baubec, *Dynamics and Context-Dependent Roles of DNA Methylation*. *J Mol Biol*, 2017. **429**(10): p. 1459-1475.
89. Edwards, J.R., et al., *Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns*. *Genome Res*, 2010. **20**(7): p. 972-80.
90. Holliday, R. and J.E. Pugh, *DNA modification mechanisms and gene activity during development*. *Science*, 1975. **187**(4173): p. 226-32.
91. Riggs, A.D., *X inactivation, differentiation, and DNA methylation*. *Cytogenet Cell Genet*, 1975. **14**(1): p. 9-25.

92. Robertson, K.D., *DNA methylation and chromatin - unraveling the tangled web*. Oncogene, 2002. **21**(35): p. 5361-79.
93. Yang, X., et al., *Gene body methylation can alter gene expression and is a therapeutic target in cancer*. Cancer Cell, 2014. **26**(4): p. 577-90.
94. Borgel, J., et al., *Targets and dynamics of promoter DNA methylation during early mouse development*. Nat Genet, 2010. **42**(12): p. 1093-100.
95. Beard, C., E. Li, and R. Jaenisch, *Loss of methylation activates Xist in somatic but not in embryonic cells*. Genes Dev, 1995. **9**(19): p. 2325-34.
96. Li, E., C. Beard, and R. Jaenisch, *Role for DNA methylation in genomic imprinting*. Nature, 1993. **366**(6453): p. 362-5.
97. Kaneda, M., et al., *Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting*. Nature, 2004. **429**(6994): p. 900-3.
98. Jahner, D., et al., *De novo methylation and expression of retroviral genomes during mouse embryogenesis*. Nature, 1982. **298**(5875): p. 623-8.
99. Walsh, C.P., J.R. Chaillet, and T.H. Bestor, *Transcription of IAP endogenous retroviruses is constrained by cytosine methylation*. Nat Genet, 1998. **20**(2): p. 116-7.
100. Bourc'his, D. and T.H. Bestor, *Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L*. Nature, 2004. **431**(7004): p. 96-9.
101. Bestor, T.H., J.R. Edwards, and M. Boulard, *Notes on the role of dynamic DNA methylation in mammalian development*. Proc Natl Acad Sci U S A, 2015. **112**(22): p. 6796-9.
102. Ioshikhes, I.P. and M.Q. Zhang, *Large-scale human promoter mapping using CpG islands*. Nat Genet, 2000. **26**(1): p. 61-3.
103. Edwards, J.R., et al., *DNA methylation and DNA methyltransferases*. Epigenetics Chromatin, 2017. **10**: p. 23.
104. Josse, J., A.D. Kaiser, and A. Kornberg, *Enzymatic synthesis of deoxyribonucleic acid. VIII. Frequencies of nearest neighbor base sequences in deoxyribonucleic acid*. J Biol Chem, 1961. **236**: p. 864-75.
105. Kornberg, A., et al., *Enzymatic Synthesis of Deoxyribonucleic Acid, Xvi. Oligonucleotides as Templates and the Mechanism of Their Replication*. Proc Natl Acad Sci U S A, 1964. **51**: p. 315-23.
106. Coulondre, C., et al., *Molecular basis of base substitution hotspots in Escherichia coli*. Nature, 1978. **274**(5673): p. 775-80.
107. Cooper, D.N., et al., *Methylation-mediated deamination of 5-methylcytosine appears to give rise to mutations causing human inherited disease in CpNpG trinucleotides, as well as in CpG dinucleotides*. Hum Genomics, 2010. **4**(6): p. 406-10.
108. Chimpanzee, S. and C. Analysis, *Initial sequence of the chimpanzee genome and comparison with the human genome*. Nature, 2005. **437**(7055): p. 69-87.
109. Doi, A., et al., *Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts*. Nat Genet, 2009. **41**(12): p. 1350-3.
110. Reik, W., *Stability and flexibility of epigenetic gene regulation in mammalian development*. Nature, 2007. **447**(7143): p. 425-32.
111. Weber, M., et al., *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells*. Nat Genet, 2005. **37**(8): p. 853-62.

112. Cohen, N.M., E. Kenigsberg, and A. Tanay, *Primate CpG islands are maintained by heterogeneous evolutionary regimes involving minimal selection*. Cell, 2011. **145**(5): p. 773-86.
113. Banovich, N.E., et al., *Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels*. PLoS Genet, 2014. **10**(9): p. e1004663.
114. Shoemaker, R., et al., *Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome*. Genome Res, 2010. **20**(7): p. 883-9.
115. Suzuki, S., et al., *The evolution of mammalian genomic imprinting was accompanied by the acquisition of novel CpG islands*. Genome Biol Evol, 2011. **3**: p. 1276-83.
116. Kanber, D., et al., *The human retinoblastoma gene is imprinted*. PLoS Genet, 2009. **5**(12): p. e1000790.
117. Ladd-Acosta, C., et al., *DNA methylation signatures within the human brain*. Am J Hum Genet, 2007. **81**(6): p. 1304-15.
118. Varki, A. and T.K. Altheide, *Comparing the human and chimpanzee genomes: searching for needles in a haystack*. Genome Res, 2005. **15**(12): p. 1746-58.
119. Rhesus Macaque Genome, S., et al., *Evolutionary and biomedical insights from the rhesus macaque genome*. Science, 2007. **316**(5822): p. 222-34.
120. Khaitovich, P., et al., *Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees*. Science, 2005. **309**(5742): p. 1850-4.
121. Caceres, M., et al., *Elevated gene expression levels distinguish human from non-human primate brains*. Proc Natl Acad Sci U S A, 2003. **100**(22): p. 13030-5.
122. Tsankova, N., et al., *Epigenetic regulation in psychiatric disorders*. Nat Rev Neurosci, 2007. **8**(5): p. 355-67.
123. Graff, J. and I.M. Mansuy, *Epigenetic codes in cognition and behaviour*. Behav Brain Res, 2008. **192**(1): p. 70-87.
124. Levenson, J.M. and J.D. Sweatt, *Epigenetic mechanisms in memory formation*. Nat Rev Neurosci, 2005. **6**(2): p. 108-18.
125. Tognini, P., D. Napoli, and T. Pizzorusso, *Dynamic DNA methylation in the brain: a new epigenetic mark for experience-dependent plasticity*. Front Cell Neurosci, 2015. **9**: p. 331.
126. McClung, C.A. and E.J. Nestler, *Neuroplasticity mediated by altered gene expression*. Neuropsychopharmacology, 2008. **33**(1): p. 3-17.
127. Schanen, N.C., *Epigenetics of autism spectrum disorders*. Hum Mol Genet, 2006. **15 Spec No 2**: p. R138-50.
128. LaSalle, J.M. and D.H. Yasui, *Evolving role of MeCP2 in Rett syndrome and autism*. Epigenomics, 2009. **1**(1): p. 119-30.
129. Nicolia, V., et al., *DNA Methylation Profiles of Selected Pro-Inflammatory Cytokines in Alzheimer Disease*. J Neuropathol Exp Neurol, 2017. **76**(1): p. 27-31.
130. Bird, A., *DNA methylation patterns and epigenetic memory*. Genes Dev, 2002. **16**(1): p. 6-21.
131. Robertson, K.D. and A.P. Wolffe, *DNA methylation in health and disease*. Nat Rev Genet, 2000. **1**(1): p. 11-9.
132. Urdinguio, R.G., J.V. Sanchez-Mut, and M. Esteller, *Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies*. Lancet Neurol, 2009. **8**(11): p. 1056-72.
133. Irwin, S.A., R. Galvez, and W.T. Greenough, *Dendritic spine structural anomalies in fragile-X mental retardation syndrome*. Cereb Cortex, 2000. **10**(10): p. 1038-44.

134. Huber, K.M., et al., *Altered synaptic plasticity in a mouse model of fragile X mental retardation*. Proc Natl Acad Sci U S A, 2002. **99**(11): p. 7746-50.
135. Eccleston, A., et al., *Epigenetics*. Nature, 2007. **447**: p. 395.
136. Robertson, K.D., *DNA methylation and human disease*. Nat Rev Genet, 2005. **6**(8): p. 597-610.
137. Feinberg, A.P. and B. Vogelstein, *Hypomethylation distinguishes genes of some human cancers from their normal counterparts*. Nature, 1983. **301**(5895): p. 89-92.
138. Feinberg, A.P. and B. Vogelstein, *Hypomethylation of ras oncogenes in primary human cancers*. Biochem Biophys Res Commun, 1983. **111**(1): p. 47-54.
139. Yoder, J.A., C.P. Walsh, and T.H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites*. Trends Genet, 1997. **13**(8): p. 335-40.
140. Widschwendter, M., et al., *DNA hypomethylation and ovarian cancer biology*. Cancer Res, 2004. **64**(13): p. 4472-80.
141. Song, F., et al., *Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression*. Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3336-41.
142. Issa, J.P., *CpG island methylator phenotype in cancer*. Nat Rev Cancer, 2004. **4**(12): p. 988-93.
143. Chan, A.O., et al., *CpG island methylation in aberrant crypt foci of the colorectum*. Am J Pathol, 2002. **160**(5): p. 1823-30.
144. Varela-Rey, M., et al., *Alcohol, DNA methylation, and cancer*. Alcohol Res, 2013. **35**(1): p. 25-35.
145. Potter, V.R., *Initiation and promotion in cancer formation: the importance of studies on intercellular communication*. Yale J Biol Med, 1980. **53**(5): p. 367-84.
146. Weinberg, R.A., *Cancer: a genetic disorder*. in The molecular basis of cancer. Elsevier, 2008: p. 3-16.
147. Riley, B. and D. Anderson, *Cancer epigenetics*. in Handbook of Epigenetics. The New Molecular and Medical Genetics, hg. Academic Press, 1, 2010: p. 521-533.
148. Stehelin, D., et al., *DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA*. Nature, 1976. **260**(5547): p. 170-3.
149. Reddy, E.P., et al., *A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene*. Nature, 1982. **300**(5888): p. 149-52.
150. Ashkenazi, R., S.N. Gentry, and T.L. Jackson, *Pathways to tumorigenesis--modeling mutation acquisition in stem cells and their progeny*. Neoplasia, 2008. **10**(11): p. 1170-82.
151. Hahn, W.C. and R.A. Weinberg, *Modelling the molecular circuitry of cancer*. Nat Rev Cancer, 2002. **2**(5): p. 331-41.
152. Vessey, C.J., C.J. Norbury, and I.D. Hickson, *Genetic disorders associated with cancer predisposition and genomic instability*. Prog Nucleic Acid Res Mol Biol, 1999. **63**: p. 189-221.
153. Duesberg, P. and D. Rasnick, *Aneuploidy, the somatic mutation that makes cancer a species of its own*. Cell Motil Cytoskeleton, 2000. **47**(2): p. 81-107.
154. Li, W., et al., *Tumor Angiogenesis as a Target for Dietary Cancer Prevention*. Journal of Oncology, 2012. **2012**.
155. Beliveau, R. and D. Gingras, *Krebszellen mögen keine Himbeeren*. Wilhelm Goldmann Verlag, München, 2010.
156. Holinski-Feder, E., *Genetik von Krebserkrankungen* in Taschenlehrbuch der Humangenetik. Thieme Verlag, 2006. **Auflage 8**: p. 513-541.

157. Wanpeng, S. and Y. Jian, *Functional mechanism for human tumor suppressors*. Journal of cancer, 2010. **1**(1): p. 136-140.
158. Colditz, G.A. and K. Bohlke, *Priorities for the primary prevention of breast cancer*. CA Cancer J Clin, 2014. **64**(3): p. 186-94.
159. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets*. Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
160. Calza, S., et al., *Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients*. Breast Cancer Res, 2006. **8**(4): p. R34.
161. Hu, Z., et al., *The molecular portraits of breast tumors are conserved across microarray platforms*. BMC Genomics, 2006. **7**: p. 96.
162. Cancer Genome Atlas, N., *Comprehensive molecular portraits of human breast tumours*. Nature, 2012. **490**(7418): p. 61-70.
163. Bertucci, F., et al., *Identification and validation of an ERBB2 gene expression signature in breast cancers*. Oncogene, 2004. **23**(14): p. 2564-75.
164. Biswas, D.K. and J.D. Iglehart, *Linkage between EGFR family receptors and nuclear factor kappaB (NF-kappaB) signaling in breast cancer*. J Cell Physiol, 2006. **209**(3): p. 645-52.
165. Foulkes, W.D., et al., *Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer*. J Natl Cancer Inst, 2003. **95**(19): p. 1482-5.
166. Sotiriou, C., et al., *Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis*. J Natl Cancer Inst, 2006. **98**(4): p. 262-72.
167. Rakha, E.A., J.S. Reis-Filho, and I.O. Ellis, *Basal-like breast cancer: a critical review*. J Clin Oncol, 2008. **26**(15): p. 2568-81.
168. Widschwendter, M. and P.A. Jones, *DNA methylation and breast carcinogenesis*. Oncogene, 2002. **21**(35): p. 5462-82.
169. Lee, E.Y. and W.J. Muller, *Oncogenes and tumor suppressor genes*. Cold Spring Harb Perspect Biol, 2010. **2**(10): p. a003236.
170. Collaborative Group on Hormonal Factors in Breast, C., *Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease*. Lancet, 2001. **358**(9291): p. 1389-99.
171. Pharoah, P.D., et al., *Familial predisposition to breast cancer in a British population: implications for prevention*. Eur J Cancer, 2000. **36**(6): p. 773-9.
172. Mavaddat, N., et al., *Genetic susceptibility to breast cancer*. Mol Oncol, 2010. **4**(3): p. 174-91.
173. Easton, D.F., et al., *Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium*. Am J Hum Genet, 1993. **52**(4): p. 678-701.
174. Turnbull, C. and N. Rahman, *Genetic predisposition to breast cancer: past, present, and future*. Annu Rev Genomics Hum Genet, 2008. **9**: p. 321-45.
175. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21*. Science, 1990. **250**(4988): p. 1684-9.
176. Ford, D., et al., *Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium*. Am J Hum Genet, 1998. **62**(3): p. 676-89.
177. King, M.C., et al., *Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2*. Science, 2003. **302**(5645): p. 643-6.

178. Garber, J.E., et al., *Follow-up study of twenty-four families with Li-Fraumeni syndrome*. *Cancer Res*, 1991. **51**(22): p. 6094-7.
179. Nelen, M.R., et al., *Localization of the gene for Cowden disease to chromosome 10q22-23*. *Nat Genet*, 1996. **13**(1): p. 114-6.
180. Hemminki, A., et al., *A serine/threonine kinase gene defective in Peutz-Jeghers syndrome*. *Nature*, 1998. **391**(6663): p. 184-7.
181. Thompson, D. and D. Easton, *The genetic epidemiology of breast cancer genes*. *J Mammary Gland Biol Neoplasia*, 2004. **9**(3): p. 221-36.
182. Ghossaini, M. and P.D. Pharoah, *Polygenic susceptibility to breast cancer: current state-of-the-art*. *Future Oncol*, 2009. **5**(5): p. 689-701.
183. Rahman, N., et al., *PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene*. *Nat Genet*, 2007. **39**(2): p. 165-7.
184. Meijers-Heijboer, H., et al., *Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations*. *Nat Genet*, 2002. **31**(1): p. 55-9.
185. Renwick, A., et al., *ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles*. *Nat Genet*, 2006. **38**(8): p. 873-5.
186. Seal, S., et al., *Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles*. *Nat Genet*, 2006. **38**(11): p. 1239-41.
187. Antoniou, A.C., et al., *A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes*. *Br J Cancer*, 2002. **86**(1): p. 76-83.
188. Antoniou, A.C., et al., *The BOADICEA model of genetic susceptibility to breast and ovarian cancer*. *Br J Cancer*, 2004. **91**(8): p. 1580-90.
189. Cox, A., et al., *A common coding variant in CASP8 is associated with breast cancer risk*. *Nat Genet*, 2007. **39**(3): p. 352-8.
190. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci*. *Nature*, 2007. **447**(7148): p. 1087-93.
191. Gold, B., et al., *Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33*. *Proc Natl Acad Sci U S A*, 2008. **105**(11): p. 4340-5.
192. Hunter, D.J., et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer*. *Nat Genet*, 2007. **39**(7): p. 870-4.
193. Stacey, S.N., et al., *Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer*. *Nat Genet*, 2007. **39**(7): p. 865-9.
194. Ahmed, S., et al., *Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2*. *Nat Genet*, 2009. **41**(5): p. 585-90.
195. Zheng, W., et al., *Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1*. *Nat Genet*, 2009. **41**(3): p. 324-8.
196. Smith, P., et al., *A genome wide linkage search for breast cancer susceptibility genes*. *Genes Chromosomes Cancer*, 2006. **45**(7): p. 646-55.
197. Easton, D.F., *How many more breast cancer predisposition genes are there?* *Breast Cancer Res*, 1999. **1**(1): p. 14-7.
198. Szpirer, C. and J. Szpirer, *Mammary cancer susceptibility: human genes and rodent models*. *Mamm Genome*, 2007. **18**(12): p. 817-31.
199. Chial, H., *Tumor suppressor genes and the two-hit hypothesis*. *Nature Education*, 2008. **1**(1): p. 177.

200. Herman, J.G. and S.B. Baylin, *Gene silencing in cancer in association with promoter hypermethylation*. N Engl J Med, 2003. **349**(21): p. 2042-54.
201. Martin, D.I., R. Ward, and C.M. Suter, *Germline epimutation: A basis for epigenetic disease in humans*. Ann N Y Acad Sci, 2005. **1054**: p. 68-77.
202. Narayan, A., et al., *Hypomethylation of pericentromeric DNA in breast adenocarcinomas*. Int J Cancer, 1998. **77**(6): p. 833-8.
203. Feinberg, A.P., et al., *Reduced genomic 5-methylcytosine content in human colonic neoplasia*. Cancer Res, 1988. **48**(5): p. 1159-61.
204. Szyf, M., *DNA methylation and cancer therapy*. Drug Resist Updat, 2003. **6**(6): p. 341-53.
205. Szyf, M., *Targeting DNA methylation in cancer*. Bull Cancer, 2006. **93**(9): p. 961-72.
206. Szyf, M., P. Pakneshan, and S.A. Rabbani, *DNA methylation and breast cancer*. Biochem Pharmacol, 2004. **68**(6): p. 1187-97.
207. Huang, T.H., et al., *Identification of DNA methylation markers for human breast carcinomas using the methylation-sensitive restriction fingerprinting technique*. Cancer Res, 1997. **57**(6): p. 1030-4.
208. Huang, T.H., M.R. Perry, and D.E. Laux, *Methylation profiling of CpG islands in human breast cancer cells*. Hum Mol Genet, 1999. **8**(3): p. 459-70.
209. Herman, J.G., et al., *Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers*. Cancer Res, 1995. **55**(20): p. 4525-30.
210. Silva, J., et al., *Concomitant expression of p16INK4a and p14ARF in primary breast cancer and analysis of inactivation mechanisms*. J Pathol, 2003. **199**(3): p. 289-97.
211. Rice, J.C., et al., *Methylation of the BRCA1 promoter is associated with decreased BRCA1 mRNA levels in clinical breast cancer specimens*. Carcinogenesis, 2000. **21**(9): p. 1761-5.
212. Murata, H., et al., *Genetic and epigenetic modification of mismatch repair genes hMSH2 and hMLH1 in sporadic breast cancer with microsatellite instability*. Oncogene, 2002. **21**(37): p. 5696-703.
213. Piva, R., et al., *Different methylation of oestrogen receptor DNA in human breast carcinomas with and without oestrogen receptor*. Br J Cancer, 1990. **61**(2): p. 270-5.
214. Sirchia, S.M., et al., *Evidence of epigenetic changes affecting the chromatin state of the retinoic acid receptor beta2 promoter in breast cancer cells*. Oncogene, 2000. **19**(12): p. 1556-63.
215. Widschwendter, M., et al., *Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer*. J Natl Cancer Inst, 2000. **92**(10): p. 826-32.
216. Graff, J.R., et al., *Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression*. J Biol Chem, 2000. **275**(4): p. 2727-32.
217. Bachman, K.E., et al., *Methylation-associated silencing of the tissue inhibitor of metalloproteinase-3 gene suggest a suppressor role in kidney, brain, and other human cancers*. Cancer Res, 1999. **59**(4): p. 798-802.
218. Alves, G., A. Tatro, and T. Fanning, *Differential methylation of human LINE-1 retrotransposons in malignant cells*. Gene, 1996. **176**(1-2): p. 39-44.
219. Weber, J., et al., *Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine*. Cancer Res, 1994. **54**(7): p. 1766-71.
220. De Smet, C., et al., *The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation*. Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7149-53.

221. Kavalari, R., et al., *Expression of MAGE tumour-associated antigens is inversely correlated with tumour differentiation in invasive ductal breast cancers: an immunohistochemical study*. *Virchows Arch*, 2001. **439**(2): p. 127-31.
222. Fisher, J.L., et al., *Urokinase plasminogen activator system gene expression is increased in human breast carcinoma and its bone metastases--a comparison of normal breast tissue, non-invasive and invasive carcinoma and osseous metastases*. *Breast Cancer Res Treat*, 2000. **61**(1): p. 1-12.
223. Lu, A., et al., *Molecular mechanisms for aberrant expression of the human breast cancer specific gene 1 in breast cancer cells: control of transcription by DNA methylation and intronic sequences*. *Oncogene*, 2001. **20**(37): p. 5173-85.
224. Gupta, A., et al., *Hypomethylation of the synuclein gamma gene CpG island promotes its aberrant expression in breast carcinoma and ovarian carcinoma*. *Cancer Res*, 2003. **63**(3): p. 664-73.
225. Girault, I., et al., *Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas*. *Clin Cancer Res*, 2003. **9**(12): p. 4415-22.
226. Nass, S.J., et al., *Expression of DNA methyl-transferase (DMT) and the cell cycle in human breast cancer cells*. *Oncogene*, 1999. **18**(52): p. 7453-61.
227. Szyf, M., *Towards a pharmacology of DNA methylation*. *Trends Pharmacol Sci*, 2001. **22**(7): p. 350-4.
228. Coleman, M.P., et al., *Cancer survival in five continents: a worldwide population-based study (CONCORD)*. *Lancet Oncol*, 2008. **9**(8): p. 730-56.
229. Matevosian, A. and S. Akbarian, *Neuronal nuclei isolation from human postmortem brain tissue*. *J Vis Exp*, 2008(20).
230. Zhang, Y. and A. Jeltsch, *The application of next generation sequencing in DNA methylation analysis*. *Genes (Basel)*, 2010. **1**(1): p. 85-101.
231. Shapiro, R., R.E. Servis and M. Welcher, *Reactions of Uracil and Cytosine Derivatives with Sodium Bisulfite. A Specific Deamination Method*. *Journal of the American Chemical Society*, 1970. **92**(2): p. 85-101.
232. Hayatsu, H., Y. Wataya, and K. Kazushige, *The addition of sodium bisulfite to uracil and to cytosine*. *J Am Chem Soc*, 1970. **92**(3): p. 724-6.
233. Hayatsu, H., *Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis--a personal account*. *Proc Jpn Acad Ser B Phys Biol Sci*, 2008. **84**(8): p. 321-30.
234. Krueger, F., et al., *DNA methylome analysis using short bisulfite sequencing data*. *Nat Methods*, 2012. **9**(2): p. 145-51.
235. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads*. *EMBnet.journal*, 2011(17): p. 10-12.
236. Krueger, F. and S.R. Andrews, *Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications*. *Bioinformatics*, 2011. **27**(11): p. 1571-2.
237. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2*. *Nat Methods*, 2012. **9**(4): p. 357-9.
238. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. *Bioinformatics*, 2009. **25**(16): p. 2078-9.
239. Kuhn, R.M., D. Haussler, and W.J. Kent, *The UCSC genome browser and associated tools*. *Brief Bioinform*, 2013. **14**(2): p. 144-61.
240. Legendre, P. and L. Legendre, *Numerical Ecology, Vol. 24 (Developments in Environmental Modelling)*. Elsevier Science, 1998.
241. Oksanen, J., et al., *Vegan: Community Ecology Package*. 2017.

242. Feng, H., K.N. Conneely, and H. Wu, *A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data*. *Nucleic Acids Res*, 2014. **42**(8): p. e69.
243. Park, Y. and H. Wu, *Differential methylation analysis for BS-seq data under general experimental design*. *Bioinformatics*, 2016. **32**(10): p. 1446-53.
244. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate - a practical and powerful approach to multiple testing*. *JR Statist Soc B*, 1995. **57**: p. 298-300.
245. Kinsella, R.J., et al., *Ensembl BioMart: a hub for data retrieval across taxonomic space*. *Database (Oxford)*, 2011. **2011**: p. bar030.
246. Lawrence, M., et al., *Software for computing and annotating genomic ranges*. *PLoS Comput Biol*, 2013. **9**(8): p. e1003118.
247. Surrallés, J., et al., *Accelerated telomere shortening in the human inactive X chromosome*. *Am J Hum Genet*, 1999. **65**(6): p. 1617-22.
248. Ding, Z., et al., *Estimating telomere length from whole genome sequence data*. *Nucleic Acids Res*, 2014. **42**(9): p. e75.
249. Choudhury, S.R., et al., *Optogenetic regulation of site-specific subtelomeric DNA methylation*. *Oncotarget*, 2016. **7**(31): p. 50380-50391.
250. Rödel, R. and R.A. Fisher, *Statistical methods for research workers*. 14. Aufl., Oliver & Boyd, Edinburgh, London 1970. XIII, 362 S., 12 Abb., 74 Tab., 40 s. *Biometrische Zeitschrift* 1971: p. 13429-30.
251. Liu, X., et al., *Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques*. *Genome Res*, 2012. **22**(4): p. 611-22.
252. Florio, M., et al., *Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion*. *Science*, 2015. **347**(6229): p. 1465-70.
253. Schneider, E., et al., *CpG sites with continuously increasing or decreasing methylation from early to late human fetal brain development*. *Gene*, 2016. **592**(1): p. 110-118.
254. Nowick, K., et al., *Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain*. *Proc Natl Acad Sci U S A*, 2009. **106**(52): p. 22358-63.
255. Shulha, H.P., et al., *Human-specific histone methylation signatures at transcription start sites in prefrontal neurons*. *PLoS Biol*, 2012. **10**(11): p. e1001427.
256. Basu, S.N., R. Kollu, and S. Banerjee-Basu, *AutDB: a gene reference resource for autism research*. *Nucleic Acids Res*, 2009. **37**(Database issue): p. D832-6.
257. Schizophrenia Working Group of the Psychiatric Genomics, C., *Biological insights from 108 schizophrenia-associated genetic loci*. *Nature*, 2014. **511**(7510): p. 421-7.
258. Bertram, L., et al., *Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database*. *Nat Genet*, 2007. **39**(1): p. 17-23.
259. Fresno, C. and E.A. Fernandez, *RDAVIDWebService: a versatile R interface to DAVID*. *Bioinformatics*, 2013. **29**(21): p. 2810-1.
260. Jones, P.A., *Functions of DNA methylation: islands, start sites, gene bodies and beyond*. *Nat Rev Genet*, 2012. **13**(7): p. 484-92.
261. Lee, H.E., S. Ayarpadikannan, and H.S. Kim, *Role of transposable elements in genomic rearrangement, evolution, gene regulation and epigenetics in primates*. *Genes Genet Syst*, 2015. **90**(5): p. 245-57.
262. Gopalakrishnan, S., B.O. Van Emburgh, and K.D. Robertson, *DNA methylation in development and human disease*. *Mutat Res*, 2008. **647**(1-2): p. 30-8.
263. Singer, T., et al., *LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes?* *Trends Neurosci*, 2010. **33**(8): p. 345-54.

264. Bock, J., et al., *Single CpG hypermethylation, allele methylation errors, and decreased expression of multiple tumor suppressor genes in normal body cells of mutation-negative early-onset and high-risk breast cancer patients*. *Int J Cancer*, 2018.
265. Hansmann, T., et al., *Constitutive promoter methylation of BRCA1 and RAD51C in patients with familial ovarian cancer and early-onset sporadic breast cancer*. *Hum Mol Genet*, 2012. **21**(21): p. 4669-79.
266. Greger, V., et al., *Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma*. *Hum Genet*, 1989. **83**(2): p. 155-8.
267. Graff, J.R., et al., *E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas*. *Cancer Res*, 1995. **55**(22): p. 5195-9.
268. Kuhtz, J., et al., *Epigenetic heterogeneity of developmentally important genes in human sperm: implications for assisted reproduction outcome*. *Epigenetics*, 2014. **9**(12): p. 1648-58.
269. Hatada, I., et al., *Astrocyte-specific genes are generally demethylated in neural precursor cells prior to astrocytic differentiation*. *PLoS One*, 2008. **3**(9): p. e3189.
270. Ma, D.K., et al., *Epigenetic choreographers of neurogenesis in the adult mammalian brain*. *Nat Neurosci*, 2010. **13**(11): p. 1338-44.
271. Takizawa, T., et al., *DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain*. *Dev Cell*, 2001. **1**(6): p. 749-58.
272. Guintivano, J., M.J. Aryee, and Z.A. Kaminsky, *A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression*. *Epigenetics*, 2013. **8**(3): p. 290-302.
273. Feng, J., et al., *Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons*. *Nat Neurosci*, 2010. **13**(4): p. 423-30.
274. Miller, C.A. and J.D. Sweatt, *Covalent modification of DNA regulates memory formation*. *Neuron*, 2007. **53**(6): p. 857-69.
275. Zeng, J., et al., *Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution*. *Am J Hum Genet*, 2012. **91**(3): p. 455-65.
276. Mendizabal, I., et al., *Comparative Methylome Analyses Identify Epigenetic Regulatory Loci of Human Brain Evolution*. *Mol Biol Evol*, 2016. **33**(11): p. 2947-2959.
277. Farcas, R., et al., *Differences in DNA methylation patterns and expression of the CCRK gene in human and nonhuman primate cortices*. *Mol Biol Evol*, 2009. **26**(6): p. 1379-89.
278. Shi, L., Q. Lin, and B. Su, *Human-specific hypomethylation of CENPJ, a key brain size regulator*. *Mol Biol Evol*, 2014. **31**(3): p. 594-604.
279. Davies, M.N., et al., *Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood*. *Genome Biol*, 2012. **13**(6): p. R43.
280. Gibbs, J.R., et al., *Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain*. *PLoS Genet*, 2010. **6**(5): p. e1000952.
281. Ghosh, S., et al., *Tissue specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues*. *Epigenetics*, 2010. **5**(6): p. 527-38.
282. Keshet, I., et al., *Evidence for an instructive mechanism of de novo methylation in cancer cells*. *Nat Genet*, 2006. **38**(2): p. 149-53.
283. Weber, M., et al., *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome*. *Nat Genet*, 2007. **39**(4): p. 457-66.

284. Illingworth, R., et al., *A novel CpG island set identifies tissue-specific methylation at developmental gene loci*. PLoS Biol, 2008. **6**(1): p. e22.
285. Rakan, V.K., et al., *An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs)*. Genome Res, 2008. **18**(9): p. 1518-29.
286. Frommer, M., et al., *A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands*. Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1827-31.
287. Clark, S.J., et al., *High sensitivity mapping of methylated cytosines*. Nucleic Acids Res, 1994. **22**(15): p. 2990-7.
288. Ando, Y. and Y. Hayashizaki, *Restriction landmark genomic scanning*. Nat Protoc, 2006. **1**(6): p. 2774-83.
289. Zhang, D., et al., *Microarray-based molecular margin methylation pattern analysis in colorectal carcinoma*. Anal Biochem, 2006. **355**(1): p. 117-24.
290. Metzker, M.L., *Sequencing technologies - the next generation*. Nat Rev Genet, 2010. **11**(1): p. 31-46.
291. Gu, H., et al., *Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution*. Nat Methods, 2010. **7**(2): p. 133-6.
292. Kozlenkov, A., et al., *Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites*. Nucleic Acids Res, 2014. **42**(1): p. 109-27.
293. Kessler, N.J., et al., *CpG methylation differences between neurons and glia are highly conserved from mouse to human*. Hum Mol Genet, 2016. **25**(2): p. 223-32.
294. Peter, C.J. and S. Akbarian, *Balancing histone methylation activities in psychiatric disorders*. Trends Mol Med, 2011. **17**(7): p. 372-9.
295. Jakovcevski, M. and S. Akbarian, *Epigenetic mechanisms in neurological disease*. Nat Med, 2012. **18**(8): p. 1194-204.
296. Slotkin, R.K. and R. Martienssen, *Transposable elements and the epigenetic regulation of the genome*. Nat Rev Genet, 2007. **8**(4): p. 272-85.
297. Hancks, D.C. and H.H. Kazazian, Jr., *Roles for retrotransposon insertions in human disease*. Mob DNA, 2016. **7**: p. 9.
298. Coufal, N.G., et al., *L1 retrotransposition in human neural progenitor cells*. Nature, 2009. **460**(7259): p. 1127-31.
299. Muotri, A.R., et al., *Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition*. Nature, 2005. **435**(7044): p. 903-10.
300. Evrony, G.D., et al., *Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain*. Cell, 2012. **151**(3): p. 483-96.
301. Bundo, M., et al., *Increased L1 retrotransposition in the neuronal genome in schizophrenia*. Neuron, 2014. **81**(2): p. 306-13.
302. Verhoeven, K., et al., *Slowed conduction and thin myelination of peripheral nerves associated with mutant rho Guanine-nucleotide exchange factor 10*. Am J Hum Genet, 2003. **73**(4): p. 926-32.
303. Wang, W., E.J. Kwon, and L.H. Tsai, *MicroRNAs in learning, memory, and neurological diseases*. Learn Mem, 2012. **19**(9): p. 359-68.
304. Preuss, T.M., *The human brain: rewired and running hot*. Ann N Y Acad Sci, 2011. **1225 Suppl 1**: p. E182-91.
305. Brawand, D., et al., *The evolution of gene expression levels in mammalian organs*. Nature, 2011. **478**(7369): p. 343-8.

306. Konopka, G., et al., *Human-specific transcriptional networks in the brain*. Neuron, 2012. **75**(4): p. 601-17.
307. Geschwind, D.H. and P. Rakic, *Cortical evolution: judge the brain by its cover*. Neuron, 2013. **80**(3): p. 633-47.
308. Rakic, P., *Evolution of the neocortex: a perspective from developmental biology*. Nat Rev Neurosci, 2009. **10**(10): p. 724-35.
309. Striedter, G., *Principles of brain evolution*. Sinauer Associates Inc., 2005.
310. Borrell, V. and I. Reillo, *Emerging roles of neural stem cells in cerebral cortex development and evolution*. Dev Neurobiol, 2012. **72**(7): p. 955-71.
311. Lui, J.H., D.V. Hansen, and A.R. Kriegstein, *Development and evolution of the human neocortex*. Cell, 2011. **146**(1): p. 18-36.
312. Lewitus, E., I. Kelava, and W.B. Huttner, *Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development*. Front Hum Neurosci, 2013. **7**: p. 424.
313. Aiello, L. and P. Wheeler, *The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution*. Curr Anthropol, 1995. **43**: p. 199-221.
314. Gibbons, A., *Solving the brain's energy crisis*. Science, 1998. **280**: p. 1345-1347.
315. Deitmer, J.W., *Strategies for metabolic exchange between glial cells and neurons*. Respir Physiol, 2001. **129**(1-2): p. 71-81.
316. Ben Achour, S. and O. Pascual, *Glia: the many ways to modulate synaptic plasticity*. Neurochem Int, 2010. **57**(4): p. 440-5.
317. Perea, G. and A. Araque, *GLIA modulates synaptic transmission*. Brain Res Rev, 2010. **63**(1-2): p. 93-102.
318. Zechner, U., et al., *A high density of X-linked genes for general cognitive ability: a runaway process shaping human evolution?* Trends Genet, 2001. **17**(12): p. 697-701.
319. Billuart, P., et al., *Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation*. Nature, 1998. **392**(6679): p. 923-6.
320. Nadif Kasri, N., et al., *The Rho-linked mental retardation protein oligophrenin-1 controls synapse maturation and plasticity by stabilizing AMPA receptors*. Genes Dev, 2009. **23**(11): p. 1289-302.
321. Hanada, R., et al., *Central control of fever and female body temperature by RANKL/RANK*. Nature, 2009. **462**(7272): p. 505-9.
322. Jang, H.S., et al., *CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function*. Genes (Basel), 2017. **8**(6).
323. Guo, J.U., et al., *Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain*. Nat Neurosci, 2014. **17**(2): p. 215-22.
324. Lister, R., et al., *Global epigenomic reconfiguration during mammalian brain development*. Science, 2013. **341**(6146): p. 1237905.
325. Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1*. Science, 2009. **324**(5929): p. 930-5.
326. Chen, T., et al., *Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b*. Mol Cell Biol, 2003. **23**(16): p. 5594-605.
327. Kriaucionis, S. and N. Heintz, *The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain*. Science, 2009. **324**(5929): p. 929-30.
328. Khare, T., et al., *5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary*. Nat Struct Mol Biol, 2012. **19**(10): p. 1037-43.

329. Spiers, H., et al., *5-hydroxymethylcytosine is highly dynamic across human fetal brain development*. BMC Genomics, 2017. **18**(1): p. 738.
330. Tfayli, A., et al., *Breast cancer in low- and middle-income countries: an emerging and challenging epidemic*. J Oncol, 2010. **2010**: p. 490631.
331. Bassil, C.F., Z. Huang, and S.K. Murphy, *Bisulfite pyrosequencing*. Methods Mol Biol, 2013. **1049**: p. 95-107.
332. Huang, Z., C.F. Bassil, and S.K. Murphy, *Bisulfite sequencing of cloned alleles*. Methods Mol Biol, 2013. **1049**: p. 83-94.
333. Beygo, J., et al., *Deep bisulfite sequencing of aberrantly methylated loci in a patient with multiple methylation defects*. PLoS One, 2013. **8**(10): p. e76953.
334. Laurentino, S., et al., *Epigenetic germline mosaicism in infertile men*. Hum Mol Genet, 2015. **24**(5): p. 1295-304.
335. Beygo, J., et al., *The molecular function and clinical phenotype of partial deletions of the IGF2/H19 imprinting control region depends on the spatial arrangement of the remaining CTCF-binding sites*. Hum Mol Genet, 2013. **22**(3): p. 544-57.
336. Zemliakova, V.V., et al., *[Abnormal methylation of several tumor suppressor genes in sporadic breast cancer]*. Mol Biol (Mosk), 2003. **37**(4): p. 696-703.
337. Kang, J.H., et al., *Methylation in the p53 promoter is a supplementary route to breast carcinogenesis: correlation between CpG methylation in the p53 promoter and the mutation of the p53 gene in the progression from ductal carcinoma in situ to invasive ductal carcinoma*. Lab Invest, 2001. **81**(4): p. 573-9.
338. Garcia, J.M., et al., *Promoter methylation of the PTEN gene is a common molecular change in breast cancer*. Genes Chromosomes Cancer, 2004. **41**(2): p. 117-24.
339. Begam, N., K. Jamil, and S.G. Raju, *Promoter Hypermethylation of the ATM Gene as a Novel Biomarker for Breast Cancer*. Asian Pac J Cancer Prev, 2017. **18**(11): p. 3003-3009.
340. Crucianelli, F., et al., *MLH1 constitutional and somatic methylation in patients with MLH1 negative tumors fulfilling the revised Bethesda criteria*. Epigenetics, 2014. **9**(10): p. 1431-8.
341. Day, T.K. and T. Bianco-Miotto, *Common gene pathways and families altered by DNA methylation in breast and prostate cancers*. Endocr Relat Cancer, 2013. **20**(5): p. R215-32.
342. Nowsheen, S., et al., *Epigenetic inactivation of DNA repair in breast cancer*. Cancer Lett, 2014. **342**(2): p. 213-22.
343. Szyf, M., *DNA methylation signatures for breast cancer classification and prognosis*. Genome Med, 2012. **4**(3): p. 26.
344. Cooper, G., *The Cell: A Molecular Approach. 2nd edition*. Sinauer Associates, Inc., 2000.
345. Wang, Y., et al., *BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures*. Genes Dev, 2000. **14**(8): p. 927-39.
346. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection*. Nat Rev Cancer, 2011. **12**(1): p. 68-78.
347. Fradet-Turcotte, A., et al., *BRCA2 functions: from DNA repair to replication fork stabilization*. Endocr Relat Cancer, 2016. **23**(10): p. T1-T17.
348. Foulkes, W.D., *Inherited susceptibility to common cancers*. N Engl J Med, 2008. **359**(20): p. 2143-53.

349. Martinez-Galan, J., et al., *ESR1 gene promoter region methylation in free circulating DNA and its correlation with estrogen receptor protein expression in tumor tissue in breast cancer patients*. BMC Cancer, 2014. **14**: p. 59.
350. Suga, Y., et al., *Quantitative p16 and ESR1 methylation in the peripheral blood of patients with non-small cell lung cancer*. Oncol Rep, 2008. **20**(5): p. 1137-42.
351. Majidinia, M. and B. Yousefi, *DNA repair and damage pathways in breast cancer development and therapy*. DNA Repair (Amst), 2017. **54**: p. 22-29.
352. Leonhardt, H. and M.C. Cardoso, *DNA methylation, nuclear structure, gene expression and cancer*. J Cell Biochem Suppl, 2000. **Suppl 35**: p. 78-83.
353. Yuan, Y., et al., *Reactivation of SYK expression by inhibition of DNA methylation suppresses breast cancer cell invasiveness*. Int J Cancer, 2005. **113**(4): p. 654-9.
354. Reik, W., W. Dean, and J. Walter, *Epigenetic reprogramming in mammalian development*. Science, 2001. **293**(5532): p. 1089-93.
355. Xi, Y. and W. Li, *BSMAP: whole genome bisulfite sequence MAPPING program*. BMC Bioinformatics, 2009. **10**: p. 232.
356. Warnecke, P.M., et al., *Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA*. Nucleic Acids Res, 1997. **25**(21): p. 4422-6.
357. Hiatt, J.B., et al., *Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation*. Genome Res, 2013. **23**(5): p. 843-54.
358. Galetzka, D., et al., *Monozygotic twins discordant for constitutive BRCA1 promoter methylation, childhood cancer and secondary cancer*. Epigenetics, 2012. **7**(1): p. 47-54.
359. El Hajj, N., et al., *Limiting dilution bisulfite (pyro)sequencing reveals parent-specific methylation patterns in single early mouse embryos and bovine oocytes*. Epigenetics, 2011. **6**(10): p. 1176-88.
360. Nguyen, C., et al., *Susceptibility of nonpromoter CpG islands to de novo methylation in normal and neoplastic cells*. J Natl Cancer Inst, 2001. **93**(19): p. 1465-72.
361. Bock, C., et al., *Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping*. Nucleic Acids Res, 2008. **36**(10): p. e55.
362. Sontag, L.B., M.C. Lorincz, and E. Georg Luebeck, *Dynamics, stability and inheritance of somatic DNA methylation imprints*. J Theor Biol, 2006. **242**(4): p. 890-9.
363. Fackler, M.J., et al., *Quantitative multiplex methylation-specific PCR assay for the detection of promoter hypermethylation in multiple genes in breast cancer*. Cancer Res, 2004. **64**(13): p. 4442-52.
364. Sharma, G., et al., *Prognostic relevance of promoter hypermethylation of multiple genes in breast cancer patients*. Cell Oncol, 2009. **31**(6): p. 487-500.
365. Dong, S.M., et al., *Promoter hypermethylation of multiple genes in carcinoma of the uterine cervix*. Clin Cancer Res, 2001. **7**(7): p. 1982-6.
366. Jahr, S., et al., *DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells*. Cancer Res, 2001. **61**(4): p. 1659-65.
367. Chan, K.C., et al., *Size distributions of maternal and fetal DNA in maternal plasma*. Clin Chem, 2004. **50**(1): p. 88-92.
368. Mouliere, F., et al., *High fragmentation characterizes tumour-derived circulating DNA*. PLoS One, 2011. **6**(9): p. e23418.
369. Delgado, P.O., et al., *Characterization of cell-free circulating DNA in plasma in patients with prostate cancer*. Tumour Biol, 2013. **34**(2): p. 983-6.

370. Roninson, I.B., E.V. Broude, and B.D. Chang, *If not apoptosis, then what? Treatment-induced senescence and mitotic catastrophe in tumor cells*. Drug Resist Updat, 2001. **4**(5): p. 303-13.
371. Lo, Y.M., et al., *Rapid clearance of fetal DNA from maternal plasma*. Am J Hum Genet, 1999. **64**(1): p. 218-24.
372. Yu, S.C., et al., *High-resolution profiling of fetal DNA clearance from maternal plasma by massively parallel sequencing*. Clin Chem, 2013. **59**(8): p. 1228-37.
373. Qin, Z., et al., *Cell-free circulating tumor DNA in cancer*. Chin J Cancer, 2016. **35**: p. 36.
374. Pantel, K., R.H. Brakenhoff, and B. Brandt, *Detection, clinical relevance and specific biological properties of disseminating tumour cells*. Nat Rev Cancer, 2008. **8**(5): p. 329-40.
375. Alix-Panabieres, C. and K. Pantel, *Circulating tumor cells: liquid biopsy of cancer*. Clin Chem, 2013. **59**(1): p. 110-8.
376. Meng, S., et al., *Circulating tumor cells in patients with breast cancer dormancy*. Clin Cancer Res, 2004. **10**(24): p. 8152-62.
377. Li, Q., et al., *Promoter methylation and expression changes of BRCA1 in cancerous tissues of patients with sporadic breast cancer*. Oncol Lett, 2015. **9**(4): p. 1807-1813.
378. Baylin, S.B., et al., *Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer*. Hum Mol Genet, 2001. **10**(7): p. 687-92.
379. Li, Z., et al., *Integrated analysis of gene expression and methylation profiles of 48 candidate genes in breast cancer patients*. Breast Cancer Res Treat, 2016. **160**(2): p. 371-383.
380. Lee, C.J., et al., *Determining the effect of DNA methylation on gene expression in cancer cells*. Methods Mol Biol, 2014. **1101**: p. 161-78.
381. Joehanes, R., et al., *Gene expression signatures of coronary heart disease*. Arterioscler Thromb Vasc Biol, 2013. **33**(6): p. 1418-26.
382. Chodavarapu, R.K., et al., *Relationship between nucleosome positioning and DNA methylation*. Nature, 2010. **466**(7304): p. 388-92.
383. Bird, A.P. and A.P. Wolffe, *Methylation-induced repression--belts, braces, and chromatin*. Cell, 1999. **99**(5): p. 451-4.
384. Ben-Porath, I. and H. Cedar, *Epigenetic crosstalk*. Mol Cell, 2001. **8**(5): p. 933-5.
385. Richards, E.J. and S.C. Elgin, *Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects*. Cell, 2002. **108**(4): p. 489-500.
386. Eden, S., et al., *DNA methylation models histone acetylation*. Nature, 1998. **394**(6696): p. 842.
387. Schubeler, D., et al., *Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation*. Mol Cell Biol, 2000. **20**(24): p. 9103-12.
388. Fuks, F., et al., *The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation*. J Biol Chem, 2003. **278**(6): p. 4035-40.
389. Tamaru, H. and E.U. Selker, *A histone H3 methyltransferase controls DNA methylation in Neurospora crassa*. Nature, 2001. **414**(6861): p. 277-83.
390. Jackson, J.P., et al., *Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase*. Nature, 2002. **416**(6880): p. 556-60.

2. Internetquellen

- Int. 1 National Institute of Neurological Disorders and Stroke: Brain Basics: The Life and Death of a Neuron (2016). [Online im Internet:] URL: <https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Life-and-Death-Neuron> [aufgerufen am 06.02.2018]
- Int. 2 World Health Organization: Breast cancer (2018). [Online im Internet:] URL: <http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/> [aufgerufen am 16.11.2017]
- Int. 3 American Cancer Society: Breast Cancer Facts & Figures 2015-2016 (2015). [Online im Internet:] URL: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2015-2016.pdf> [aufgerufen am 16.11.2017]
- Int. 4 Deutsche Krebsgesellschaft: Brustkrebs Diagnose: Klassifikation (2018). [Online im Internet:] URL: <https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/brustkrebs/klassifikation.html> [aufgerufen am 06.03.2018]
- Int. 5 Plastische und Ästhetische Chirurgie, Universitätsklinikum Bonn: Aufbau der Brust. [Online im Internet:] URL: <https://www.meinebrust.de/brustvergroesserung-aufbau-der-brust/> [aufgerufen am 06.03.2018]
- Int. 6 Deutsche Krebsgesellschaft: Erblicher Brustkrebs – wenn der Krebs in den Genen liegt (2017). [Online im Internet:] URL: <https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/basis-informationen-krebs-allgemeine-informationen/erblicher-brustkrebs-wenn-der-k.html> [aufgerufen am 16.11.2017]
- Int. 7 EpiBeat: The Basics of Bisulfite Conversion for DNA Methylation Analysis (2013). [Online im Internet:] URL: <http://www.epibeat.com/tools-technology/the-basics-of-bisulfite-conversion-for-dna-methylation-analysis/721/> [aufgerufen am 09.10.2017]
- Int. 8 CoreGenomics: How do SPRI beads work (2012). [Online im Internet:] URL: <http://coregenomics.blogspot.de/2012/04/how-do-spri-beads-work.html> [aufgerufen am 05.03.2018]
- Int. 9 Broad Institute: Broad/Illumina Genome Analyzer Boot Camp (2018). [Online im Internet:] URL: <https://www.broadinstitute.org/files/shared/illumina/sequencingSlides.pdf> [aufgerufen am 05.03.2018]
- Int. 10 Illumina: MiSeq System: Denature and Dilute Libraries Guide (2017). [Online im Internet:] URL: https://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/miseq-denature-dilute-libraries-guide-15039740-03.pdf [aufgerufen am 08.03.2018]
- Int. 11 Illumina: Reagent Preparation Guide For Single-Read Runs (2012) [Online nicht mehr verfügbar] URL: http://support.illumina.com/sequencing/sequencing_kits/truseq_sr_cluster_kit_v5-cs-ga/documentation.ilmn [aufgerufen am 06.09.2014]

- Int. 12 Illumina: Reagent Preparation Guide (2012) [Online nicht mehr verfügbar] URL: http://support.illumina.com/sequencing/sequencing_kits/truseq_sbs_kit_v5-ga/documentation.ilmn [aufgerufen am 06.09.2014]
- Int. 13 Babraham Bioinformatics: FastQC (2014) [Online im Internet:] URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [aufgerufen am 05.03.2018]
- Int. 14 Babraham Bioinformatics: Trim Galore! (2015) [Online im Internet:] URL: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [aufgerufen am 05.03.2018]
- Int. 15 Cutadapt (2014) [Online im Internet:] URL: <https://cutadapt.readthedocs.io/en/stable/> [aufgerufen am 05.03.2018]
- Int. 16 Ea-utils: Command-line tools for processing biological sequencing data (2011) [Online im Internet:] URL: <https://expressionanalysis.github.io/ea-utils/> [aufgerufen am 05.03.2018]
- Int. 17 Babraham Bioinformatics: Bismark (2015) [Online im Internet:] URL: <https://www.bioinformatics.babraham.ac.uk/projects/bismark/> [aufgerufen am 05.03.2018]
- Int. 18 Bowtie 2: Fast and sensitive read alignment (2012) [Online im Internet:] URL: bowtie-bio.sourceforge.net/bowtie2/ [aufgerufen am 05.03.2018]
- Int. 19 Samtools (2017) [Online im Internet:] URL: <http://htslib.org/> [aufgerufen am 05.03.2018]
- Int. 20 UCSC Genome Browser: Annotation Database hg19 (2013) [Online im Internet:] URL: <http://hgdownload.soe.ucsc.edu/goldenPath/hg19> [aufgerufen am 05.03.2018]
- Int. 21 UCSC Genome Browser: Annotation Database panTro4 (2013) [Online im Internet:] URL: <http://hgdownload.soe.ucsc.edu/goldenPath/panTro4/> [aufgerufen am 05.03.2018]
- Int. 22 UCSC Genome Browser: LiftOver (2013) [Online im Internet:] URL: http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/liftOver [aufgerufen am 05.03.2018]
- Int. 23 UCSC Genome Browser: LiftOver (2013) [Online im Internet:] URL: <http://hgdownload.cse.ucsc.edu/goldenPath/panTro4/liftOver/panTro4ToHg19.over.chain.gz> [aufgerufen am 05.03.2018]
- Int. 24 Vegan: Community Ecology Package (2018) [Online im Internet:] URL: <https://cran.r-project.org/web/packages/vegan/index.html> [aufgerufen am 05.03.2018]
- Int. 25 AutDB: Autism Informatics Portal (2017) [Online im Internet:] URL: <http://autism.mindspec.org/autdb> [aufgerufen am 11.05.2017]
- Int. 26 BDgene: Genetic Database for Bipolar Disorder (2016) [Online im Internet:] URL: <http://bdgene.psych.ac.cn> [aufgerufen am 11.05.2017]
- Int. 27 Alzforum: AlzGene (2013) [Online im Internet:] URL: <http://www.alzgene.org> [aufgerufen am 25.04.2017]

- Int. 28 Qiagen: Pyrosequencing (2018) [Online im Internet:] URL: <https://www.qiagen.com/sg/resources/technologies/pyrosequencing-resource-center/technology-overview/#comparison> [aufgerufen am 10.10.2017]
- Int. 29 ThermoFisher: How TaqMan Assays Work (2018) [Online im Internet:] URL: <https://www.thermofisher.com/de/de/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/how-taqman-assays-work.html> [aufgerufen am 05.03.2017]
- Int. 30 Cancer Research UK: Worldwide cancer incidence statistics (2014) [Online im Internet:] URL: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer/incidence#heading-One> [aufgerufen am 16.11.2017]
- Int. 31 Deutsche Krebsgesellschaft: Brustkrebs: Diagnoseverfahren im Überblick (2017) [Online im Internet:] URL: <https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/brustkrebs/diagnose.html> [aufgerufen am 16.11.2017]
- Int. 32 GTEX: Gene Page (2017) [Online im Internet:] URL: <https://www.gtportal.org/home/gene/ACTB> [aufgerufen am 22.01.2018]
- Int. 33 GTEX: Gene Page (2017) [Online im Internet:] URL: <https://www.gtportal.org/home/gene/GAPDH> [aufgerufen am 22.01.2018]

VERZEICHNISSE

1. Abkürzungsverzeichnis

	°C	Grad Celsius
	µl	Mikroliter
	µM	Mikromolar
	∅	Durchschnitt
A	A	Adenin
	APS	Adenosinphosphosulfat
	ATP	Adenosintriphosphat
B	BC	Brustkrebs (engl. <i>breast cancer</i>)
	bp	Basenpaar
	bzw.	beziehungsweise
C	C	Cytosin
	ca.	circa
	CpG	Cytosin-Phospho-Guanin-Dinukleotid
	CTOB	engl. <i>complementary to the original bottom strand</i>
	CTOT	engl. <i>complementary to the original top strand</i>
D	dATP	Desoxyadenosintriphosphat
	DBAS	engl. <i>Deep Bisulfite Amplicon Sequencing</i>
	dCTP	Desoxycytidintriphosphat
	dGTP	Desoxyguanosintriphosphat
	DMR	differenziell methylierte Region
	DNA	Desoxyribonukleinsäure (engl. <i>desoxyribonucleic acid</i>)
	dNTP	Desoxyribonukleosidtriphosphat
	ds	doppelsträngig
E	EB	engl. <i>Elution Buffer</i>
	EO	engl. <i>early-onset</i>
	EMR	Epimutations-Rate
	EtOH	Ethanol
F	Fwd	engl. <i>forward</i>
G	g	Erdbeschleunigung

	G	Guanin
	g/mol	Gramm pro mol
	GAI	Genome Analyzer II
	gbisDNA	genomische bisulfit-konvertierte DNA
	gDNA	genomische DNA
H	h	Stunde (lat. <i>hora</i>)
	H ₂ O	Wasser
	HR	engl. <i>high-risk</i>
	HSA	lat. <i>Homo sapiens</i>
I	IQB	Interquartilbereich
J	Jhd.	Jahrhundert
K	K	Kontrolle
L	LED	Leuchtdiode (engl. <i>light-emitting diode</i>)
M	M	Molare Masse
	MAF	engl. <i>minor allele frequency</i>
	MgCl ₂	Magnesiumchlorid
	min	Minuten
	ml	Milliliter
	mM	Millimolar
	MYA	Millionen Jahre (engl. <i>million years ago</i>)
N	N+	neuronale Zellen
	N-	nicht-neuronale Zellen
	N	Normal
	NaAc	Natriumacetat
	NaCl	Natriumchlorid
	NaOH	Natriumhydrogensulfat
	ng	Nanogramm
	NGS	engl. <i>Next Generation Sequencing</i>
	nM	Nanomolar
O	OB	engl. <i>original bottom strand</i>
	OT	engl. <i>original top strand</i>

P	PCR	Polymerase-Kettenreaktion (engl. <i>polymerase chain reaction</i>)
	pM	Pikomolar
	PP _i	Pyrophosphat
	PTR	lat. <i>Pan troglodytes</i>
R	Rev	engl. <i>reverse</i>
	RNA	Ribonukleinsäure (engl. <i>ribonucleic acid</i>)
	rpm	Umdrehungen pro Minute (engl. <i>revolutions per minute</i>)
	RRBS	engl. <i>Reduced Representation Bisulfite Sequencing</i>
	RT	Raumtemperatur
S	s	Sekunde
	SBS	engl. <i>Sequencing by Synthesis</i>
	SD	Standardabweichung (engl. <i>standard deviation</i>)
	SNP	Einzelnukleotid-Polymorphismus (engl. <i>single nucleotide polymorphism</i>)
	SPRI	engl. <i>Solid Phase Reversible Immobilisation</i>
	ss	einzelsträngig (engl. <i>single stranded</i>)
T	T	Thymin
	TSG	Tumorsuppressorgen
U	U	Uracil
	u.a.	unter anderem
	U/μl	Units per Mikroliter
	ÜN	über Nacht
	USER	engl. <i>Uracil-Specific Excision Reagent</i>
	UV	Ultraviolett
V	V	Volt

2. Abbildungsverzeichnis

Abbildung 1: Taxonomische Klassifizierung der Primaten und phylogenetische Beziehung rezenter großer Primaten.	7
Abbildung 2: Stammbaum der <i>Hominini</i>	8
Abbildung 3: Lage des Brodmann-Areals 10 im präfrontalen Cortex.	7
Abbildung 4: Schematische Darstellung der typischen Neuronen bei Säugern.	9
Abbildung 5: Schematische Darstellung verschiedener Gliazelltypen.	10
Abbildung 6: Schematische Darstellung der drei wesentlichen epigenetischen Mechanismen.	13
Abbildung 7: Vereinfachte schematische Darstellung der DNA-Methylierung und Demethylierung.	15
Abbildung 8: Verteilung der DNA-Methylierung im menschlichen Genom.	16
Abbildung 9: Schematische Darstellung der DNA-Methylierungsmuster in einer normalen Zelle und einer Krebszelle.	21
Abbildung 10: Anzahl an aufgetretenen <i>in situ</i> und invasiven Brustkrebserkrankungen in den USA.	24
Abbildung 11: Aufbau der weiblichen Brust, schematisch dargestellt.	25
Abbildung 12: Verteilung der aktuell bekannten Brustkrebs susceptibilitäts gene bzw. -loci für familiären Brustkrebs.	29
Abbildung 13: Die „Two-Hit“ Hypothese von Knudson.	30
Abbildung 14: Methylierungsveränderungen bei Brustkrebs.	31
Abbildung 15: Chemische Reaktion der Bisulfit-vermittelten Desaminierung.	50
Abbildung 16: Schematische Darstellung des Ablaufs der Probenpräparation beim DBAS.	52
Abbildung 17: Beispielhafte Sequenzen nach einer Amplifikation der gBisDNA.	53
Abbildung 18: Aufreinigung mittels SPRI (Solid Phase Reversible Immobilisation) beads.	55
Abbildung 19: Schematische Darstellung der Platten- bzw. Reagenzien-Anordnung auf dem JANUS Pipettierroboter.	57
Abbildung 20: Elektropherogramm und Gelbild eines finalen Pools beim DBAS.	64
Abbildung 21: Schematische Übersicht über einen NGS-Zyklus bei der SBS-Technologie.	65
Abbildung 22: Schematische Darstellung des Arbeitsablaufs zur Generierung der RRBS-Daten.	67
Abbildung 23: Elektropherogramm und Gelbild einer finalen RRBS Library.	73
Abbildung 24: Übersicht über die enzymatische Reaktionskaskade bei der Pyrosequenzierung.	79
Abbildung 25: Prinzip der Real-Time TaqMan Technologie.	81
Abbildung 26: Anzahl an sequenzierten Reads und abgedeckten CpG-Stellen je Library.	84
Abbildung 27: Korrespondenzanalyse basierend auf den 10.000 variabelsten Methylierungsloci.	85
Abbildung 28: Verteilung der abgedeckten CpG-Stellen auf Genom- und Genebene.	86
Abbildung 29: Anzahl an Intra- und interspezifischen DMRs.	87
Abbildung 30: Chromosomale Verteilung der intra- und interspezifischen DMRs.	90
Abbildung 31: Chromosomale Verteilung der mittels RRBS abgedeckten CpG-Stellen.	90
Abbildung 32: Signifikant angereicherte „Terms“ der Ontologie „biological process“.	92
Abbildung 33: Assoziierung verschiedener Repeat-Klassen mit intra- und interspezifischen DMRs.	105
Abbildung 34: Validierung der RRBS-Daten mittels Pyrosequenzierung.	107
Abbildung 35: Durchschnittliche Coverage je Amplikon inklusive Standardabweichung (SD).	110
Abbildung 36: Durchschnittliche Methylierungslevel je Amplikon.	112
Abbildung 37: Verteilung der Epimutations-Raten (EMR).	114
Abbildung 38: Methylierungsprofil auf Einzel-Allel-Ebene von Proben mit und ohne Epimutation.	117
Abbildung 39: Methylierungsprofil, exemplarisch dargestellt für eine HR BC Probe mit „single CpG errors“.	120
Abbildung 40: Durchschnittliche Methylierung je CpG-Stelle für <i>BRCA1</i> , <i>RAD51C</i> und <i>TP53</i> Region 2.	123
Abbildung 41: Durchschnittliche Methylierung je CpG-Stelle für <i>BRCA2</i> , <i>ATM</i> Region 1 und 2, <i>PTEN</i> , <i>TP53</i> Region 1, <i>MLH1</i> und <i>RB1</i>	128
Abbildung 42: Kummulative Hypermethylierung in Proben mit Epimutationen bzw. Einzel-CpG-Fehlern.	130

Abbildung 43: Beziehung zwischen der Anzahl an Genen mit single CpG errors einer Probe und ihrem Alter bei der Krebsdiagnose. 131

Abbildung 44: Ergebnisse der Expressionsanalyse in den Tumorsuppressorgenen 132

3. Tabellenverzeichnis

Tabelle 1: Körperliche Merkmale und Besonderheiten der Hominiden. 10

Tabelle 2: Übersicht über die analysierten Brustkrebspatienten- und Kontrollproben..... 35

Tabelle 3: Auflistung der Brustkrebspatienten- und Kontrollproben 37

Tabelle 4: Übersicht über die Individuen für die vergleichende Methylierungsanalyse zwischen Mensch und Schimpanse. 37

Tabelle 5: Genspezifische PCR-Primer für das DBAS. 38

Tabelle 6: Indexprimer für das DBAS 39

Tabelle 7: Adaptersequenz für DBAS 40

Tabelle 8: Indexprimer für RRBS 40

Tabelle 9: Adaptersequenz für RRBS..... 41

Tabelle 10: Sonden für die TaqMan RT-PCR..... 41

Tabelle 11: PCR-Primer und die korrespondierenden Sequenzierprimer für die Pyrosequenzierung..... 42

Tabelle 12: Kommerziell erworbene Kit-Systeme mit Herstellerangaben. 43

Tabelle 13: Verwendete Reagenzien mit Angaben zum Hersteller..... 43

Tabelle 14: Verwendete Enzyme mit Angaben zum Hersteller..... 44

Tabelle 15: Chemikalien mit Angaben zum Hersteller. 44

Tabelle 16: Verwendete Puffer und Lösungen mit Angaben zur Zusammensetzung. 44

Tabelle 17: Verwendete Geräte mit Angaben zum Hersteller. 45

Tabelle 18: Auflistung der wichtigsten Software- und Analysetools..... 46

Tabelle 19: Mit der Neurogenese assoziierte Gene, die mit interspezifischen DMRs überlappen..... 93

Tabelle 20: Enrichmentanalyse der DMRs in Genlisten, die mit neuropsychiatrischen Erkrankungen assoziiert sind. 95

Tabelle 21: Inter- und intraspezifische DMRs assoziiert mit Mentaler Retardierung..... 96

Tabelle 22: DMR-Abgleich mit human-spezifisch regulierten Genen. 98

Tabelle 23: Interspezifische DMRs assoziiert mit human-spezifischer Regulation im Gehirn. 99

Tabelle 24: Ausgewählte LINE1-assoziierte DMRs. 102

Tabelle 25: Durchschnittliche Methylierungslevel 110

Tabelle 26: Anzahl an Epimutationen (>50% methylierte CpGs) 113

Tabelle 27: Durchschnittliche Epimutations-Raten (EMR) 115

Tabelle 28: Anzahl an Einzel-CpG-Hypermethylierungen 119

Anhang-Tabelle 1: Gene Ontology Term Analyse der humanen intraspezifischen DMRs. 193

Anhang-Tabelle 2: Gene Ontology Term Analyse der intraspezifischen DMRs im Schimpansen. 196

Anhang-Tabelle 3: Auflistung der 138 Gene assoziiert mit Mentaler Retardierung 197

Anhang-Tabelle 4: Inter- und intraspezifische DMRs assoziiert mit neuropsychiatrischen Erkrankungen 197

Anhang-Tabelle 5: Intraspezifische DMRs, assoziiert mit human-spezifischer Regulation im Gehirn. 202

Anhang-Tabelle 6: Übersicht aller DMR-assoziierten L1 Elemente..... 206

PUBLIKATIONEN & KONGRESSBEITRÄGE

1. Publikationen

Allele-specific methylation analysis of imprinted genes reveals parental age and BMI effects in cord bloods.

Potabatulla R, Dittrich M, Böck J, Schorsch M, Hahn T, Haertle L, Müller T, El Hajj N, Haaf T.
(Submitted to Epigenomics)

Cell type and species-specific methylation patterns in neuronal and non-neuronal cells of human and chimpanzee cortex.

Böck J*, Remmele CW*, Dittrich M, Müller T, Kondova I, Persengiev S, Bontrop RE, Ade CP, Kraus T, Giese A, El Hajj N, Schneider E, Haaf T.

* Contributed equally (Submitted to Cerebral Cortex)

Single CpG hypermethylation, allele methylation errors, and decreased expression of multiple tumor suppressor genes in normal body cells of mutation-negative early-onset and high-risk breast cancer patients.

Böck J, Appenzeller S, Haertle L, Schneider T, Gehrig A, Schröder J, Rost S, Wolf B, Bartram CR, Sutter C, Haaf T.

Int J Cancer. 2018 Apr 16. doi: 10.1002/ijc.31526. [Epub ahead of print]

Hypermethylation of the non-imprinted maternal MEG3 and paternal MEST alleles is highly variable among normal individuals.

Haertle L, Maierhofer A, Böck J, Lehnen H, Böttcher Y, Blüher M, Schorsch M, Potabattula R, El Hajj N, Appenzeller S, Haaf T.

PLoS One, 2017 Aug 30; 12(8):e0184030. doi: 10.1371/journal.pone.0184030. eCollection 2017.

CpG sites with continuously increasing or decreasing methylation from early to late human fetal brain development.

Schneider E, Dittrich M, Böck J, Nanda I, Müller T, Seidmann L, Tralau T, Galetzka D, El Hajj N, Haaf T.

Gene, 2016 Oct 30; 592(1):110-8. doi: 10.1016/j.gene.2016.07.058. Epub 2016 Jul 25.

Epigenetic dysregulation in the developing Down syndrome cortex.

El Hajj N, Dittrich M, Böck J, Kraus TF, Nanda I, Müller T, Seidmann L, Tralau T, Galetzka D, Schneider E, Haaf T.

Epigenetics, 2016 Aug 2; 11(8):563-78. doi: 10.1080/15592294.2016.1192736. Epub 2016 May 31.

Sterile alpha motif containing 7 (samd7) is a novel crx-regulated transcriptional repressor in the retina.

Hlawatsch J, Karlstetter M, Aslanidis A, Lückoff A, Walczak Y, Plank M, Böck J, Langmann T.

PLoS One. 2013; 8(4):e60633. doi: 10.1371/journal.pone.0060633. Epub 2013 Apr 2.

2. Kongressbeiträge

2.1 Erstauthor

Single CpG hypermethylation, allele methylation errors, and decreased expression of multiple tumor suppressor genes in normal body cells of mutation-negative early-onset and high-risk breast cancer patients. **Böck J**, Appenzeller S, Haertle L, Schneider T, Gehrig A, Schröder J, Rost S, Wolf B, Bartram CR, Sutter C, Haaf T. Poster presentation (nominated for poster prize): 29th Conference of the German Society of Human Genetics, Münster, Germany (2018)

P-BasEpi-004: https://www.gfhev.de/de/kongress/Tagungsband_Abstracts_medgen_1_2018.pdf

DNA methylation patterns of cortical neurons and glia distinguish humans from chimpanzees – implications for human brain evolution. **Böck J**, Remmele C, Dittrich M, Schneider E, El Hajj N, Kondova I, Kraus T, Haaf T. Oral presentation: 27th Conference of the German Society of Human Genetics, Lübeck, Germany (2016)

WS9-01: http://beat.wolf.home.hefr.ch/documents/GfH_2016_Tagungsband_final.pdf#page=59

Deep bisulfite sequencing for quantification of constitutive epimutations in tumor suppressor genes. **Böck J**, Haertle L, Appenzeller S, Wolf B, Schneider T, Sutter C, Haaf T. Poster presentation (nominated): 27th Conference of the German Society of Human Genetics, Lübeck, Germany (2016)

P-Nomin.-002:

http://beat.wolf.home.hefr.ch/documents/GfH_2016_Tagungsband_final.pdf#page=59

2.2 Ko-Author

Paternal body mass index (BMI) effects on sperm methylation and transmission into the next generation. Potabattula R, Schorsch M, Dittrich M, **Böck J**, Haertle L, Hahn T, El Hajj N, Haaf T. Oral presentation: 28th Conference of the German Society of Human Genetics, Bochum, Germany (2017)

WS9-001: http://www.gfhev.de/de/kongress/009_Abstractband_gesamt.pdf

The unmethylated allele of oppositely imprinted (i.e. *MEST* and *MEG3*) genes is highly susceptible to epimutations during early development and may contribute an additional layer of complexity to phenotypic variation. Haertle L, Maierhofer A, **Böck J**, Lehnen H, Böttcher Y, Blüher M, Schorsch M, Pottabattula R, Kuhtz J, Wolf B, Appenzeller S, El Hajj N, Haaf T. Oral presentation: 27th Conference of the German Society of Human Genetics, Lübeck, Germany (2016)

WS9-03: http://beat.wolf.home.hefr.ch/documents/GfH_2016_Tagungsband_final.pdf#page=59

Epigenetic dysregulation in the developing Down syndrome brain. El Hajj N, Dittrich M, **Böck J**, Kraus TF, Nanda I, Müller T, Tralau T, Galetzka D, Schneider E, Haaf T. Oral presentation: 27th Conference of the German Society of Human Genetics, Lübeck, Germany (2016)

WS9-05: http://beat.wolf.home.hefr.ch/documents/GfH_2016_Tagungsband_final.pdf#page=59

Optimization and evaluation of differential methylation analyses for RRBS data. Remmele C, **Böck J**, Kondova I, Kraus T, Müller T, Haaf T, Schneider E, Dittrich M. Poster presentation: 27th Conference of the German Society of Human Genetics, Lübeck, Germany (2016)

P-Techno-257:

http://beat.wolf.home.hefr.ch/documents/GfH_2016_Tagungsband_final.pdf#page=59

EIDESSTATTLICHE VERSICHERUNG

Die vorliegende Arbeit wurde von Januar 2014 bis Mai 2018 am Institut für Humangenetik der Universität Würzburg unter der Betreuung von Herrn Univ.-Prof. Dr. med. Thomas Haaf angefertigt.

Hiermit erkläre ich an Eides statt, die Dissertation „Epigenetische Modifikationen und deren Einfluss auf die phänotypische Ausprägung: Differenzielle Methylierungsanalysen mittels verschiedener Next-Generation Sequencing-basierter Techniken“ selbständig und ohne Hilfe einer kommerziellen Promotionsberatung sowie ausschließlich unter Verwendung der angegebenen Quellen und Hilfsmittel verfasst zu haben. Die, den herangezogenen Werken wörtlich oder sinngemäß entnommenen Stellen sind als solche gekennzeichnet.

Ich erkläre hiermit, dass die Regeln der Universität Würzburg über gute wissenschaftliche Praxis eingehalten wurden. Die Dissertation wurde weder in gleicher noch in ähnlicher Form in einem anderen Prüfungsverfahren vorgelegt.

Am 10.08.2011 wurde mir von der Universität Regensburg der akademische Grad „Bachelor of Science“ und am 21.11.2013 von der Julius-Maximilians-Universität Würzburg der akademische Grad „Master of Science“ verliehen. Weitere akademische Grade habe ich nicht erworben.

.....
Ort, Datum

.....
Unterschrift (Böck Julia)

ANHANG

Anhang-Tabelle 1: Gene Ontology Term Analyse der humanen intraspezifischen DMRs. Dargestellt sind die signifikant (FDR adjusted p-value < 0,05) angereicherten „Terms“ für den intraspezifischen Vergleich (N+ vs. N-) im Menschen. **A)** Ontologie „Biological process“. **B)** Ontologie „Cellular component“. **C)** „Molecular function“.

A) „Biological process“

Term	Enrichment	p-value*
GO:0099536~synaptic signaling	2,34	2,77E-08
GO:0007268~chemical synaptic transmission	2,34	2,77E-08
GO:0098916~anterograde trans-synaptic signaling	2,34	2,77E-08
GO:0099537~trans-synaptic signaling	2,34	2,77E-08
GO:0007399~nervous system development	1,52	5,61E-06
GO:0022008~neurogenesis	1,67	6,37E-06
GO:0023052~signaling	1,27	1,02E-05
GO:0043087~regulation of GTPase activity	2,06	1,23E-05
GO:0048666~neuron development	1,83	1,23E-05
GO:0031344~regulation of cell projection organization	2,11	2,15E-05
GO:0031175~neuron projection development	1,89	3,32E-05
GO:0043547~positive regulation of GTPase activity	2,07	4,89E-05
GO:0044700~single organism signaling	1,26	4,94E-05
GO:0030030~cell projection organization	1,66	5,54E-05
GO:0050803~regulation of synapse structure or activity	2,81	8,77E-05
GO:0010975~regulation of neuron projection development	2,28	8,96E-05
GO:0030182~neuron differentiation	1,66	0,0001
GO:0007154~cell communication	1,25	0,0002
GO:0050804~modulation of synaptic transmission	2,61	0,0002
GO:0048812~neuron projection morphogenesis	2,05	0,0003
GO:0048699~generation of neurons	1,61	0,0003
GO:0007267~cell-cell signaling	1,59	0,0004
GO:0045664~regulation of neuron differentiation	1,99	0,0006
GO:0048468~cell development	1,47	0,0008
GO:0051960~regulation of nervous system development	1,79	0,0018
GO:0051179~localization	1,22	0,0020
GO:0030154~cell differentiation	1,31	0,0023
GO:0051049~regulation of transport	1,49	0,0026
GO:0043269~regulation of ion transport	1,98	0,0033
GO:0034765~regulation of ion transmembrane transport	2,19	0,0035
GO:0098655~cation transmembrane transport	1,86	0,0037
GO:0048167~regulation of synaptic plasticity	3,12	0,0037
GO:0050767~regulation of neurogenesis	1,81	0,0040
GO:0050773~regulation of dendrite development	3,19	0,0079
GO:0070588~calcium ion transmembrane transport	2,49	0,0083
GO:0034220~ion transmembrane transport	1,68	0,0091

GO:0034762~regulation of transmembrane transport	2,11	0,0094
GO:0048869~cellular developmental process	1,27	0,0150
GO:0048667~cell morphogenesis involved in neuron differentiation	1,92	0,0177
GO:0006810~transport	1,24	0,0199
GO:0050806~positive regulation of synaptic transmission	3,33	0,0224
GO:0065007~biological regulation	1,10	0,0301
GO:0065009~regulation of molecular function	1,32	0,0313
GO:0006811~ion transport	1,49	0,0359
GO:0051336~regulation of hydrolase activity	1,53	0,0362
GO:0030001~metal ion transport	1,70	0,0375
GO:0044763~single-organism cellular process	1,08	0,0406
GO:0051345~positive regulation of hydrolase activity	1,64	0,0415
GO:0051234~establishment of localization	1,22	0,0426
GO:0055085~transmembrane transport	1,52	0,0460
GO:0048858~cell projection morphogenesis	1,66	0,0469
GO:0098660~inorganic ion transmembrane transport	1,73	0,0512

* FDR adjusted p-value

B) „Cellular component“

Term	Enrichment	p-value*
GO:0045202~synapse	2,45	1,67E-13
GO:0071944~cell periphery	1,47	3,33E-13
GO:0005886~plasma membrane	1,48	6,77E-13
GO:0097458~neuron part	2,02	1,51E-12
GO:0044459~plasma membrane part	1,64	3,48E-10
GO:0043005~neuron projection	2,04	8,16E-09
GO:0044456~synapse part	2,32	1,92E-08
GO:0016020~membrane"	1,21	5,94E-08
GO:0036477~somatodendritic compartment	2,27	5,94E-08
GO:0098590~plasma membrane region	1,97	6,78E-07
GO:0034702~ion channel complex	2,88	3,69E-06
GO:1902495~transmembrane transporter complex	2,67	1,49E-05
GO:0030425~dendrite	2,32	1,58E-05
GO:0042995~cell projection	1,58	1,61E-05
GO:1990351~transporter complex	2,61	2,82E-05
GO:0043025~neuronal cell body	2,39	3,25E-05
GO:0030054~cell junction	1,64	7,56E-05
GO:0044297~cell body	2,20	0,0002
GO:0098794~postsynapse	2,35	0,0002
GO:0097060~synaptic membrane	2,56	0,0005
GO:0031226~intrinsic component of plasma membrane	1,57	0,0008
GO:0044425~membrane part	1,22	0,0014
GO:0060076~excitatory synapse	2,77	0,0026
GO:0030424~axon	2,12	0,0027

GO:0005887~integral component of plasma membrane	1,55	0,0030
GO:0099572~postsynaptic specialization	2,75	0,0073
GO:0014069~postsynaptic density	2,75	0,0073
GO:0098797~plasma membrane protein complex	1,97	0,0109
GO:0045211~postsynaptic membrane	2,57	0,0165
GO:0098793~presynapse	2,25	0,0205
GO:0034703~cation channel complex	2,74	0,0292

* FDR adjusted p-value

C) „Molecular function“

Term	Enrichment	p-value*
GO:0022836~gated channel activity	2,52	0,0001
GO:0005261~cation channel activity	2,61	0,0002
GO:0005216~ion channel activity	2,28	0,0003
GO:0022838~substrate-specific channel activity	2,22	0,0006
GO:0015085~calcium ion transmembrane transporter activity	3,50	0,0007
GO:0015267~channel activity	2,13	0,0015
GO:0022803~passive transmembrane transporter activity	2,13	0,0016
GO:0046873~metal ion transmembrane transporter activity	2,19	0,0017
GO:0072509~divalent inorganic cation transmembrane transporter activity	2,91	0,0041
GO:0005262~calcium channel activity	3,49	0,0052
GO:0005516~calmodulin binding	2,78	0,0094
GO:0008324~cation transmembrane transporter activity	1,89	0,0102
GO:0015075~ion transmembrane transporter activity	1,73	0,0175
GO:0022843~voltage-gated cation channel activity	3,09	0,0191
GO:0005088~Ras guanyl-nucleotide exchange factor activity	2,43	0,0245
GO:0005085~guanyl-nucleotide exchange factor activity	2,14	0,0462
GO:0022857~transmembrane transporter activity	1,60	0,0475
GO:0005244~voltage-gated ion channel activity	2,59	0,0494
GO:0022832~voltage-gated channel activity	2,59	0,0494

* FDR adjusted p-value

Anhang-Tabelle 2: Gene Ontology Term Analyse der intraspezifischen DMRs im Schimpansen. Dargestellt sind die signifikant (FDR adjusted p-value < 0,05) angereicherten „Terms“ für den intraspezifischen Vergleich (N+ vs. N-) im Schimpansen. **A)** Ontologie „Biological process“. **B)** Ontologie „Cellular component“. **C)** „Molecular function“.

A) „Biological process“

Term	Enrichment	p-value*
GO:0044699~single-organism process	1,13	0,0044
GO:0023052~signaling	1,38	0,0044
GO:0007154~cell communication	1,37	0,0052
GO:0007155~cell adhesion	1,89	0,0078
GO:0022610~biological adhesion	1,89	0,0084
GO:0044700~single organism signaling	1,36	0,0156
GO:0044763~single-organism cellular process	1,14	0,0288

* FDR adjusted p-value

B) „Cellular component“

Term	Enrichment	p-value*
GO:0044459~plasma membrane part	1,82	0,0002
GO:0005886~plasma membrane	1,54	0,0002
GO:0071944~cell periphery	1,53	0,0003
GO:0036477~somatodendritic compartment	2,75	0,0006
GO:0097458~neuron part	2,02	0,0042
GO:0030054~cell junction	1,95	0,0095
GO:0045202~synapse	2,31	0,0210
GO:0043005~neuron projection	2,14	0,0235
GO:0044297~cell body	2,77	0,0276
GO:0043025~neuronal cell body	2,90	0,0344
GO:0045211~postsynaptic membrane	3,93	0,0405
GO:0060076~excitatory synapse	3,89	0,0460

* FDR adjusted p-value

C) „Molecular function“

Term	Enrichment	p-value*
GO:0022836~gated channel activity	3,25	0,0352

* FDR adjusted p-value

Anhang-Tabelle 3: Auflistung der 138 Gene assoziiert mit Mentaler Retardierung, die als Referenzliste für den Abgleich mit den DMRs verwendet wurden.

ABI2, ACSL4, ADAT3, ADGRG2, AFF3, ANK3, AP1S2, ARHGEF6, ARX, ATRX, BCAP31, BDNF, BRWD3, C12orf4, CACNG2, CAPS, CC2D1A, CCDC82, CDH15, CHAMP1, COL4A3BP, CRADD, CRBN, CTNNB1, DCPS, DEAF1, DLG3, DMBT1, DPP6, DYNC1H1, DYRK1A, EDC3, EEF1A2, ELP2, EPB41L1, EXTL3, FBX031, FBXO31, FBXO47, FMN2, FTSJ1, GATAD2B, GDI1, GNB1, GPT2, GRIA3, GRIK2, GRIN1, GRIN2B, HERC2, HIVEP2, HNMT, HUWE1, IL1RAPL1, IMPA1, IQSEC2, KAT6A, KDM5C, KIAA0442, KIF1A, KIRREL3, KPTN, LAMC1, LINS1, LMAN2L, LRP2, MAGEA11, MAN1B1, MAP3K7, MAPK8, MBD5, MBOAT7, MECP2, MED23, MEF2C, METTL23, MPDZ, MYT1L, NAGS, NDST1, NSUN2, OPHN1, PAK3, PGAP1, PGAP2, PHF8, PIDD1, PIGG, PPP2R1A, PPP2R5D, PQBP1, PRPS1, PRSS12, PURA, PUS3, RAB39B, RBMX, RPS6KA3, SDK2, SETBP1, SETD5, SLAIN1, SLC6A17, SOX11, SOX3, SPATA13, ST3GAL3, SUMF2, SYNGAP1, SYNRG, TAF2, TBC1D23, TBL1XR1, TECR, TET1, THOC2, TMEM135, TNIK, TRAPPC6B, TRAPPC9, TRIO, TRMT1, TSPAN7, TTI2, TUSC3, UBA7, UPF3B, USP44, VPS35, WASHC4, ZBTB11, ZBTB18, ZC3H14, ZDHHC15, ZMYND11, ZNF41, ZNF81.

Anhang-Tabelle 4: Inter- und intraspezifische DMRs assoziiert mit neuropsychiatrischen Erkrankungen (Alzheimer, Autismus, Bipolare Störungen, Schizophrenie). Angegeben sind die chromosomale Lage der DMRs, die Länge (bp) und Anzahl der abgedeckten CpG Stellen (# CG), der Genname des DMR-überlappenden Gens und die durchschnittliche Methylierung dieser Region in Prozent. **A)** Alzheimer. **B)** Autismus. **C)** Bipolare Störungen. **D)** Schizophrenie.

A) Alzheimer

Spezies	Chromosom	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
							neuronal	nicht-neuronal
Mensch	chr2	10471492	10471644	153	10	<i>HPCAL1</i>	23	67
	chr2	149478091	149478225	135	5	<i>EPC2</i>	12	67
	chr2	222428762	222428939	178	6	<i>EPHA4</i>	5	85
	chr2	222320270	222320370	101	4		15	65
	chr2	234093058	234093207	150	3	<i>INPP5D</i>	7	65
	chr8	27221186	27221340	155	3	<i>PTK2B</i>	54	62
	chr9	79281044	79281199	156	5	<i>PRUNE2</i>	4	53
	chr9	90166704	90166852	156	5	<i>DAPK1</i>	11	75
	chr10	11298767	11298886	120	8	<i>CELF2</i>	8	62
	chr10	131714021	131714143	123	3	<i>EBF3</i>	12	73
	chr14	92800636	92800746	111	3	<i>SLC24A4</i>	11	71
	chr15	75090834	75090971	138	7	<i>CSK</i>	16	61
	chr19	56686224	56686338	115	3	<i>GALP</i>	17	73
chr22	23604903	23605184	282	13	<i>BCR</i>	12	57	
Schimpanse	chrX	91409756	91410029	274	11	<i>PCDH11X</i>	76	6

B) Autismus

Zelltyp	Chromosom	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
							Mensch	Schimpanse
neuronal	chr2	1817284	1818177	894	10	MYT1L	69	32
nicht-neuronal	chr1	27875811	27876033	223	5	AHDC1	57	5
	chr1	7693634	7693747	114	5	CAMTA1	77	21
	chr9	140854361	140854509	149	5	CACNA1B	73	24
	chr15	23892617	23892778	162	9	MAGEL2	44	8
	chr16	1219628	1219845	218	11	CACNA1H	67	29
	chr17	79045375	79045513	139	8	BAIAP2	70	6
	chr17	79022690	79023511	822	5		69	19
	chr20	57429362	57429533	172	5	GNAS	64	6
	chr22	51169861	51172813	2953	4	SHANK3	83	21
	chrX	28632961	28633099	139	5	IL1RAPL1	71	20

Spezies							neuronal	nicht-neuronal
Mensch	chr1	27874287	27875853	1567	33	AHDC1	15	63
	chr1	7407189	7407311	123	9	CAMTA1	21	84
	chr2	1828372	1828483	112	11	MYTL1	1	66
	chr2	2297359	2300439	3081	4		9	61
	chr2	2295414	2295587	174	11		7	60
	chr2	60718637	60718781	145	10	BCL11A	30	87
	chr2	60740385	60740509	125	3		10	80
	chr2	162284524	162284649	126	10	SLC4A10	32	76
	chr3	10420739	10420875	137	3	ATP2B2	17	87
	chr3	10375982	10376307	326	8		4	49
	chr3	10438927	10439047	121	5		27	88
	chr7	127670359	127670506	148	5	SND1	10	59
	chr7	154065566	154072194	6629	11	DPP6	16	72
	chr9	140938140	140939430	1291	5	CACNA1B	22	77
	chr9	141008052	141008445	394	14		5	61
	chr9	140777319	140783025	5707	23		10	73
	chr9	90166704	90166852	149	6	DAPK1	11	75
	chr9	140052457	140052647	191	8	GRIN1	2	36
	chr9	140036019	140036167	149	8		3	64
	chr9	79281044	79281199	156	5	PRUNE2	4	53
	chr9	130438696	130439055	360	11	STXBP1	1	57
	chr10	88027900	88028081	182	7	GRID1	68	34
	chr11	19893172	19893288	117	4	NAV2	5	64
	chr12	2412124	2412251	128	4	CACNA1C	7	75
	chr12	2789263	2790700	1438	8		3	74
	chr16	1271171	1271388	218	19	CACNA1H	8	60
	chr16	1205362	1205469	108	5		21	85
	chr16	1206326	1208560	2235	24		13	70
	chr16	1219628	1219891	264	16		15	56
	chr16	1231958	1232221	264	13		15	57
	chr16	2135987	2136270	284	6	TSC2	21	77
	chr16	6413542	6413691	150	5	RBFOX1	17	67
	chr16	6331221	6331349	129	3		17	84
	chr16	81694699	81694839	141	6	CMIP	3	53
	chr16	10271788	10274042	2255	8	GRIN2A	3	50
	chr17	79050932	79052330	1399	40	BAIAP2	17	75

	chr17	79019527	79023511	3985	4		11	71
	chr17	48703972	48704077	106	5	<i>CACNA1G</i>	36	79
	chr17	8046729	8050801	4073	3	<i>PER1</i>	9	61
	chr19	13525493	13525634	142	6	<i>CACNA1A</i>	10	81
	chr22	24823822	24823951	130	3	<i>ADORA2A</i>	72	14
Schimpanse	chr2	1818160	1818263	104	5	<i>MYTL1</i>	14	95
	chr3	53784412	53784560	149	5	<i>CACNA1D</i>	95	36
	chr5	1446774	1446882	109	3	<i>SLC6A3</i>	51	8
	chr7	127670407	127670630	224	15	<i>SND1</i>	2	51
	chr9	140781066	140782809	1744	5	<i>CACNA1B</i>	9	86
	chr16	2135465	2135987	523	4	<i>TSC2</i>	6	76
	chr16	10001773	10001887	115	6	<i>GRIN2A</i>	11	74
	chr17	79011820	79011920	101	4	<i>BAIAP2</i>	7	72
	chr17	79032885	79033025	141	9		9	82
	chr17	79049134	79049251	118	9		11	53
	chr19	13525493	13525634	142	5	<i>CACNA1A</i>	20	88
	chr19	46274251	46274370	120	7	<i>DMPK</i>	54	17
	chr22	24823822	24823951	130	3	<i>ADORA2A</i>	74	11
	chrX	28632996	28633099	104	5	<i>IL1RAPL1</i>	67	21

C) Bipolare Störungen

Zelltyp	Chromosom	Start	Ende	Länge (bp)	# CG	Genname	Metylierung (%)	
							Mensch	Schimpanse
nicht-neuronal	chr1	177001756	177001904	149	7	<i>ASTN1</i>	89	5
	chr1	236511571	236511701	131	8	<i>EDARADD</i>	81	6
	chr5	390627	390730	104	4	<i>AHRR</i>	76	25
	chr6	33989204	33990187	984	3	<i>GRM4</i>	83	21
	chr7	42004666	42004966	301	14	<i>GLI3</i>	69	24
	chr9	140854361	140854509	149	5	<i>CACNA1B</i>	73	24
	chr10	14623230	14623577	348	6	<i>FAM107B</i>	60	3
	chr12	1957458	1957596	139	8	<i>CACNA2D4</i>	71	28
	chr16	85936631	85939292	2662	10	<i>IRF8</i>	64	9
chr17	42989090	42989653	564	4	<i>GFAP</i>	76	32	

Spezies	Chromosom	Start	Ende	Länge (bp)	# CG	Genname	neuronal	nicht-neuronal
							neuronal	nicht-neuronal
Mensch	chr1	5923501	5925501	2001	11	<i>NPHP4</i>	6	76
	chr1	6046227	6046872	646	6		15	86
	chr1	5969661	5969783	123	7		9	81
	chr1	2038356	2040615	2260	11	<i>PRKCZ</i>	5	66
	chr1	2077728	2077844	117	5		14	76
	chr1	53785299	53788107	2809	3	<i>LRP8</i>	13	73
	chr1	154544691	154548824	4134	4	<i>CHRNA2</i>	8	82
	chr1	208391222	208391329	108	11	<i>PLXNA2</i>	14	70
	chr2	1828372	1828483	112	11	<i>MYT1L</i>	1	66
	chr2	2297359	2300439	3081	4		9	61
	chr2	2295414	2295587	174	11		7	60
	chr2	60740385	60740509	125	3	<i>BCL11A</i>	10	80
	chr2	60718637	60718781	145	10		30	87
	chr2	106509009	106510137	1129	4	<i>NCK2</i>	16	79
	chr3	37536589	37540885	4297	10	<i>ITGA9</i>	18	76

	chr3	66002247	66002949	703	5	MAG11	15	80
	chr3	178096892	178097157	266	5	KCNMB2	23	83
	chr4	54430900	54431064	165	8	LNX1	54	91
	chr5	664529	666800	2272	10	CEP72	17	91
	chr5	168649978	168650196	219	4	SLIT3	21	79
	chr6	162715005	162715186	182	3	PARK2	70	10
	chr6	35346474	35346610	137	5	PPARD	30	55
	chr6	155577670	155577788	119	8	TIAM2	16	54
	chr6	44243763	44244079	317	44	TMEM151 B	3	54
	chr7	151422838	151422951	114	4	PRKAG2	3	79
	chr7	151492360	151492487	128	6		20	89
	chr7	4299840	4300306	467	11	SDK1	28	88
	chr7	18614381	18614560	180	6	HDAC9	21	92
	chr7	158887190	158887333	144	3	VIPR2	26	93
	chr8	134471925	134472027	103	6	ST3GAL1	19	86
	chr9	140777319	140783025	5707	23	CACNA1B	10	73
	chr9	140938140	140939430	1291	5		22	77
	chr9	141008052	141008445	394	14		5	61
	chr9	137990580	137993107	2528	4	OLFM1	3	45
	chr9	137997125	137997293	169	7		3	40
	chr9	138000340	138001189	850	9		12	70
	chr10	127887036	127887198	163	4	ADAM12	3	74
	chr10	123952255	123954658	2404	3	TACC2	78	25
	chr11	19893172	19893288	117	4	NAV2	5	64
	chr11	64391574	64391698	125	4	NRXN2	33	86
	chr12	1943676	1943854	179	13	CACNA2D 4	30	90
	chr12	111800972	111801126	155	16	FAM109A	6	74
	chr14	105236377	105238511	2135	5	AKT1	5	63
	chr16	10271788	10274042	2255	8	GRIN2A	3	50
	chr16	89255474	89255720	247	6	CDH15	21	78
	chr20	62051078	62052738	1661	16	KCNQ2	13	66
	chr20	62073091	62076065	2975	5		9	65
	chr20	62046217	62048848	2632	13		24	82
	chr20	62097766	62097991	226	20		26	64
	chr20	62100956	62101145	190	9		5	48
	chr20	36769674	36771872	2199	8	TGM2	7	74
	chr22	37022212	37022374	163	8	CACNG2	2	82
	chr22	19798368	19799734	1367	7	GNB1L	2	64
	chr22	26857119	26858160	1042	4	HPS4	8	71
Schimpanse	chr1	5921766	5925436	3671	12	NPHP4	5	76
	chr1	2046744	2047580	837	11	PRKCZ	9	47
	chr1	230849519	230849661	143	4	AGT	66	13
	chr2	1818160	1818263	104	5	MYT1L	14	95
	chr3	53784412	53784560	149	5	CACNA1D	95	36
	chr5	151261260	151261371	112	4	GLRA1	83	7
	chr7	4299623	4300026	404	10	SDK1	22	88
	chr7	151422839	151422943	105	3	PRKAG2	2	77
	chr8	134471925	134472027	103	6	ST3GAL1	23	93
	chr9	140781066	140782809	1744	5	CACNA1B	9	86
	chr9	137996152	137996925	774	5	OLFM1	7	60
	chr9	137997110	137998034	925	14		10	62
	chr10	114752132	114752348	217	4	TCF7L2	73	3
	chr15	93598561	93600276	1716	4	RGMA	85	21
	chr16	10001773	10001887	115	6	GRIN2A	11	74

	chr20	62100956	62101138	183	8	<i>KCNQ2</i>	6	52
	chr22	50187974	50189095	1122	4	<i>BRD1</i>	15	78

D) Schizophrenie

Zelltyp	Chromosom	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
							Mensch	Schimpanse
neuronal	chr11	57591001	57591112	112	5	<i>LRP1</i>	17	70
nicht-neuronal	chr1	243637718	243637844	127	10	<i>SDCCAG8</i>	89	42
	chr5	137593424	137593714	291	18	<i>GFRA3</i>	67	18
	chr7	1902882	1903056	175	8	<i>MAD1L1</i>	75	27
	chr11	113283246	113283362	117	6	<i>DRD2</i>	59	4
	chr16	67918866	67919928	1063	5	<i>NRN1L</i>	67	5
	chr17	2271238	2271389	152	3	<i>SGSM2</i>	73	9
	chr19	19368779	19368990	212	25	<i>HAPLN4</i>	72	17
Spezies							neuronal	nicht-neuronal
Mensch	chr1	2371057	2371183	127	5	<i>PLCH2</i>	15	59
	chr1	2366178	2366284	107	6		33	83
	chr2	233766535	233766659	125	7	<i>NGEF</i>	14	64
	chr2	233757595	233759265	1671	15		30	86
	chr2	233784952	233785142	191	13		17	73
	chr2	233746844	233749813	2970	4		29	86
	chr2	233750056	233750522	467	8		6	67
	chr2	162284524	162284649	126	10	<i>SLC4A10</i>	32	76
	chr2	149478091	149478225	135	5	<i>EPC2</i>	12	67
	chr3	52524923	52526534	1612	13	<i>NISCH</i>	6	57
	chr3	17734783	17734904	122	5	<i>TBC1D5</i>	28	94
	chr7	2170480	2170615	136	9	<i>MAD1L1</i>	12	79
	chr8	143369845	143372211	136	9	<i>TSNARE1</i>	10	52
	chr8	143334442	143334684	243	13		19	82
	chr8	143414301	143414431	131	8		18	84
	chr11	46395976	46396096	121	7	<i>DGKZ</i>	1	76
	chr11	46373975	46374094	120	4		3	43
	chr11	46374944	46378048	3105	3		20	87
	chr12	2412124	2412251	128	4	<i>CACNA1C</i>	7	75
	chr12	2789263	2790700	1438	8		3	74
	chr14	103996769	103999006	2238	4	<i>TRMT61A</i>	19	74
	chr14	104000054	104000631	578	31	<i>MAU2</i>	4	38
	chr14	104191360	104191555	196	6	<i>ZFYVE21</i>	16	66
	chr14	104158553	104158687	135	6	<i>KLC1</i>	25	72
	chr14	104153351	104153523	173	10		3	54
	chr15	91419253	91419490	238	5	<i>FURIN</i>	0	51
chr16	10271788	10274042	2255	8	<i>GRIN2A</i>	3	50	
chr16	30017306	30017447	142	6	<i>DOC2A</i>	13	62	
chr16	30018671	30018802	132	3		6	64	
chr17	2282729	2282959	231	13	<i>SGSM2</i>	12	56	
chr19	19612190	19612306	117	12	<i>GATAD2A</i>	9	67	
chr22	40005520	40005621	102	3	<i>CACNA1I</i>	8	70	
Schimpanse	chr2	233784974	233785142	169	10	<i>NGEF</i>	14	69
	chr11	133777787	133778939	1153	6	<i>IGSF9B</i>	8	65
	chr12	123472556	123472750	195	3	<i>PITPNM2</i>	24	91
	chr14	103998923	104000404	1482	12	<i>TRMT61A</i>	5	71

chr14	104191347	104196144	4798	3	ZFYVE21	6	69
chr15	40544483	40544590	108	10	PAK6	2	40
chr16	10001773	10001887	115	6	GRIN2A	11	74
chr19	19639903	19640017	115	6	NDUFA13	0	46
chr22	40058564	40058723	160	13	CACNA1I	41	3

Anhang-Tabelle 5: Intraspezifische DMRs, assoziiert mit human-spezifischer Regulation im Gehirn. Angegeben sind die chromosomale Lage der DMRs, die Länge (bp), Anzahl der abgedeckten CpG Stellen (# CG), der Genname des DMR-überlappenden Gens und die durchschnittliche Methylierung dieser Region in Prozent.

Spezies	human-spezifische Regulation	Chromosom	Start	Ende	Länge (bp)	# CG	Genname	Methylierung (%)	
								neuronal	nicht-neuronal
Mensch	H3K4me3	chr1	1022574	1023456	883	12	C1orf159	2	61
	H3K4me3	chr1	2119615	2119773	159	5	C1orf86	21	70
	H3K4me3	chr1	6143628	6149641	6014	14	KCNCB2	15	61
		chr1	6121160	6121388	229	6		7	68
		chr1	6131984	6132186	203	4		17	80
	Expression	chr1	9791279	9796183	4905	26	CLSTN1	20	78
	H3K4me3	chr1	10799025	10799303	279	12	CASZ1	43	89
		chr1	10705100	10705208	109	9		2	77
		chr1	10699452	10699552	101	17		14	46
		chr1	10695607	10697616	2010	12		12	49
	Expression	chr1	32201183	32203138	1956	5	BAI2	13	62
		chr1	32208440	32208618	179	14		26	77
	H3K4me3	chr1	33231570	33236472	4903	16	KIAA1522	17	77
	Expression	chr1	51727483	51727621	139	7	RNF11	4	74
	H3K4me3	chr1	167095044	167096117	1074	29	DUSP27	7	76
	H3K4me3	chr2	1828372	1828483	112	11	MYT1L	1	66
		chr2	2297359	2300439	3081	4		9	61
		chr2	2295414	2295587	174	11		7	60
	Expression	chr2	9352875	9353041	167	8	ASAP2	10	59
	H3K4me3	chr2	3296387	3297422	1036	4	TSSC1	21	87
	Expression	chr3	9799147	9799362	216	7	OGG1,CAMKI	27	75
	Expression	chr3	40137679	40137807	129	4	MYRIP	29	75
	Expression, H3K4me3	chr3	49689366	49690213	848	34	BSN	17	83
	Expression	chr3	66002247	66002949	703	5	MAG1	15	80
	H3K4me3	chr3	75827718	75827965	248	4	ZNF717	17	78
		chr3	75827263	75827404	142	5		17	80
	Expression	chr3	93716756	93716892	137	5	ARL13B	17	90
	Expression	chr3	127425757	127425861	105	5	MGLL	14	76
	Expression	chr3	129317949	129318057	109	10	PLXND1	4	53
	Expression	chr3	178096892	178097157	266	5	KCNMB2	23	83
	Expression	chr4	57532179	57532317	139	3	HOPX	92	26
	Expression	chr4	183247002	183247133	132	5	TENM3	14	60
	Expression	chr5	73931954	73932130	177	5	ENC1	9	61
	Expression	chr5	168649978	168650196	219	4	SLIT3	21	79
	Expression	chr6	162715005	162715186	182	3	PARK2	70	10
	Expression	chr7	4980395	4980565	171	7	MMD2	66	9
	Expression	chr7	44278050	44278169	120	8	CAMK2B	5	79
		chr7	44279101	44279853	753	25		19	81
		chr7	44306515	44306619	105	8		26	90

	chr7	44308486	44308887	402	4		9	71
	chr7	44317254	44321187	3934	3		11	87
Expression	chr7	73129312	73130555	1244	9	<i>STX1A</i>	6	82
Expression	chr7	154065566	154072194	6629	11	<i>DPP6</i>	16	72
H3K4me3	chr7	157199249	157199386	138	6	<i>DNAJB6</i>	1	47
	chr7	157209657	157210569	913	6		29	93
Expression	chr7	157451655	157451851	197	3	<i>PTPRN2</i>	21	74
	chr7	157494734	157494934	201	7		25	89
	chr7	157516154	157516303	150	6		24	66
	chr7	157744646	157746218	1573	15		5	68
	chr7	158019266	158019386	121	7		10	72
	chr7	158061718	158061821	104	5		3	87
	chr7	158136324	158136487	164	6		4	83
	chr7	158197857	158199151	1295	8		24	84
	chr7	158281139	158281656	518	16		19	80
	chr7	158363369	158363481	113	18		12	85
Expression	chr8	1781830	1782233	404	4	<i>ARHGEF10</i>	93	46
	chr8	1849599	1853008	3410	9		4	60
	chr8	1878933	1879119	187	7		34	82
	chr8	1879708	1879810	103	9		34	86
	chr8	1896437	1896570	134	8		4	67
H3K4me3	chr8	2032571	2033532	962	9	<i>MYOM2</i>	21	89
Expression	chr9	101366335	101366440	106	6	<i>GABBR2</i>	4	35
	chr9	101079307	101079571	265	5		8	66
H3K4me3	chr9	133892451	133892722	272	19	<i>LAMC3</i>	2	52
Expression	chr9	135084726	135084899	174	15	<i>NTNG2</i>	27	59
	chr9	135118487	135118631	145	4		7	68
H3K4me3	chr10	574573	575141	569	4	<i>DIP2C</i>	8	66
H3K4me3	chr10	697178	698645	138	5		30	83
Expression	chr10	18278569	18278704	136	5	<i>SLC39A12</i>	37	86
Expression	chr10	43758300	43759269	970	9	<i>RASGEF1A</i>	6	56
Expression	chr10	133449980	133450090	111	8	<i>TCERG1L</i>	5	64
	chr10	133039313	133039413	101	3		36	86
H3K4me3	chr11	856982	859580	2599	39	<i>TSPAN4</i>	13	74
H3K4me3	chr11	1268577	1269840	1264	4	<i>MUC5B</i>	10	71
H3K4me3	chr11	3169396	3174129	4734	4	<i>OSBPL5</i>	11	67
	chr11	3147548	3147966	419	7		9	66
Expression	chr11	45924061	45924235	175	10	<i>MAPK8IP1</i>	4	59
Expression	chr11	46374944	46378048	3105	3	<i>DGKZ</i>	20	87
	chr11	46395976	46396096	121	7		1	76
	chr11	46373975	46374094	120	4		3	43
Expression	chr11	72326271	72330016	3746	7	<i>PDE2A</i>	18	72
Expression	chr11	131758173	131758278	106	4	<i>NTM</i>	18	70
H3K4me3	chr12	2412124	2412251	128	4	<i>CACNA1C</i>	7	75
	chr12	2789263	2790700	1438	8		3	74
Expression	chr12	49951280	49951382	103	3	<i>KCNH3, MCRS1</i>	17	72
Expression	chr12	125621302	125621477	176	24	<i>AACS</i>	2	74
H3K4me3	chr12	132682609	132687677	5069	15	<i>GALNT9</i>	24	72
	chr12	132698314	132698952	639	20		14	75
	chr12	132872656	132872760	105	6		23	82
	chr12	132879850	132879968	119	4		31	83
	chr12	132884996	132885148	153	11		27	69
	chr12	132885650	132887039	1390	19		42	55
	chr12	132896253	132898054	1802	16		34	76
H3K4me3	chr13	27699763	27699923	161	8	<i>USP12</i>	11	83
H3K4me3	chr13	113719802	113719917	116	6	<i>MCF2L</i>	5	53
	chr13	113706721	113706838	118	6		3	53

	Expression	chr14	24542756	24542900	145	8	<i>CPNE6</i>	13	82
	Expression	chr14	104191360	104191555	196	6	<i>ZFYVE21</i>	16	66
	Expression	chr15	26874224	26874350	127	13	<i>GABRB3</i>	8	57
	Expression	chr15	77948268	77950378	2111	8	<i>LINGO1</i>	15	51
		chr15	77906235	77908806	2572	34		22	76
	H3K4me3	chr15	101483002	101483212	211	8	<i>LRRK1</i>	9	79
	H3K4me3	chr15	101945022	101946475	1454	3	<i>PCSK6</i>	10	61
	Expression	chr16	1412007	1412116	110	13	<i>GNPTG</i>	12	66
	H3K4me3	chr16	2761934	2765210	3277	3	<i>PRSS27</i>	5	49
	H3K4me3	chr16	6331221	6331349	129	3	<i>RBFOX1</i>	17	84
		chr16	6413542	6413691	150	5		17	67
	Expression	chr16	61995143	61995304	162	5	<i>CDH8</i>	34	94
	H3K4me3	chr16	81694699	81694839	141	6	<i>CMIP</i>	3	53
	Expression	chr16	83191266	83193656	2391	6	<i>CDH13</i>	12	72
	H3K4me3	chr16	83979824	83984819	4996	8	<i>OSGIN1</i>	18	66
	H3K4me3	chr16	87741116	87741293	178	6	<i>KLHDC4</i>	29	89
		chr16	87736879	87737395	517	11		14	81
	Expression	chr17	2703783	2703900	118	6	<i>RAP1GAP2</i>	14	68
		chr17	2712338	2713453	1116	10		11	78
		chr17	2726263	2726441	179	3		25	90
		chr17	2796953	2797054	102	8		5	67
		chr17	2811426	2811526	101	6		23	91
		chr17	2867220	2867343	124	6		28	89
	Expression	chr17	12821287	12821431	145	5	<i>ARHGAP44</i>	4	75
	Expression	chr17	19315865	19318778	2914	17	<i>RNF112</i>	21	74
	H3K4me3	chr17	8046729	8050801	4073	3	<i>PER1</i>	9	61
	Expression	chr17	27903098	27903557	460	4	<i>GIT1</i>	27	79
	Expression	chr17	43192784	43192898	115	5	<i>PLCD3</i>	2	62
	H3K4me3	chr19	651964	652201	238	14	<i>RNF126</i>	5	74
		chr19	652442	652786	345	12		10	64
		chr19	658017	658270	259	10		18	56
		chr19	658398	659071	674	6		11	66
		chr19	660357	660499	143	5		1	66
	Expression	chr19	3635017	3638543	3527	6	<i>PIP5K1C</i>	18	78
		chr19	3659912	3660027	1167	7		33	91
		chr19	3694278	3696527	2250	14		16	75
	Expression	chr19	47968997	47969100	104	5	<i>SLC8A2</i>	8	62
	Expression	chr19	55687454	55690188	2735	3	<i>SYT5</i>	8	44
		chr19	55705057	55705180	124	4		5	48
	Expression	chr19	55785660	55785939	280	18	<i>HSPBP1</i>	6	58
		chr19	55783844	55785549	1706	3		4	58
	H3K4me3	chr20	371097	371257	161	6	<i>TRIB3</i>	14	84
	H3K4me3	chr20	3759857	3759960	104	11	<i>SPEF1</i>	3	43
	H3K4me3	chr20	62917579	62917777	199	5	<i>PCMTD2</i>	5	86
	Expression	chr22	26829939	26831422	1484	6	<i>ASPHD2</i>	22	75
	Expression	chr22	36973023	36973376	354	6	<i>CACNG2</i>	11	80
		chr22	37022212	37022374	163	8		21	82
	Expression	chr22	38622697	38622839	143	6	<i>TMEM184B</i>	15	64
	H3K4me3	chr22	46850547	46851605	1059	4	<i>CELSR1</i>	8	69
	H3K4me3	chr22	47367948	47369318	1371	10	<i>TBC1D22A</i>	37	69
Schimpanse	H3K4me3	chr1	10742351	10746979	4629	4	<i>CASZ1</i>	84	54
	Expression	chr1	97,91,300	9796183	4884	11	<i>CLSTN1</i>	19	92
	Expression	chr1	156616814	156617858	1045	11	<i>BCAN</i>	75	32
	Expression	chr1	230849519	230849661	143	4	<i>AGT</i>	66	13
	H3K4me3	chr2	1818160	1818263	104	5	<i>MYT1L</i>	14	95
	H3K4me3	chr3	75826348	75827894	1547	3	<i>ZNF717</i>	54	3
		chr3	75829292	75829467	176	6		57	2

Expression	chr3	127425757	127425861	105	5	<i>MGLL</i>	12	80
Expression	chr5	73931985	73932130	146	4	<i>ENC1</i>	9	81
Expression	chr5	155389707	155389835	129	3	<i>SGCD</i>	80	9
H3K4me3	chr7	4299623	4300026	404	10	<i>SDK1</i>	22	88
Expression	chr8	1788181	1788341	161	3	<i>ARHGEF10</i>	88	31
H3K4me3	chr10	393762	393980	219	5	<i>DIP2C</i>	40	93
H3K4me3	chr10	114752132	114752348	217	4	<i>TCF7L2</i>	73	3
Expression	chr10	132974344	132974451	108	3	<i>TCERG1L</i>	7	88
H3K4me3	chr11	856970	858836	1867	17	<i>TSPAN4</i>	5	82
	chr11	859379	859671	293	10		6	79
H3K4me3	chr12	132900117	132901119	1003	15	<i>GALNT9</i>	5	76
	chr12	132901578	132903510	1933	9		16	82
	chr12	132904685	132905552	868	17		5	47
Expression	chr14	104191347	104196144	4798	3	<i>ZFYVE21</i>	6	69
Expression	chr15	27087554	27087664	111	3	<i>GABRB3</i>	21	77
Expression	chr15	67467407	67467507	101	3	<i>SMAD3</i>	33	90
Expression	chr16	27863895	27864072	178	3	<i>GSG1L</i>	23	77
Expression	chr16	70680815	70680985	171	8	<i>IL34</i>	4	50
H3K4me3	chr16	87736879	87737001	123	7	<i>KLHDC4</i>	17	65
Expression	chr19	55785549	55785700	152	5	<i>HSPBP1</i>	12	91

Anhang-Tabelle 6: Übersicht aller DMR-assoziierten L1 Elemente. Angegeben ist jeweils die chromosomale Lage (basierend auf HG37) der repetitiven Elemente (REs), sowie der differenziell methylierten Regionen (DMRs). Zu jeder DMR sind neben dem in der Region gelegenen Gen auch die upstream und downstream lokalisierten Gene angegeben. Alle dick hervorgehobenen Gene können mittels Literatur in funktionellen Zusammenhang mit dem Gehirn gebracht werden.

Name RE	Lage REs	Lage DMRs	upstream	überlappend	downstream	Methylierung (%)		
						Neuronal	Nicht-Neuronal	
Mensch NUM	L1M4	chr3:37538303-37538732;chr3:37538962-37539161	chr3:37536589-37540885	C3orf35	ITGA9	RNU7-73P	18	76
		chr7:1037881-1038090;chr7:1038381-1038587;chr7:1038873-1039072	chr7:1528740-1533963	CYP2W1	C7orf50	MIR339	6	78
	L1M4b	chr22:49653106-49653293	chr:2249651658-49654164	RPL35P8		C22orf34	18	85
		L1M5	chr1:1973799-1973912	chr1:1972875-1974300	GABRD		RP11-547D24.3	19
	chr10:134117112-134117348		chr10:3319189-3319377	RP11-195B3.1		RP11-482E14.1	12	86
	L1MA1	chr10:134117562-134117711	chr10:134115790-134120496	DPYSL4	STK32C	LRRC27	5	36
		chr22:49652164-49652468	chr22:49651658-49654164	RPL35P8		C22orf34	18	85
	L1MA6	chr22:21082283-21082544	chr22:21081518-21083979	AC007308.6	PI4KA	SERPIND1	5	73
	L1MB2	chr17:921593-921652	chr17:920227-923771	TIMM22	ABR	MIR3183	8	74
	L1MB3	chr17:77080155-77080415	chr17:77079383-77081799	CIQTNF1	ENGASE	ENGASE	18	81
		chr2:3132460-3132738	chr2:31326733136412	ACO19118.2	ACO19118.3	ACO19118.2	43	82
	L1MB5	chr6:21317253-21317418	chr6:21313624-21318374	CDKAL1		RNU6-150P	12	75
		chr11:64634307-64634565	chr11:64631980-64635884	CDC42BPG	EHD1	AP001187.11	46	59
		chr17:74915816-74916009	chr17:74906329-74920018	LINC00868	MGAT5B	RP11-87G24.3	5	66
		chr3:66002407-66002828	chr3:66002247-66002949	AC106827.1	MAG11	SLC25A26	15	80
	L1MB7	chr7:154070706-154071073	chr7:154065566-154072194	ACO06019.4	DPP6	RP11-476H24.1	16	72
		chr10:134797102-134797613	chr10:134797550-134797728	LINC01168		GPR123	1	58
		chr12:121030154-121030701	chr12:121029499-121034287	POP5	RPL11P5	RP11-728G15.1	5	83
		chr12:132326998-132327036;chr12:132327337-132327547;chr12:132328092-132328323	chr12:132326992-132328729	RNU6-1017P	MMP17	RP11-417L19.2	40	89
		chr16:1704511-1704836	chr16:1704611-1704742	LA16c-395F10.1	CRAMP1L;LA16c-431H6.6	HN1L	7	51
chr3:37537180-37537307		chr3:37536589-37540885	C3orf35	ITGA9	RNU7-73P	18	76	
L1MB8	chr4:8068162-8068650	chr4:8068238-8069458	RP11-338K13.1	RP11-1258F18.1;ABLIM2	GMPSP1	7	74	
	chr17:29022335-29022618	chr17:29021589-29024229	AC005562.1		RN75L316P	13	60	
	chr17:79033969-79034383;chr17:79034418-79034514	chr17:79032885-79038320	BAIAP2-AS1	BAIAP2	AATK	10	82	
	chr17:79050785-79050991	chr17:79050932-79052330	BAIAP2-AS1	BAIAP2	AATK	17	75	
	chr19:2765865-2765936	chr19:2765358-2766071	SLC39A3	SGTA	THOP1	7	72	
	chr19:6232302-6232602;chr19:6232682-6232891	chr19:6229139-6232929	CTC-503J8.6	MLLT1	CTC-503J8.4	4	56	
	chr19:55783462-55783926;chr19:55784237-55784355;chr19:55784658-55784729	chr19:55783844-55785549	PPP6R1	HSPBP1	BRSK1	4	58	
	chr22:47368278-47368683	chr22:47367948-47369318	Z97351.1	TBC1D22A	CITF22-49D8.1	37	69	
	chr8:94901377-94901507	chr8:94901267-94901386	RP3-388N13.2	PDP1	MIR378D2	35	88	
	chr14:100215026-100215780	chr14:100215528-100215652	CYP46A1	EML1	RNU1-47P	12	78	
L1MC1	chr17:79062172-79062553	ch17:79060745-79066325	BAIAP2-AS1	BAIAP2	AATK	13	79	
	chr7:213433-214043;chr7:214339-214381	chr7:213696-214921	AC093627.12	FAM20C	AC145676.2	14	61	
	chr7:1530893-1531246;chr7:1531556-1532242	chr7:1528740-1533963	RP11-1246C19.1	INTS1	AC102953.6	18	68	
	chr9:132981261-132981346	chr9:132975862-132982367	AL360004.1	NCS1	HMCN2	19	78	
L1MC2	chr8:1850670-1850956	chr8:1849599-1853008	MIR596	ARHGFB10	KBTBD11-OT1	4	60	
L1MC3	chr17:58231854-58232001	chr17:58231197-58233639	CTD-2319I12.5	CA4	FAM106DP	9	70	
	chr22:49062154-49062309;chr22:49062416-49062539	chr22:49060730-49062509	CTA-299D3.8	FAM19A5	MIR4535	11	73	
L1MC4	chr8:145814201-145814624	chr8:145813527-145814465	CTD-2517M22.9	ARHGAP39	AF186192.1	8	75	
	chr1:154546653-154547023	chr1:154544691-154548824	UBE2Q1	CHRN2	RP11-61L14.6	8	81	
L1MC4	chr11:70089773-70090008	chr11:70086652-70090124	RP11-805I14.5;FADD		PPFIA1	14	74	
	chr2:99157465-99157508;chr2:99157806-99157902	chr2:99155916-99159945	CNGA3	INPP4A	COA5	14	77	
	chr7:674216-674280;chr7:674577-674763	chr7:672307-674963	AC147651.4	PRKAR1B	HEATR2	6	53	
L1MC4	chr9:140724526-140725011;chr9:140725684-140725847	chr9:140723061-140727855	RP11-188C12.2	EHMT1	MIR602	2	56	

Mensch	NUM	L1MC4a	chr10:133777999-133778557;chr10:133779451-133779690 chr13:21293695-21293917 chr7:672787-672982	chr10:133777895-133782644 chr13:21291275-21295529 chr7:672307-674963	PPP2R2D AL512652.1 AC147651.4	BNIP3 IL17D PRKAR1B	JAKMIP3 N6AMT2 HEATR2	7 13 6	72 62 53	
		L1MC5	chr2:233757914-233758067 chr7:44318660-44318802;chr7:44318825-44318915 chr9:140724419-140724469	chr2:233757595-233759265 chr7:44317254-44321187 chr9:140723061-140727855	C2orf82 VK76 RP11-188C12.2	NGEF CAMK2B EHMT1	AC106876.2 NUDCD3 MIR602	30 11 2	86 87 56	
		L1MD	chr10:134055029-134055141 chr10:134803881-134804164	chr10:134053498-134056965 chr10:134803507-134806093	DPYSL4 LINCO1168	STK32C	LRRC27 GPR123	18 13	78 77	
		L1MD2	chr4:1228774-1229113 chr9:140615055-140615585;chr9:140616142-140616545	chr4:1228936-1229135 chr9:140614916-140617481	SPON2 SETP5	CTBP1 EHMT1	CTBP1-AS2 RP11-188C12.2	19 9	73 79	
		L1ME1	chr17:79035129-79035500;chr17:79035513-79035746;chr17:79036049-79036149 chr4:3204237-3204456;chr4:3204751-3204905 chr9:139661749-139661860;chr9:139662184-139662516;chr9:139662581-139662880	chr17:79032885-79038320 chr4:3201467-3205977 chr9:139661597-139663524	BAIAP2-AS1 HTT-AS LCN15	BAIAP2 HTT	AATK MSANTD1 ATP6V1G1P3	10 3 36	82 60 79	
		L1ME2z	chr10:134802874-134803871 chr20:44808197-44808282 chr9:140723775-140723886	chr10:134803507-134806093 chr20:44806684-44810899 chr9:140723061-140727855	LINCO1168 RP5-998H6.2 RP11-188C12.2	CDH22 EHMT1	SLC35C2 MIR602	11 2	80 56	
		L1ME3	chr17:29022631-29022800;chr17:29023295-29023672 hr22:49060427-49061116;chr22:49061135-49061462	chr17:29021589-29024229 chr22:49060730-49062509	AC005562.1 CTA-299D3.8	FAM19A5	RN75L316P MIR4535	13 11	60 73	
		L1ME3A	chr2:10972269-10972509;chr2:10972690-10972764	chr2:10971398-10975309	RNU7-176P	PDIA6	AC092687.4	5	63	
		L1ME3C	chr11:46376210-46376636	chr11:46374944-46378048	CREB3L1	DGKZ	MIR4688	20	87	
		L1ME3E	chr4:2723382-2723846	chr4:27234192723553	RNF4	FAM193A	TNIP2	12	73	
		L1ME3F	chr17:79037096-79037258	chr17:79032885-79038320	BAIAP2-AS1	BAIAP2	AATK	10	82	
		L1ME4a	chr7:3326042-3326174	chr7:3325441-3326373	AC073316.2	AC073316.1	SDK1	19	83	
		L1MEa	chr9:140725959-140726208	chr9:140723061-140727855	RP11-188C12.2	EHMT1	MIR602	2	56	
		L1MEd	chr1:1906735-1906976;chr1:1907182-1907237 chr6:21314987-21315122	chr1:1906591-1907388 chr6:21313624-21318374	TMEM52 CDKAL1	C1orf222	RP11-547D24.1 RP1-135L22.1	18 12	86 75	
		L1MEg	chr16:2329693-2329743;chr16:2330049-2330178 chr7:728576-728758;chr7:729057-729113	chr16:2329106-2333054 chr7:727590-732295	MIR4717 AC147651.4	ABCA3 PRKAR1B	RP11-715J22.1 HEATR2	11 8	81 90	
		L1P2	chr13:27699624-27700568	chr13:27699763-27699923	RPS20P32	USP12	USP12-AS1	11	83	
		L1PA11	chr2:149472902-149478507	chr2:149478091-149478225	AC017080.1	EPC2	RNU2-9P	12	67	
		L1PA15	chr11:64633184-64633885;chr11:64634177-64634291	chr11:64631980-6435884	CDC42BPG	EHD1	AP001187.11	46	59	
		L1PA16	chr7:729949-730113	chr7:727590-732295	AC147651.4	PRKAR1B	HEATR2	8	90	
		L1PA3	chr3:17734606-17741044 chr4:54430805-54436865 chr4:107134512-107140537	chr3:17734783-17734904 chr4:54430900-54431064 chr4:107134511-107134615	U7 LNK1-AS1 RP11-710F7.3	TBC1D5 FIP1L1;LNK1 TBCK	AC104451.2 RP11-18M17.1 AIMP1	28 54 42	94 91 94	
		L1PA5	chr11:44946303-44946453	chr11:44945002-44949785	RP11-45A12.2	TSPAN18;TP5311	AC068858.1	46	70	
		L1PA8	chr2:3133292-3133578 chr8:1850154-1850381	chr2:3132673-3136412 chr8:1849599-1853008	AC019118.2 MIR596	AC019118.3 ARHGEF10	AC019118.4 KBTBD11-OT1	43 4	82 60	
		L1PB4	chr12:101880933-101881146	chr12:101880173-101881553	RNA5SP367	SPIC	MYBPC1	13	86	
		N-NUM	L1M5	chr12:114244393-114244708	chr12:114243041-114245986	RP11-438N16.2		RBM19	72	18
			L1MC4	chr19:838927-839059	chr19:838837-839019	AZU1		PRTN3	75	28
			L1MD2	chr8:10417673-10417834;chr8:10418020-10418604	chr8:10417271-10422048	PRSS55		RP1L1	77	29
			L1ME2z	chr6:163754029-163754170;chr6:163754218-163754394;chr6:163754687-163755044	chr6:163752822-163756922	PACRG		RP3-495O10.1	87	29
			L1PBa	chr12:101880933-101881146	chr12:31176229-31176335	RP11-551L14.6	DDX11-AS1	DDX11	69	14
Schimpanse	NUM	L1M2	chr4:1276066-1276493	chr4:1276483-1277541	CTBP1	CTBP1-AS2	MAEA	17	92	
		L1M5	chr21:47812712-47812919	chr21:47809153-47813461	C21orf58	PCNT	RPL18AP2	36	62	
		L1MA1	chr22:21082283-21082544	chr22:21081518-21083723	AC007308.6	PI4KA	SERPIND1	3	82	
		L1MB3	chr14:104856430-104856621	chr14:104854711-104859567	RP11-260M19.1		RNU6-684P	23	56	
		L1MB8	chr12:131263207-131263349 chr16:1333199-1333415	chr12:131261257-131265721 chr16:1333388-1333501	R1MBP2;RP11-662M24.2 RP11-616M22.9		STX2 TPSP2	19 13	66 52	
		L1MC	chr21:47812004-47812127	chr21:47809153-47813461	C21orf58	PCNT	RPL18AP2	36	62	
		L1MC4	chr2:99157465-99157508;chr2:99157806-99157902	chr2:99156051-99159855	CNGA3	INPP4A	COA5	17	95	
		L1MD2	chr3:13093407-13094275	chr3:13093944-13094080	RP11-767C1.1	IQSEC1	RP11-715I4.1	7	54	
		L1MD3	chr11:1320068-1320429	chr11:1319585-1324221	AC136297.1	TOLLIP	TOLLIP-AS1	1	70	
		L1ME1	chr12:131813263-131814363	chr12:131814119-131814247	RNA5SP376		RP13-507P19.2	9	80	
L1PA4	chr2:42014295-42014829	chr2:42012522-42017257	snoZ247		LDHAP3	24	64			

	N-NUM							Schimpanse	
Schimpanse	L1MB3	chr15:20457780-20458267;chr15:20458733-20458830	chr15:20456739-20461236	BMS1P15	RP11-492D6.3	CHEK2P2	66	7	
	L1MB7	chr18:77621583-77623131	chr18:77622798-77622914	RP11-154H12.3		KCNG2	88	49	
	L1MC	chr12:131918999-131919089;chr12:131919636-131919820	chr12:131917878-131921744	RP13-507P19.1		RP11-897M7.1	92	29	
	L1MC4a	chr19:43653450-43653768	chr19:43653560-43653671	CEACAMP9		AC005392.13	61	3	
		chr20:55972958-55973276	chr20:55970900-55974770	RP4-800J21.3		RBM38	67	2	
	L1MC5	chr17:37905265-37905712;chr17:37905798-37905820;chr17:37906114-37906253	chr17:37903730-37907972	GRB7		RP11-94L15.2	85	38	
	L1ME1	chr17:37904221-37904417	chr17:37903730-37907972	GRB7		RP11-94L15.2	85	38	
	L1ME3	chr11:2280119-2280319	chr11:2278774-2283048	MIR4686		ASCL2	72	16	
	L1ME4a	chr22:48534108-48534719	chr22:48533788-48537344	CTA-280A3_B.2	LL22NC03-121E8.4	LL22NC03-121E8.3	80	32	
	L1P4b	chr14:20267817-20269960	chr14:20268726-20268856	OR4N1P	OR4N2	OR4K6P	73	1	
	L1PA11	chr11:2279544-2279610	chr11:2278774-2283048	MIR4686		ASCL2	72	16	
	L1PA16	chr17:37905713-37905797	chr17:37903730-37907972	GRB7		RP11-94L15.2	85	38	
	L1PBa1	chr4:146953901-146954697	chr4:146954414-146954532	Y_RNA		RP11-203B7.2	60	8	

	HUM							Mensch	Schimpanse
Neuronale Zellen	L1PBa1	chr8:2584506-2585024	chr8:2584733-2585625	RP11-378A12.1	RP11-134O21.1;GS1-57L11.1	C5MD1	34	95	
	L1MC3	chr4:188953092-188953245	chr4:188953133-188953241	TRIML2		ZFP42	64	10	
	L1MC4	chr8:144788059-144788457	chr8:144787532-144788522	CCDC166		RP11-429J17.2	81	4	
HUM	L1MB5	chr16:1160593-1160798;chr16:1160900-1160999	chr16:1158680-1162581	LA16c-349E11.1	LA16c-381G6.1	CACNA1H	27	84	
	L1ME2z	chr17:1032842-1032929;chr17:1033240-1033407	chr17:1029918-1033700	MIR3183		AC016292.1	27	62	
	L1PA8	chr17:1032409-1032518;chr17:1032825-1032837	chr17:1029918-1033700	MIR3183	ABR	AC016292.1	27	62	
Nicht-Neuronale Zellen	L1M1	chrX:153390703-153392264	chrX:153391129-153391259	MECP2		OPN1LW	45	8	
	L1M5	chr5:1228149-1228320	chr5:1228061-1228444	SLC6A19		TERT	71	29	
		chr8:1882726-1883281	chr8:1882222-1885011	MIR596	ARHGEF10	KBTBD11-OT1	89	11	
	L1MA10	chr20:60543018-60543194;chr20:60543305-60543411	chr20:60541116-60545298	MIR1257	TAF4	RP11-11M20.2	60	27	
		chr9:13677655-136776706	chr9:136776366-136780525	AL590710.1	VAV2	LINC00094	68	16	
	L1MA2	chr9:135388415-135388600	chr9:135385719-135389883	RNU5D-2P	C9orf171	BARHL1	77	18	
	L1MA9	chr16:34297502-34297545	chr16:34295965-34300712	CTD-2144E22.6	CCNYL3	CTD-2144E22.2	63	6	
	L1MB3	chr13:114802752-114803175	chr13:114802848-114802961	LINC00565	RASA3	RASA3-IT1	65	1	
		chr17:77241511-77241625;chr17:77241626-77241710	chr17:77239436-77242699	RP11-398J5.1	RBFOX3	MIR4739	60	12	
	L1MB4	chr16:34296723-34297502;chr16:34297594-34298124;chr16:34298421-34298941	chr16:34295965-34300712	CTD-2144E22.6	CCNYL3	CTD-2144E22.2	63	6	
		chr16:88539071-88539149	chr16:88535740-88539534	MIR5189	ZFPM1	RP11-21B21.4	79	36	
		chr20:61627161-61627271	chr20:6162674761631444	SLC17A9		BHLHE23	61	21	
	L1MB5	chr9:136779317-136779633	chr9:136776366-136780525	AL590710.1	VAV2	LINC00094	68	16	
	L1MB7	chr16:1075961-1076317	chr16:1075594-1076037	RP11-161M6.3		SSTR5-AS1	64	16	
		chr8:142162340-142163148	chr8:142159766-142164167	RP11-809O17.1	DENND3	SLC45A4	74	29	
	L1MB8	chr16:88617633-88617919	chr16:88615054-88618763	ZFPM1		ZC3H18	74	28	
	L1MC1	chr8:1883277-1883653	chr8:1882222-1885011	MIR596	ARHGEF10	KBTBD11-OT1	89	11	
	L1MC3	chr17:77241497-77241510	chr17:77239436-77242699	RP11-398J5.1	RBFOX3	MIR4739	60	12	
	L1MC4	chr8:144788059-144788457	chr8:144787532-14789414	RP11-429J17.2	ZNF707;CCDC166	RP11-429J17.4	71	8	
		chrX:152069993-152070097	chrX:152066920-152070848	NSDHL		ZNF185	54	19	
	L1MC5	chr12:1957365-1957928	chr12:1957458-1957596	LRTM2	CACNA2D4	RP5-1096D14.3	71	28	
		chr21:46233356-46233470	chr21:46232703-46237148	AL773604.8	SUMO3	PTTG1P	70	19	
	L1MD2	chr2:240617499-240618316	chr2:240617264-240619544	ACO79612.2		RP11-315F22.1	87	25	
		chr9:138558869-138559062	chr9:138553850-138559100	GLT6D1	LCN9	SOHLH1	78	15	
	L1MD3	chr14:101944628-101944919	chr14:101944555-101944656	RP11-1029J19.2		DIO3O5	65	13	
	L1ME1	chr20:61626910-61627082	chr20:61626747-61631444	SLC17A9		BHLHE23	61	21	
	L1ME2z	chr9:135388638-135388715	chr9:135385719-135389883	RNU5D-2P	C9orf171	BARHL1	77	18	
	L1ME3	chr20:60545138-60545654	chr20:60541116-60545298	MIR1257	TAF4	RP11-11M20.2	60	27	
	L1ME3C	chr9:135388766-135388865	chr9:135385719-135389883	RNU5D-2P	C9orf171	BARHL1	77	18	
	L1ME4a	chr12:6496319-6496549	chr12:6494060-6497866	SCNN1A	LTBR	RP1-102E24.8	53	8	
	chr20:36049445-36049515	chr20:3604592536049848	SRC		RPL7A.P14	69	15		
L1MEg	chr12:133028903-133029317	chr12:133026845-133030351	RP13-895J2.7		RP11-503G7.1	88	30		
	chr2:127896371-127896573;chr2:127896844-127897268;chr2:127897461-127898235	chr2:127896177-127899176	BIN1		NIFKP9	87	15		
	chr22:50319982-50320141	chr22:50319331-50320783	ALG12	CRELD2	PIM3	70	11		
L1P3	chr2:127896705-127896829	chr2:127896177-127899176	BIN1		NIFKP9	87	15		
L1PA11	chr9:136777042-136777214	chr9:136776366-136780525	AL590710.1	VAV2	LINC00094	68	16		
L1PA15-16	chr8:120779447-120780204	chr8:120779789-120779900	RN7SKP153	TAF2	DSCC1	58	13		
L1PREC2	chr3:140737039-140737833	chr3:140736740-140738237	SLC25A36		SPSB4	70	14		