

**Auswirkungen
ausgewählter Einflussfaktoren
auf die genomweite DNA-Methylierung
beim Menschen**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Jana Gutwein

Kiel, 2015

Erste Gutachterin:	Prof. Dr. rer. nat. Manuela Dittmar
Zweiter Gutachter:	Prof. Dr. rer. nat. Ole Ammerpohl
Tag der mündlichen Prüfung:	26.03.2015
Zum Druck genehmigt:	26.03.2015

gez. Prof. Dr. rer. nat. Axel Scheidig, Prodekan

Inhaltsverzeichnis

Inhaltsverzeichnis

Abkürzungen	I
1 Einleitung	1
1.1 Genregulation und Epigenetik.....	1
1.2 Epigenetische Mechanismen.....	1
1.2.1 Histonmodifikationen	2
1.2.2 RNA-basierte Mechanismen	3
1.2.3 DNA-Methylierung.....	4
1.2.4 DNA-Hydroxymethylierung.....	10
1.3 Epigenetische Veränderungen aufgrund von Umweltfaktoren	11
1.3.1 Ernährung und Adipositas.....	11
1.3.2 Zigarettenkonsum	13
1.3.3 Der Alterungsprozess.....	13
1.3.4 Schichtarbeit und zirkadianer Rhythmus	15
1.4 Das diskordante Zwillingsmodell zur Untersuchung von Umwelteinflüssen	16
1.5 Zielsetzung.....	18
2 Material.....	20
2.1 Probenmaterial.....	20
2.2 Kits.....	20
2.3 Chemikalien.....	21
2.4 Verbrauchsmaterialien.....	21
2.5 Instrumente und Geräte.....	21
2.6 Software und Datenbanken	22
2.7 Primersequenzen für die Bisulfit-Pyrosequenzierung.....	23
3 Studienkollektiv und Methoden	24
3.1 Studienkollektiv.....	24
3.2 Bisulfitkonvertierung genomischer DNA.....	27
3.3 Infinium HumanMethylation450 BeadChip	28

Inhaltsverzeichnis

3.3.1	Methodischer Ablauf	31
3.3.2	Datenvorbereitung für Auswertung der Methylierungsdaten	37
3.3.2.1	Normalisierung mit GenomeStudio	37
3.3.2.2	Ausschluss von Loci mit nicht quantifizierbaren Signalintensitäten.....	38
3.3.2.3	Reproduzierbarkeit der hybridisierten DNA-Proben	38
3.3.2.4	Hintergrundkorrektur und Normalisierung mit RnBeads	39
3.3.2.5	Ausschlussbedingungen für die finale Analyse.....	40
3.3.3	Auswertung der Infinium HumanMethylation450 BeadChip-Daten	40
3.3.3.1	Identifizierung differentiell methylierter Genorte über die <i>False Discovery Rate</i> (FDR)	40
3.3.3.2	Identifizierung differentiell methylierter Genorte über p-Wert und $\Delta\beta$	41
3.3.3.3	Identifizierung differentiell methylierter Genorte im diskordanten Zwillingsmodell	41
3.3.4	Bestimmung des epigenetischen Alters nach Steve Horvath	42
3.3.5	Biologische Charakteristika differentiell methylierter Genorte.....	42
3.3.5.1	Genontologie-Analyse.....	42
3.3.5.2	Analyse von angereicherten Signalwegen	43
3.3.5.3	Weitere Anreicherungsanalysen.....	43
3.4	Bisulfit-Pyrosequenzierung	44
3.4.1	Methodischer Ablauf der Bisulfit-Pyrosequenzierung.....	45
3.5	Statistische Verfahren	51
4	Ergebnisse	53
4.1	Etablierung des Infinium HumanMethylation450 BeadChips.....	53
4.2	Anwendung des Infinium HumanMethylation450 BeadChips zur Analyse ausgewählter Einflussfakoren	55
4.2.1	Qualitätskontrolle der Methylierungswerte	55
4.2.1.1	Ausschluss von Loci mit nicht quantifizierbaren Signalintensitäten.....	55
4.2.1.2	Reproduzierbarkeit der hybridisierten DNA-Proben	56
4.2.1.3	Ausschlussbedingungen für die finale Analyse.....	56

Inhaltsverzeichnis

4.2.1.4	Identifizierung technischer Artefakte mittels Korrelationsanalyse des genomweiten Methylierungsmusters.....	56
4.2.1.5	Identitätsüberprüfung mittels SNP-Analyse	58
4.2.1.6	Identifizierung potentieller <i>Batch</i> -Effekte mittels <i>Principal Component Analyse</i> (PCA)	59
4.2.1.7	Zusammenfassung des Qualitätskontrollprozesses.....	61
4.2.2	Normalisierung der Signalintensitäten mittels RnBeads	62
4.2.3	Auswirkung des chronologischen Alters auf die genomweite DNA-Methylierung	63
4.2.3.1	Identifizierung altersabhängig methylierter CpG-Loci über die FDR.....	64
4.2.3.2	Eigenschaften der altersabhängig differentiell methylierten CpG-Regionen.....	68
4.2.3.3	GO-Analyse der 63 altersabhängig methylierten Genorte	69
4.2.3.4	Erweiterte Analyse der 63 altersabhängig methylierten Genorte mittels WebGestalt	71
4.2.4	Auswirkung des <i>Body Mass Index</i> (BMI) auf die genomweite DNA-Methylierung	72
4.2.4.1	Identifizierung der zwischen dem adipösen und normalgewichtigen Phänotyp differentiell methylierter Genorte über die FDR	73
4.2.4.2	Identifizierung differentiell methylierter Genorte über p-Wert und $\Delta\beta$	76
4.2.4.3	Eigenschaften der 80 zwischen den adipösen und normalgewichtigen Zwillingen differentiell methylierten CpG-Regionen	78
4.2.4.4	GO-Analyse der 43 mit hohem BMI korrelierenden differentiell methylierten Genorte.....	80
4.2.4.5	Erweiterte Analyse der zwischen normalgewichtigen und fettleibigen Zwillingen differentiell methylierten Genorte (n=43) mittels WebGestalt	81
4.2.5	Auswirkung des Rauchens auf die genomweite DNA-Methylierung.....	81
4.2.5.1	Identifizierung der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte über die FDR	81
4.2.5.1.1	Eigenschaften der 22 zwischen Rauchern und Nichtrauchern differentiell methylierten CpG-Regionen.....	84
4.2.5.1.2	GO-Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte (n=14)	85
4.2.5.1.3	Erweiterte Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte (n=14) mittels WebGestalt	87
4.2.5.2	Identifizierung differentiell methylierter Genorte zwischen Rauchern und Nichtrauchern mittels p-Wert und $\Delta\beta$	87

Inhaltsverzeichnis

4.2.5.2.1	Eigenschaften der zwischen Raucher und Nichtraucher differentiell methylierten CpG-Regionen (n=131)	90
4.2.5.2.2	GO-Analyse der zwischen Raucher und Nichtraucher differentiell methylierten Genorte (n=68)	91
4.2.5.2.3	Erweiterte Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte (n=78) mittels WebGestalt	91
4.2.5.2.4	Identifizierung von krebisrelevanten Genen innerhalb der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene	92
4.2.5.3	Prüfung der mit dem Rauchen assoziierten über die FDR ermittelten 22 CpG-Signatur in Lungen-Proben	92
4.2.6	Zwillingsstudie: Auswirkung von Schichtarbeit mit Nachtschichtdienst auf die genomweite DNA-Methylierung	95
4.2.6.1	Identifizierung differentiell methylierter Loci zwischen Schichtarbeitern mit Nachtschicht und Tagarbeitern	95
4.2.7	Zwillingsstudie: Auswirkung von Nachtarbeit auf die genomweite DNA-Methylierung	98
4.2.7.1	Identifizierung differentiell methylierter Loci zwischen Nacht- und Tagarbeitern	98
4.2.7.2	Eigenschaften der zwischen Nacht- und Tagarbeitern differentiell methylierten CpG-Regionen	99
4.2.7.3	GO-Analyse der zwischen Nacht- und Tagarbeitern differentiell methylierten Genorte	99
4.2.7.4	Erweiterte Analyse der 11 zwischen Nacht- und Tagarbeitern differentiell methylierten Genorte mittels WebGestalt	101
4.2.7.5	Validierung der Nachtarbeit-Signatur für ausgewählte Genorte	103
4.2.8	Zusammenfassung der DNA-Methylierungsanalysen	106
5	Diskussion	108
5.1	Validierung des Infinium HumanMethylation450 BeadChip und der Bisulfit-Pyrosequenzierung	108
5.2	Altersassoziierte Veränderungen des DNA-Methylierungsmusters	110
5.2.1	Eigenschaften der neu identifizierten mit zunehmendem Alter schwächer methylierten Genorte	112
5.2.2	Eigenschaften der neu identifizierten mit zunehmendem Alter stärker methylierten Genorte	113
5.2.3	Signifikante Anreicherungen altersassoziiierter Gene in biologischen Prozessen	114

Inhaltsverzeichnis

5.2.3.1	Anreicherung altersassoziierter Gene in Prozessen des Blutdrucks und der Herzkontraktion	115
5.2.3.2	Anreicherung altersassoziierter Gene in Prozessen des Nervensystems	115
5.2.4	Stärken und Limitationen der Alters-Analyse	115
5.2.5	Verifizierung der Alterssignatur basierend auf der Arbeit von Steve Horvath	117
5.2.6	Fazit der altersassozierten DNA-Methylierungsanalyse	118
5.3	Mit hohem BMI korrelierende Veränderungen des genomweiten DNA-Methylierungsmusters	118
5.3.1	Hoher BMI ist mit differentieller Methylierung von Loci in Genen, die in Prozessen des Nervensystems involviert sind, assoziiert	120
5.3.2	Identifizierte mit hohem BMI korrelierende differentiell methylierte Gene mit bekannter Assoziation zu Adipositas	120
5.3.3	Identifizierte mit hohem BMI korrelierende differentiell methylierte Gene mit Assoziation zu Typ-2-Diabetes (T2D)	123
5.3.4	Limitationen der BMI-Analyse	124
5.3.5	Fazit der DNA-Methylierungsanalyse für hohen BMI	125
5.4	Veränderungen des genomweiten DNA-Methylierungsmusters aufgrund des Zigarettenrauchens	126
5.4.1	Eigenschaften der über die FDR neu identifizierten zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte	126
5.4.2	Eigenschaften der aus dem zweiten Ansatz neu identifizierten zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte	128
5.4.3	Differentiell methylierte mit dem Rauchen assoziierte Gene sind im kardiovaskulären System involviert	132
5.4.3.1	Anreicherungen der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene in Herzerkrankungen	132
5.4.3.2	Anreicherung der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene im Renin-Angiotensin-Signalweg.....	133
5.4.3.3	Anreicherung der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene im Krankheitsbild der Thrombose	135
5.4.4	Prüfung der über die FDR bei Rauchern identifizierten 22 CpG-Methylierungssignatur in Lungen-Proben	136
5.4.4.1	Differentielle Methylierung in <i>AHRR</i> (<i>aryl hydrocarbon receptor (AHR) repressor</i>)	137
5.4.4.2	Differentielle Methylierung in <i>F10</i> (<i>coagulation factor X</i>).....	140

Inhaltsverzeichnis

5.4.5	Stärke und Limitation des Vergleichs des genomweiten DNA-Methylierungsmusters von Rauchern und Nichtrauchern	141
5.4.6	Fazit der genomweiten DNA-Methylierungsanalyse von Rauchern und Nichtrauchern	142
5.5	Nacharbeit und der Einfluss auf das DNA-Methylom	142
5.5.1	Limitationen der Nacharbeits-Methylierungsanalyse	146
5.5.2	Fazit der zwillingsbasierten Analyse zur Untersuchung des Einflusses von Nacharbeit auf die genomweite DNA-Methylierung.....	147
5.6	Synthese der Einzelbefunde und Ausblick der verschiedenen DNA-Methylierungsanalysen	148
6	Zusammenfassung	151
7	Summary	153
8	Literaturverzeichnis	155
9	Anhang.....	177
10	Danksagung.....	196
11	Publikationsliste	197
12	Eidesstattliche Erklärung.....	201

Abkürzungen

Abkürzungen

5caC	5-Carboxylcytosin
5fC	5-Formylcytosin
5hmC	5-Hydroxymethylcytosin
5mC	5-Methylcytosin
A (Nukleotid)	Adenin
A ^{vy}	<i>Agouti Viable Yellow</i>
APS	Adenosin-Phosphosulfat
ATP	Adenosin-Triphosphat
bp	<i>base pair</i> (Basenpaar)
BMI	<i>Body Mass Index</i>
BPA	Bisphenol A
BSPS	Bisulfit-Pyrosequenzierung
C (Nukleotid)	Cytosin
CNG	nicht-CpG-Stelle
CpG	Cytosin-(Phosphat-)Guanin-Dinukleotid
DHS	DNase I-hypersensitive Stelle
DMR	differenziell methylierte Region
DNA	Desoxyribonukleinsäure
DNMT	DNA-Methyltransferase
DNP	Dinitrophenyl
dNTP	Desoxyribonukleosidtriphosphat
DZ	dizygoter Zwilling
ER	endoplasmatisches Retikulum
FDR	<i>False Discovery Rate</i>
GO	Genontologie
H3K9	Histons 3, Lysinrest 9
HM450K BC	Infinium HumanMethylation450 BeadChip

Abkürzungen

IPD	Intra-Paar-Differenz
kb	<i>kilo base pair</i> (Kilo-Basenpaar)
MAF	<i>Minor Allele Frequency</i> (minimale Allelfrequenz)
MBD	Methyl-CpG-bindende Domäne
M-Bead	für das methylierte Allel korrespondierender Bead-Typ
miRNA	mikro-RNA
mRNA	<i>messenger RNA</i> (Boten-RNA)
MZ	monozygoter Zwilling
MZP	monozygotes Zwillingspaar
NaOH	Natronlauge
ncRNA	<i>non-coding RNA</i> (nicht kodierende RNA)
OR	<i>Odds Ratio</i>
PCA	<i>Principal Component Analysis</i> (Hauptkomponentenanalyse)
PCR	<i>Polymerase Chain Reaction</i> (Polymerasekettenreaktion)
PPi	Pyrophosphat
pRNA	Promotor-assoziierte RNA
RNA	Ribonukleinsäure
RNAi	RNA-Interferenz
SAM	S-Adenosyl-Methionin
SNP	<i>Single Nucleotide Polymorphism</i> (Einzelnukleotid-Polymorphismus)
SWAN	<i>Subset-Quantile Within Array Normalisation</i>
T (Nukleotid)	Thymin
T2D	Typ-2-Diabetes
TBE	Tris-Borat-EDTA
TF	Transkriptionsfaktor
TRDMT	tRNA-Asparaginsäure-Methyltransferase
tRNA	Transfer-RNA
TSG	Tumorsuppressorgen
TSS	Transkriptionsstartpunkt

Abkürzungen

U (Nukleotid)	Uracil
U-Bead	für das unmethylierte Allel korrespondierender Bead-Typ
UTR	untranslatierte Region
WHO	<i>World Health Organization</i> (Weltgesundheitsorganisation)

1 Einleitung

1.1 Genregulation und Epigenetik

Der menschliche Organismus besteht aus ca. zehn Billionen (10^{13}) Zellen [1], die mit einigen Ausnahmen (z.B. haploide Keimzellen, kernlose Erythrozyten sowie spezialisierte Immunzellen) alle das gleiche Erbgut, den gleichen Genotyp, besitzen, sich aber aufgrund unterschiedlicher Genaktivität zu über 220 verschiedenen Zelltypen [2] differenzieren, die sich sowohl in Funktion, Eigenschaft als auch Struktur unterscheiden. Gewährleistet wird der spezifische Phänotyp durch verschiedene Mechanismen der Genregulation, welche die Aktivität der Gene steuert, indem sie Zeitpunkt und Ausmaß der Genexpression bestimmt und somit für das einzigartige Genexpressionsmuster einer jeden Zelle verantwortlich ist. Während Kontrollmechanismen der Genregulation beispielsweise auf transkriptionaler Ebene in Form von Transkriptionsfaktoren (TFs), auf posttranskriptionaler Ebene durch RNA-Prozessierung oder auf posttranslationaler Ebene durch Modifikation von Proteinen eingreifen können, ist z.B. für die Embryonalentwicklung und Zelldifferenzierung die Regulation der Genexpression durch epigenetische Modifikationen essentiell. Die Epigenetik wird als Prozess der mitotisch und meiotisch vererbaren Veränderungen in der Genexpression definiert, die nicht auf Veränderungen in der DNA-Sequenz beruhen [3] und zugleich stabil, aber potentiell reversibel sind [4]. Dabei wird dem Organismus eine flexible Anpassung an veränderte Umweltbedingungen ermöglicht. Durch die Gesamtheit der epigenetischen Regulationsmechanismen wird das Epigenom jeder Zelle charakterisiert. Somit hat jeder Zelltyp und möglicherweise auch jede Zelle ihr eigenes einzigartiges Epigenom [5].

1.2 Epigenetische Mechanismen

Zu den wichtigen Genregulationsmechanismen der Epigenetik zählen die posttranslationalen Modifikationen von Histonen, RNA-basierte Mechanismen sowie die postreplikative Methylierung der DNA (siehe Abbildung 1). Weiterhin sind *nuclear positioning*, die Positionierung von Genen im Zellkern [6] sowie der Einbau spezifischer Histonvarianten [7] und die Änderung der Position von Nukleosomen auf der DNA [8] als epigenetische Mechanismen bekannt.

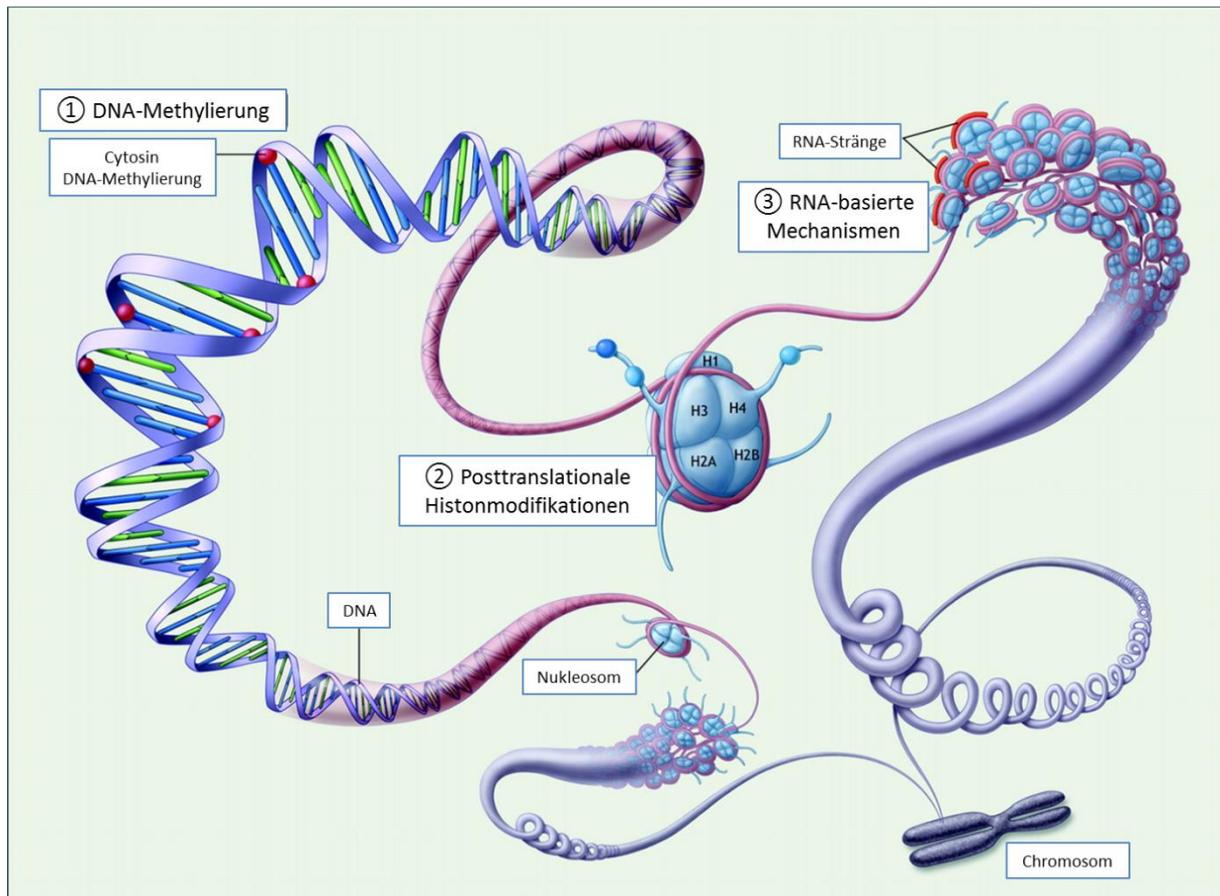


Abbildung 1: Epigenetische Mechanismen der Genregulation. Zu den wichtigen epigenetischen Mechanismen in der Kontrolle der Genexpression gehören drei zueinander in Wechselbeziehung stehende Mechanismen: DNA Methylierung, posttranslationale Histonmodifikationen und RNA-basierte Mechanismen. **(1)** DNA Methylierung (dargestellt als rote Bälle) beinhaltet die kovalente Modifikation des Cytosins innerhalb von CpG-Dinukleotiden unter der Bildung von 5-Methylcytosin. **(2)** Verschiedene posttranslationale Modifikationen der terminalen Histonschwänze (dargestellt als hell- und dunkelblaue Bälle) können die physikalischen Eigenschaften und die hohe Verdichtung des Chromatins entscheidend beeinflussen. **(3)** RNA-basierte Mechanismen sind wichtige Regulatoren der Chromatinstruktur und Genexpression (dargestellt als rote Stränge, die das Chromatin ummanteln). Abbildung verändert nach [9].

Da der Mechanismus der DNA-Methylierung wesentlicher Bestandteil der vorliegenden Arbeit ist, wird zunächst ein kurzer Überblick über die posttranslationalen Histonmodifikationen und RNA-basierten Mechanismen gegeben und anschließend die DNA-Methylierung ausführlich betrachtet.

1.2.1 Histonmodifikationen

Histone sind als Bestandteil des Chromatins für die kompakte Verpackung der DNA im Zellkern eukaryotischer Zellen mitverantwortlich und bestimmen durch die Veränderung der Chromatinstruktur infolge verschiedener Histonmodifikationen die Zugänglichkeit von bestimmten DNA-Sequenzen für die Genexpression [10,11]. Basierend auf dem Grad der

Kondensation des Chromatins, kann dieses in zwei verschiedene Typen eingeteilt werden: das dicht gepackte Heterochromatin, welches durch eine „geschlossene“ Chromatinstruktur mit einer transkriptionellen Inaktivität der Chromatinregion assoziiert ist und das locker gepackte Euchromatin, welches durch eine „offene“ Chromatinstruktur entsprechend mit einer transkriptionellen Aktivität einhergeht [12]. Jeweils zwei Kopien der vier basischen Kernhistone (H2A, H2B, H3 und H4) bilden das Histon-Oktamer, um welches ca. 146–147bp DNA gewunden sind [13,14]. Der Komplex aus DNA und Histonproteinen bildet das Nukleosom. Jedes Kernhistonprotein besitzt eine globuläre Domäne und einen langen N-terminalen Polypeptidschwanz, der aus dem Nukleosom herausragt und Angriffspunkt von einer Vielzahl an verschiedenen Modifikationen ist [11,15]. Zu den reversiblen Histonmodifikationen gehören beispielsweise die Azetylierung, Methylierung, Phosphorylierung, Ubiquitinierung, Sumoylierung [16] und Biotinylierung [17] der Aminosäurereste. Einige dieser Modifikationen, z.B. die von Histonazetylasen vermittelte Azetylierung von Lysinresten haben das Potenzial direkt die Chromatinstruktur zu beeinflussen und somit die Genexpression zu fördern. Durch die Azetylierung wird die positive Ladung der Histonproteine neutralisiert, woraufhin die Interaktion mit der negativ geladenen DNA abgeschwächt wird [18–21]. Dies führt folglich zu einer lockeren Chromatinstruktur, die somit für die Transkriptionsmaschinerie zugänglich ist. Analog führt eine Deazetylierung von Lysinresten zu einer kompakteren Chromatinstruktur und korreliert im Allgemeinen mit einer transkriptionellen Repression. Es konnte jedoch in Hefezellen gezeigt werden, dass auch die Deazetylierung eine transkriptionsaktivierende Funktion ausübt [22]. Die Methylierung von Lysinresten wird sowohl mit transkriptioneller Aktivierung als auch Repression assoziiert, was von der Position des methylierten Lysins am Histonschwanz und der Anzahl der hinzugefügten Methylgruppen abhängig ist [23]. So ist eine Methylierung des Histons 3 an den Lysinresten 9 und 27 (H3K9, H3K27) mit einer transkriptionellen Repression, an den Lysinresten 4, 36 und 79 (H3K4, H3K36, H3K79) hingegen mit einer Aktivierung assoziiert [24].

1.2.2 RNA-basierte Mechanismen

Es gibt zahlreiche bekannte Formen der RNA-vermittelten Genregulation. Hierzu zählen z.B. nicht kodierende RNAs (ncRNAs) und die RNA-Interferenz (RNAi). Als ncRNAs werden RNA-Moleküle bezeichnet, die zwar durch Transkription entstehen, aber nicht in Proteine

translatiert werden. Zu ihnen zählen beispielsweise die ribosomalen und Transfer-RNAs, die an der mRNA-Translation beteiligt sind sowie die kleinen nukleären RNAs, die für das Spleißen der Introns aus der prä-mRNA verantwortlich sind [25]. Durch Interaktion mit DNA oder mRNA spielen ncRNAs bei der epigenetischen Regulation der Genexpression eine wesentliche Rolle. So können sich andere ncRNAs, beispielsweise pRNAs (Promotor-assoziierte RNAs), an komplementäre Abschnitte der DNA anlagern und dadurch DNA-Methyltransferasen (DNMTs) an diese regulatorischen DNA-Sequenzen dirigieren. Diese sorgen dort für eine gezielte RNA-abhängige DNA-Methylierung der Promotoren von bestimmten Genen, was in einer Geninaktivierung resultiert [26,27]. RNAi ist ein Mechanismus, bei dem kleine doppelsträngige RNAs für eine sequenzspezifische Regulation der Genexpression genutzt werden. Die ncRNAs binden dabei entweder an kodierende oder Promotor-Regionen eines komplementären mRNA-Transkripts, was schließlich durch spezifische Degradation zur Inhibierung der Genexpression führt [28,29].

1.2.3 DNA-Methylierung

Der Prozess der DNA-Methylierung kommt sowohl bei Prokaryoten als auch bei Eukaryoten vor [30] und bezeichnet die kovalente Bindung einer Methylgruppe an bestimmte Basen der DNA. Die DNA-Methylierung gehört zu den bestuntersuchten epigenetischen Modifikationen und wurde in Wirbeltieren bereits 1980 als Mechanismus zur Repression von Genen beschrieben [31].

Bei Säugern und anderen Wirbeltieren findet die Methylierung hauptsächlich am Kohlenstoffatom in Position 5 des Pyrimidinrings von Cytosin statt, überwiegend innerhalb von CpG-Dinukleotiden (CpGs) [32]. Die Bildung von 5-Methylcytosin wird durch die hochkonservierten DNMTs katalysiert, die hierfür als Methyl donor S-Adenosyl-Methionin (SAM) [33] verwenden (Abbildung 2).

Einleitung

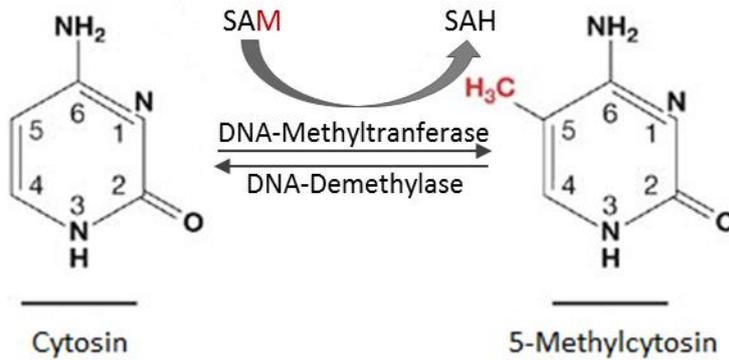


Abbildung 2: Die Methylierung von Cytosin zu 5-Methylcytosin. DNA Methyltransferasen katalysieren den Transfer der Methylgruppe von SAM auf die Position 5 des Cytosinrings in der DNA unter der Bildung von 5-Methylcytosin und SAH. Die Linien unter den Cytosin-Basen zeigen an, dass die Basen keine freien Moleküle sind; sie sind Bestandteile der DNA. SAM: S-Adenosyl-Methionin; SAH: S-Adenosyl-Homocystein. Abbildung verändert nach [34].

Für das globale Cytosin-Methylierungsmuster sind hierbei bei Säugern drei unabhängig voneinander kodierte DNMTs verantwortlich: DNMT1, DNMT3A und DNMT3B [35]. Es konnte eine weitere potenzielle DNMT (DNMT2) identifiziert werden, die jedoch den Transfer von Methylgruppen auf RNA statt DNA katalysiert [36] und daher kürzlich in tRNA-Asparaginsäure-Methyltransferase 1 (TRDMT1) umbenannt wurde [37]. Es können zwei Klassen von DNMTs unterschieden werden. Die *de novo* DNMTs sind für die Etablierung des Methylierungsmusters an spezifischen DNA-Sequenzen in der frühen Embryonalentwicklung verantwortlich. Sie werden entsprechend in frühen Embryonalzellen verstärkt exprimiert, während sie in ausgereiften somatischen Geweben nur schwach exprimiert vorliegen [37]. Die *de novo* Methylierung wird von der DNMT3-Familie bestehend aus zwei aktiven DNMTs (DNMT3A und DNMT3B) und dem DNMT3-ähnlichen Protein (DNMT3L), einem regulatorischen Faktor ohne Methyltransferase-Domäne [38], katalysiert. Hingegen ist die Erhaltungs-DNMT (DNMT1) vorwiegend für die Aufrechterhaltung des Methylierungsmusters während der Zellteilung durch gezielte Methylierung von hemimethylierten CpG-Dinukleotiden, die durch semikonservative DNA-Replikation generiert wurden, verantwortlich [39]. DNMT1 ist daher während der S-Phase des Zellzyklus in der Nähe der Replikationsgabel lokalisiert [40]. In nachstehender Abbildung ist ein Überblick über die Funktionen der DNMTs dargestellt.

Einleitung

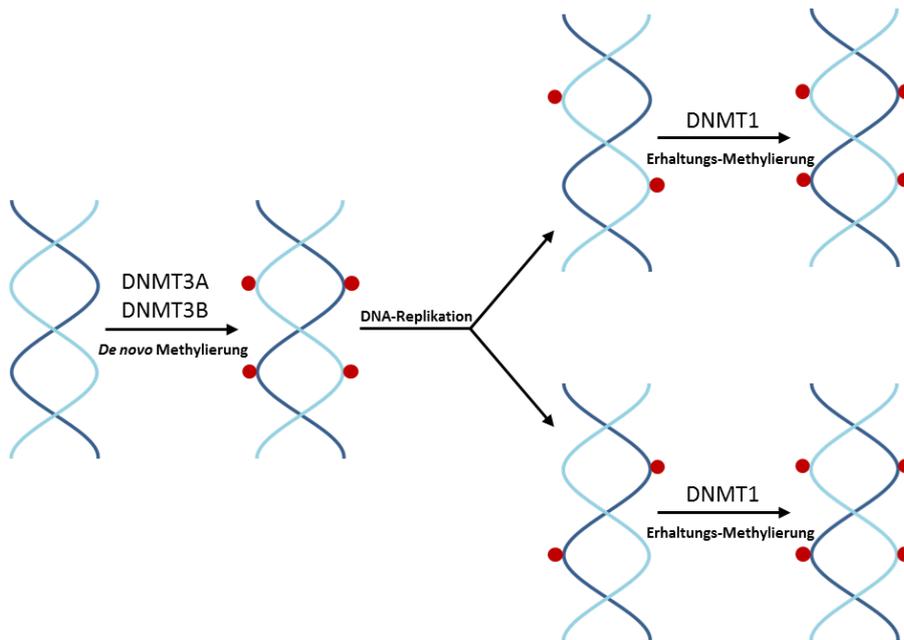


Abbildung 3: De novo Methylierung und Erhaltungs-Methylierung der DNA. Die unmethylierte DNA wird durch die DNA-Methyltransferasen DNMT3A und DNMT3B „de novo“ methyliert, was zu einem symmetrischen Methylierungsmuster an CpG-Dinukleotiden führt (Cytosin-Methylierung ist als roter Ball dargestellt). Wird die DNA repliziert (semikonservative Replikation), erkennt die DNMT1 die Methylgruppe an dem Elternstrang und methyliert daraufhin den komplementären Tochterstrang. Somit wird die hemimethylierte DNA postreplikativ wieder vollständig methyliert, wobei das ursprüngliche Methylierungsmuster aufrecht erhalten bleibt.

Neben dem dominanten CG-Methylierungsmuster ist aber auch die Methylierung von Nicht-CG-Sequenzmotiven wie CHG und CHH (H=A, C oder T) beschrieben worden. So wurde eine Nicht-CG-Methylierung z.B. in Pflanzen [41] und in Säugetieren, dort insbesondere in embryonalen Stammzellen [42,43], Oozyten von Mäusen [44] und heranreifenden menschlichen Neuronen [45] gefunden.

Im menschlichen Genom sind etwa 70–80% aller CpG-Dinukleotide methyliert [46,47], was einer Gesamtmethylierung von etwa 4% aller Cytosine im Genom entspricht [46]. Die Mehrheit der CpGs (ca. 98%) ist über das gesamte Genom ungleichmäßig verteilt, wo sie in differenzierten Zellen in intergenischen und intronischen Regionen meist methyliert vorzufinden sind [48]. Abzugrenzen von diesen CpG-armen Regionen sind Regionen, die eine große Anzahl an CpG-Wiederholungen beinhalten, die sogenannten CpG-Inseln (ca. 2%). Sie werden durch eine Mindestlänge von 500bp, einem CG-Gehalt von mindestens 55% und einem Quotienten von beobachteten zu statistisch erwarteten CpG-Dinukleotiden von mindestens 0,65 charakterisiert [49]. Es gibt schätzungsweise 50000 CpG-Inseln, von denen etwa 21000 in repetitiven Sequenzen lokalisiert sind [50].

Einleitung

Von den CpG-Inseln entfernt und funktionell mit diesen assoziiert liegen *Shores* (Regionen, die bis zu 2kb von einer CpG-Insel entfernt liegen) sowie *Shelves* (Regionen, die 2–4kb von einer CpG-Insel entfernt liegen) [51]. Weiter entfernte Bereiche werden als *Open Sea* bezeichnet [52].

CpG-Inseln liegen meist in Promotorbereichen aktiver Gene und sind dort in gesunden Zellen bis auf wenige Ausnahmen gewöhnlich nicht methyliert [53]. Darin eingeschlossen sind die Transkriptionsstartpunkte vieler Haushaltsgene, ein Teil der gewebespezifischen Gene und an der Entwicklung beteiligte regulatorische Gene [54]. Insgesamt können etwa 70% aller Promotoren mit einer CpG-Insel assoziiert werden [55].

CpG-Dinukleotide sind im Genom höherer Eukaryoten generell nur etwa ein fünftel so häufig anzutreffen als statistisch aus den Kombinationsmöglichkeiten der Basen zu erwarten wäre [49,50,56,57]. Dieser Unterrepräsentation liegt die spontane Desaminierung des 5-Methylcytosins zu Thymin zugrunde. Die daraus resultierende T-G-Fehlpaarung wird nur ineffizient repariert [58], woraufhin das mutierte Thymin durch DNA-Replikation als C/T-Transition an die Tochterzellen weitergegeben wird. Diese Art der Mutation kann bei einigen genetisch bedingten Erkrankungen, z.B. Hämophilie A, Protein-C-Defizienz oder Phenylketonurie beobachtet werden [59–62]. Nur in den CpG-Inseln wird die zu erwartende Häufigkeit der CpGs nahezu erreicht. Das liegt darin begründet, dass CpG-Inseln gewöhnlich nicht methyliert sind und somit nicht der spontanen Desaminierung unterliegen [49].

Die DNA-Methylierung reguliert viele biologische Schlüsselprozesse. Dabei erfüllt sie verschiedene Funktionen, die sich in Prokaryoten und Eukaryoten fundamental unterscheiden. So dient die DNA-Methylierung in Bakterien dem Schutz vor Phagen-DNA durch das bakterielle Modifikations-Restriktions-System, welches zwischen zelleigener und fremder DNA aufgrund des Methylierungsmusters unterscheiden kann [63,64]. Ferner spielt die DNA-Methylierung in der Fehlpaarungskorrektur bei der DNA-Synthese eine wichtige Rolle [64]. Die DNA-Methylierung kann bei Prokaryoten dabei sowohl an den Positionen 4 bzw. 5 des Cytosins als auch an der Position 6 des Adenins stattfinden [65].

In den meisten Eukaryoten dient die DNA-Methylierung hingegen hauptsächlich der Unterdrückung von Transposons und endogener Retroviren [66] sowie der Repression bestimmter Gene [3]. Die auf der Promotor-Methylierung basierende transkriptionelle Repression kann durch drei mögliche Mechanismen erklärt werden:

a) Das direkte Modell geht von einer sterischen Hinderung durch die Anwesenheit von Methylgruppen aus, indem diese die Bindung von bestimmten Transkriptionsfaktoren inhibieren können [67]. Einige Transkriptionsfaktoren, wie AP-2, CREB, E2F und NF- κ B, erkennen spezielle DNA-Sequenzen, die CpG-Dinukleotide beinhalten [67–70]. Werden diese CpGs in den DNA-Bindestellen methyliert, kommt es folglich zur Inhibierung der DNA-Protein-Interaktion. Hingegen wurden aber auch Transkriptionsfaktoren beschrieben, wie CTF, SP1 und YY1, die methylierungsunabhängig an ihre Erkennungssequenzen binden können, da sie gegenüber dieser Methylierung unempfindlich sind [71–73]. Demgegenüber steht die Mehrheit der TFs, die in der Sequenz ihrer Bindestellen keine CpG-Dinukleotide aufweisen und somit durch die Methylierung nicht beeinflusst werden können [74].

b) Das indirekte Modell besagt, dass die Methylgruppe selbst in der Lage ist, Methyl-CpG-bindende Proteine zu rekrutieren. Von ihnen sind bisher drei Proteine bekannt (MBD1, MBD2 und MeCP2), die eine homologe Methyl-CpG-bindende Domäne (MBD) besitzen, mit der sie an die methylierte DNA-Sequenz binden und wiederum Co-Repressoren und Histondeazetylasen rekrutieren können. Zusammen mit ihnen vermitteln sie die Umstrukturierung des Chromatins, was infolge der Chromatinkondensation zur Inhibierung der Transkription führt [75,76]. Zudem konnten auch einige Zinkfinger-Proteine, z.B. ZBTB4, ZBTB33 und ZBTB38 identifiziert werden, die zwar nicht aus der MBD-Familie stammen, aber ebenfalls an methylierte CpG-Dinukleotide binden und dabei als Transkriptionsrepressoren wirken [77].

c) Der dritte potentielle Mechanismus basiert auf der Tatsache, dass methylierte und unmethylierte DNA zu unterschiedlichen Chromatinstrukturen führen. Während unmethylierte DNA-Regionen eine offene Konformation des Chromatins bewirken, so dass diese für die Transkriptionsmaschinerie zugänglich ist und somit die Genexpression aktiviert werden kann, führt methylierte DNA zu einer verdichteten Chromatinstruktur, die Transkriptionsfaktoren daran hindern kann auf die DNA zuzugreifen [78,79].

Es ist dabei wahrscheinlich, dass die Modelle b) und c) in Wechselbeziehung stehen und interagieren [80].

Einleitung

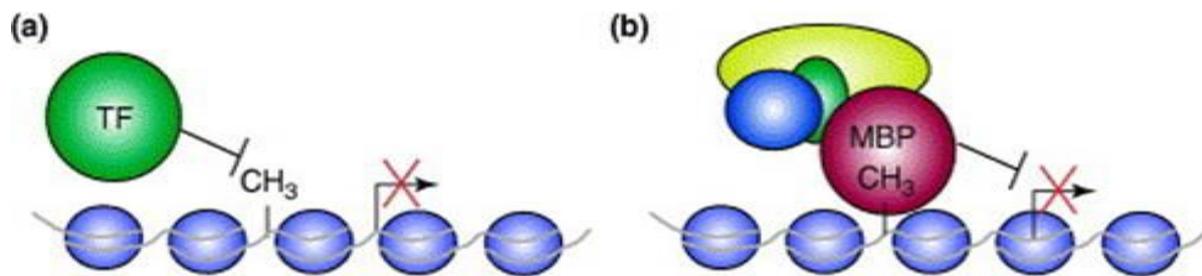


Abbildung 4: Der direkte und indirekte Mechanismus der durch DNA-Methylierung vermittelten transkriptionellen Repression eines Gens. a) Eine DNA-Methylierung (CH₃) in Bindungsstellen einiger Transkriptionsfaktoren (TF) blockiert diese für die Transkriptionsmaschinerie und verhindert somit die transkriptionelle Aktivierung eines Gens. **b)** Methyl-CpG-bindende Proteine (MBPs) erkennen methylierte DNA und rekrutieren Co-Repressoren, mit denen sie durch Modifikation des Chromatins die Transkription verhindern. Abbildung nach [30].

In Säugern befindet sich die große Mehrheit der CpG-Methylierung in parasitären Sequenzen [81]. Der Großteil dieser Methylierung ist in Transposons zu finden [82]. Transposons sind DNA-Sequenzen, die die Fähigkeit haben, sich innerhalb der Zelle an einer neuen Stelle im Genom zu integrieren und sind durch terminale invertierte Wiederholungen charakterisiert [83]. Sie machen dabei mehr als 40% des Säuger-genoms aus [84]. Viele dieser repetitiven Elemente werden durch Hypermethylierung der Promotoren in somatischen Zellen stillgelegt [82,85], wo sie im Laufe der Zeit durch Akkumulation der C/T-Transition, die in der spontanen Desaminierung methylierter Cytosine zu Thyminen begründet liegt, irreversibel inaktiviert werden [39]. Daher kann sich nur noch ein geringer Teil der Transposons (<0,05%) innerhalb des Genoms bewegen [84], wodurch die Integrität des Genoms gewahrt wird.

Ferner wird in Säugern eine Methylierung von CpG-Inseln in Promotoren genutzt, um die Transkription und somit die Expression des jeweiligen Zielgens zu inhibieren [86]. Klassisches Beispiel hierfür ist die X-Chromosom-Inaktivierung während der frühen Embryonalentwicklung bei Frauen, wobei Promotor-assoziierte CpG-Inseln methyliert werden und somit zur Geninaktivierung auf diesem Chromosom im Zuge der Dosiskompensation beitragen [87,88]. Als weiteres Beispiel sind die inaktiven Allele geprägter Gene zu nennen, die ebenfalls auf einer Methylierung der Promotoren dieser Gene beruhen, was zu einer allelspezifischen Genexpression paternaler oder maternaler Gene führt [89,90].

Während der Tumorentstehung und -progression sind viele CpG-Inseln der Hypermethylierung ausgesetzt, während die CpG-armen Regionen hypomethyliert werden [48]. Diese Änderung im DNA-Methylierungsmuster, wie die Promotor-Hypermethylierung bzw. die

Hypomethylierung von Nicht-Promotor-Regionen und strukturellen Elementen, führt zu Änderungen in der Chromatinstruktur, was zur Inaktivierung von Tumorsuppressorgenen (TSG) und zur Aktivierung schädlicher DNA-Elemente und somit zur Instabilität des Genoms führt [91].

1.2.4 DNA-Hydroxymethylierung

Während 5-Hydroxymethylcytosin (5hmC) erstmals in Bakteriophagen-DNA vor etwa 60 Jahren entdeckt werden konnte [92], wurden signifikante Mengen in der DNA von Säugern erst 2009 nachgewiesen [93,94]. Dabei konnte gezeigt werden, dass diese Base durch Oxidation von 5-Methylcytosin (5mC) durch Mitglieder der TET Enzym-Familie (TET1–3) katalysiert wird [93,95]. Es wird vermutet, dass 5hmC ein Zwischenprodukt des aktiven Demethylierungsprozesses von 5mC zu Cytosin ist [96]. Die Menge an 5hmC variiert in verschiedenen Gewebearten und Entwicklungsstadien stark, mit dem höchsten Vorkommen in Geweben des zentralen Nervensystems, insbesondere in Regionen mit kognitiven Funktionen (Großhirnrinde und Hippocampus) [94,97–99]. Dies lässt auf eine Rolle von 5hmC in neurobiologischen Funktionen schließen. Darüber hinaus wurde 5hmC auch in embryonalen Stammzellen detektiert [93,95,100], wo es die Aktivität von spezifischen regulatorischen Regionen reprimiert und daher als epigenetischer Marker für stillgelegte Enhancer vorgeschlagen wird [101].

Zudem sind noch zwei weitere Cytosin-Modifikationen bekannt, die ebenfalls von TET-Proteinen katalysiert werden und Teil des Demethylierungsprozesses sein könnten. Dabei wird 5hmC über 5-Formylcytosin (5fC) weiter zu 5-Carboxylcytosin (5caC) oxidiert. 5fC und 5caC könnten durch die Thymin-DNA-Glykosylase und die Basenextensionsreparatur oder 5caC durch Decarboxylierung wieder zu unmethyliertem Cytosin konvertiert werden [102,103]. Insgesamt konnte bereits gezeigt werden, dass die oxidierten Derivate von 5mC in bestimmten Zelltypen sowohl verschiedene Transkriptionsregulatoren als auch DNA-Reperatur-Proteine rekrutieren, die in der aktiven DNA-Methylierung eine Rolle spielen könnten [104].

1.3 Epigenetische Veränderungen aufgrund von Umweltfaktoren

Im Gegensatz zum stabilen Genom ist das Epigenom dynamisch, da es durch verschiedenste Umweltfaktoren, insbesondere während der fetalen und frühen postnatalen Entwicklung, beeinflusst werden kann. Wirken Umwelteinflüsse während der frühen Embryogenese auf den Organismus ein, können erhebliche, somatische Zellen betreffende Modifikationen der DNA-Methylierung verursacht werden, während sie im späteren Verlauf des Lebens weniger umfangreich und zudem gewebespezifische Modifikationen der DNA-Methylierung bewirken [105]. Die Epigenetik wird daher als Vermittler zwischen Genom und Umwelt angesehen [106]. Demnach kommt es in der Zelle zu epigenetischen Modifikationen, wenn eine Adaption an veränderte Umweltbedingungen erfolgt. Diese epigenetischen Veränderungen können nicht nur mitotisch an die Tochtergeneration, sondern teilweise auch über die Meiose an die nächste Generation als sogenannter Transgenerationseffekt vererbt werden [107,108].

Der Lebensstil ist ein wichtiger Parameter, der das Epigenom auf vielerlei Weise beeinflussen kann und beschreibt die Lebensweise, die Einstellungen und Werte einer Person oder Gruppe widerspiegelt (<http://www.thefreedictionary.com>). Diese schließt verschiedene Faktoren ein, wie Ernährung, Alkohol- und Zigaretten-Konsum sowie (psychische) Belastung, körperliche Aktivität und Arbeitsgewohnheiten [109].

1.3.1 Ernährung und Adipositas

Eine übermäßige Gewichtszunahme geht aus dem Zusammenspiel von Umweltfaktoren, genetischer Disposition und dem individuellen Lebensstil hervor [110]. Zu den verschiedenen Mechanismen, die zu interindividuellen Unterschieden in der Adipositas führen könnten, hat sich die epigenetische Regulation der Genexpression in den letzten Jahren als potentiell wichtiger mitwirkender Prozess herausgestellt. Daher können Veränderungen im DNA-Methylierungsmuster das Ergebnis des Zusammenspiels von verschiedenen Ernährungs- und Umweltfaktoren sowie die Ursache der interindividuellen Unterschiede sein, die die Anfälligkeit zur Ausbildung von Adipositas und anderer metabolischen Erkrankungen betreffen [110,111].

Veränderungen sowohl im qualitativen als auch quantitativen Nahrungsangebot können zu verschiedenen Modifikationen im Epigenom führen. Hinweise verdichten sich, dass die

Neigung zu Adipositas im Erwachsenenalter ihren Ursprung in der frühen Entwicklung innehaben könnte. Der Einfluss einer suboptimalen intrauterinen nutritiven Umgebung auf das Epigenom und den Phänotyp des Nachwuchses konnte in Tierversuchen bereits bestätigt werden [112]. So konnte eine Unterernährung in Ratten und Schafen während ihrer Trächtigkeit mit metabolischen Dysfunktionen wie Bluthochdruck, Insulinresistenz und Adipositas des Nachwuchses in Verbindung gebracht werden, welche auf epigenetischen Modifikationen beruhen [113,114]. Ebenso konnte der epigenetische Effekt eines qualitativen Nahrungsangebotes in Form einer methylreichen Ernährung auf die Nachkommen untersucht werden. Waterland und Jirtle zogen für ihr Experiment trächtige *Agouti Viable Yellow* (A^{vy})-Mäuse heran, bei denen im *Agouti*-Gen ein transposables Genelement inseriert wurde, welches u.a. die Ausprägung der Fellfarbe kontrolliert [115]. A^{vy} -Mäuse bilden neben der charakteristischen gelben Fellfarbe zudem einen adipösen und für Krebs und Diabetes anfälligen Phänotyp aus. Fütterten sie die trächtigen gelben A^{vy} -Mäuse zusätzlich mit Methylendonatoren wie Folsäure, Vitamin B12, Cholin und Betain, so zeigte der Nachwuchs wie bei normalen *Agouti*-Mäusen eine braune Fellfarbe sowie einen schlanken und gesunden Phänotyp. Für die normale Entwicklung der Jungtiere war eine verstärkte Methylierung des *Agouti*-Transposons verantwortlich, welches daraufhin inaktiviert und die A^{vy} -Expression somit während der Embryonalentwicklung reprimiert wurde.

Darüber hinaus konnte in epidemiologischen Studien auch beim Menschen gezeigt werden, dass ein suboptimales intrauterines quantitatives Nahrungsangebot, welches sowohl auf einen Überschuss als auch Mangel der maternalen Nahrungszufuhr beruhen kann, mit einem erhöhten Risiko für einige chronische Erkrankungen, wie Adipositas, Typ-2-Diabetes (T2D) und Erkrankungen des kardiovaskulären Systems im späteren Leben assoziiert waren [116,117]. Weiterhin konnten verschiedene Studien zeigen, dass die Mangelernährung infolge des Hungerwinters 1944/45 in den Niederlanden während Schwangerschaft und Kindheit lebenslange Auswirkungen auf die Gesundheit hatte und diese Effekte abhängig vom Zeitpunkt der Unterernährung variierten [118]. So konnte beispielsweise gezeigt werden, dass eine mütterliche Mangelernährung während des ersten Trimesters der Schwangerschaft mit Adipositas, koronarer Herzkrankheit und einer verminderten DNA-Methylierung des geprägten *insulin-like growth factor II*-Gens im Erwachsenenalter assoziiert war [119–121].

Bisphenol A (BPA) wird in der industriellen Produktion unter anderem als Weichmacher eingesetzt und stellt einen Umweltfaktor dar, der sich auf das Körpergewicht auswirken kann und ebenfalls mit Änderungen in der DNA-Methylierung interferiert. So konnte eine Zunahme im Körpergewicht des Nachwuchses von Mäusen und Ratten aufgrund einer *in utero* Exposition mit BPA beobachtet werden [122,123]. Ebenso konnte beim Menschen durch Kontakt mit BPA, das unter anderem in vielen Plastik-Trinkflaschen einschließlich Babyflaschen zu finden ist, ein erhöhtes Körpergewicht und Adipositas assoziiert werden, welches auf einer gesteigerten Adipogenese basierte [124]. Dabei kann BPA sowohl eine Hyper- als auch eine Hypomethylierung an CpG-Inseln in der Nähe von Gen-Promotor-Regionen induzieren [zusammengefasst in 125].

1.3.2 Zigarettenkonsum

Der Konsum von Zigaretten gilt als einer der schwerwiegendsten und zudem vermeidbaren Risikofaktoren für die Gesundheit und wird weltweit als wichtige Ursache für ein frühzeitiges Versterben angesehen. Dabei ist das Rauchen von Zigaretten mit einer großen Anzahl an Krankheiten assoziiert, einschließlich vieler Krebsentitäten des Atmungssystems und Magen-Darm-Trakts sowie verschiedenen Lungen- und Herzkreislauferkrankungen [126]. Das Rauchen von Zigaretten stellt sowohl bei Männern als auch bei Frauen die häufigste Ursache für alle histologischen Typen des Lungenkrebses dar [127,128]. Rauchen ist einer der mächtigsten Umweltfaktoren, der für Modifikationen im DNA-Methylierungsmuster verantwortlich gemacht wird [129–134]. Zudem konnten bei Neugeborenen Alterationen in der DNA-Methylierung bestimmter Genorte identifiziert werden, deren Mütter während ihrer Schwangerschaft rauchten [135–138]. Diesen epigenetischen Veränderungen könnten einige der bekannten Auswirkungen des aktiven mütterlichen Rauchens auf den Nachwuchs unterliegen, beispielsweise ein geringes Geburtsgewicht, eine eingeschränkte Lungenfunktion sowie frühe Atemwegserkrankungen [127].

1.3.3 Der Alterungsprozess

Das Altern ist ein Prozess, der durch einen zeitabhängigen globalen Rückgang der funktionalen Leistungsfähigkeit und Stressresistenz des Organismus charakterisiert werden kann und mit einem erhöhten Risiko für verschiedene Krankheiten wie Krebs und

letztendlich dem Tod assoziiert ist [139,140]. In diesem komplexen Prozess sind viele Faktoren auf molekularer Ebene involviert, die z.B. Telomerverkürzung, Akkumulation genetischer Mutationen, oxidativer Stress und molekulare Stoffwechselwege, die durch quantitative und qualitative Veränderungen in der Ernährung geändert werden, einschließen. Ebenso spielen epigenetische Modifikationen im Alterungsprozess eine wichtige Rolle. So könnten epigenetische Veränderungen während des Alterns potentiell zelluläre Funktionen beeinflussen und dadurch Auswirkungen auf die Entwicklung verschiedener altersassoziierter Phänotypen und/oder Krankheiten, wie neurologische Funktionsstörungen, Diabetes, Arthrose und Krebs, innehaben [141,142].

Es wird vermutet, dass epigenetische Faktoren, einschließlich der DNA-Methylierung, auch eine zentrale Rolle in kontrollierten Veränderungen in der Genexpression und genomischen Instabilität während des Alterungsprozesses innehaben könnten [143]. So nimmt mit dem Altern im Allgemeinen die globale DNA-Methylierung (z.B. von repetitiven und parasitären Sequenzen) ab, während die zelltyp-spezifische lokale Promotor-Methylierung (z.B. von TSG) zunimmt. Diese epigenetischen Veränderungen können ebenfalls in malignen Tumoren beobachtet werden [144], was in einem direkten Zusammenhang mit der Tatsache stehen könnte, dass fast alle Tumorentitäten bei älteren Menschen häufiger auftreten als bei jüngeren Menschen.

Die epigenetische Variation ist für Unterschiede in der Krankheitsanfälligkeit monozygoter Zwillinge mitverantwortlich und es konnte zudem gezeigt werden, dass junge monozygote Zwillingspaare sich epigenetisch ähnlicher sind als ältere monozygote Zwillingspaare [145]. Somit können sowohl der Alterungsprozess als auch verschiedene Umwelteinflüsse signifikante Änderungen in DNA-Methylierungsprofilen hervorrufen, die ferner mit fortschreitendem Alter akkumuliert werden können [146].

Es ist bekannt, dass die individuelle Wahl eines bestimmten Lebensstils, wie das Rauchen und die körperliche Aktivität, den Altersprozess beschleunigen oder verzögern kann [147,148].

Inzwischen konnte in diversen Studien ein Zusammenhang zwischen dem chronologischen Alter und dem DNA-Methylierungsmuster belegt werden, indem altersabhängige CpG-Loci identifiziert wurden, die entweder mit fortschreitendem Alter eine stärkere oder geringere Methylierung aufwiesen [141,146,149–154]. Einige CpG-Positionen zeigten dabei während

des Alterungsprozesses eine nahezu lineare Änderung in ihrem Methylierungsgrad und wurden daher als biologische Marker für verschiedene Modelle der Altersvorhersage verwendet [148,152,154–157]. So entwickelte Steve Horvath von der Universität in Kalifornien, Los Angeles, einen Algorithmus zur Bestimmung des epigenetischen Alters, wofür er 353 CpG-Stellen des Genoms nutzte [152]. Mit dieser Signatur konnte das epigenetische Alter in fast allen Geweben des Körpers bestimmt werden. Es korrelierte dabei signifikant mit dem chronologischen Alter und reflektierte nicht die Veränderungen in der Zelltyp-Zusammensetzung, sondern die für das Altern spezifischen Veränderungen im DNA-Methylom.

1.3.4 Schichtarbeit und zirkadianer Rhythmus

Zu den vom Menschen künstlich geschaffenen Umweltbedingungen zählt vor Allem die zivilisatorische Umwelt, also Umweltbedingungen, die auf zivilisatorischen Einflüssen und/oder Zwängen beruhen. Dazu zählt beispielsweise die Schichtarbeit, die im Laufe der Industrialisierung eine immer mehr verbreitete Arbeitsform darstellt. Nach den Leitlinien der Deutschen Gesellschaft für Arbeitsmedizin und Umweltmedizin ist Schichtarbeit als eine Form der Tätigkeit mit Arbeit zu wechselnden Zeiten (z.B. Früh-, Spät- und Nachtschicht) oder zu konstanter, aber ungewöhnlicher Tageszeit (z.B. permanente Spät- oder Nachtschicht) definiert [158]. Viele Dienstleistungen werden rund um die Uhr angeboten, darunter Sparten des Gesundheitswesens (Notfalldienste, Krankenschwestern), der Verkehrsbetriebe (Piloten, Taxifahrer, Zugführer), des Gastgewerbes (Hotel, Restaurant, Bar) sowie einige Verbrauchermärkte und Fastfoodketten. Heutzutage sind etwa 20% aller Arbeiter in Europa im Schichtdienst einschließlich Nachtarbeit tätig [159,160]. Für den Großteil der Schichtarbeiter kann diese Art der Arbeit dabei nicht frei gewählt werden, sondern stellt vielmehr eine berufliche Anforderung dar.

Schichtarbeit wird mit vielen gesundheitlichen Problemen assoziiert, z.B. Schlafstörungen, Übermüdungen, Stimmungsschwankungen, verringerter Leistungsfähigkeit, des metabolischen Syndroms, kardiovaskulären Krankheiten, Magen-Darm-Erkrankungen und Krebs [159,161–165]. Diese Probleme resultieren überwiegend aus der Desynchronisation der zirkadianen (~24-Stunden-) Rhythmik [162,165–167]. Der zugrunde liegende pathophysiologische Mechanismus ist dabei die Störung des normalen Schlaf-Wach-

Rhythmus durch Exposition von Licht und Dunkelheit zu unüblichen Zeiten [162]. Dadurch wird das zirkadiane System gestört bzw. verschoben, das u.a. Hormonproduktion, Körperkerntemperatur und Melatoninsekretion periodisch reguliert und überwiegend von Hell-Dunkel-Perioden sowie endogenen Zeitgebern beeinflusst wird [168,169]. An der Regulation der zirkadianen Rhythmen sind sogenannte Uhren-Gene und deren Genprodukte beteiligt, deren periodische transkriptionelle und translationale Aktivität dem 24-Stunden-Rhythmus folgt und durch epigenetische Mechanismen beeinflusst wird [170–172]. Studien konnten bereits demonstrieren, dass in Mäusen ein Leben gegen den inneren Rhythmus die epigenetisch gesteuerte rhythmische Aktivität des Fettstoffwechsels in der Leber störte, was in einer Fettleber resultierte [173,174]. Ferner konnte bereits ein Zusammenhang zwischen Schichtarbeit und aberranter Promotor-Methylierung von mikro-RNAs [175], zirkadianen und krebsrelevanten [176] und geprägten Genen [177] sowie eine aberrante Methylierung repetitiver Elemente [178] im Blut von Schichtarbeitern identifiziert werden.

1.4 Das diskordante Zwillingsmodell zur Untersuchung von Umwelteinflüssen

Man unterscheidet grundsätzlich zwei Formen von Zwillingen: monozygote Zwillinge (MZ) und dizygote Zwillinge (DZ). MZ entstehen, wenn eine einzige Eizelle von einem Spermium befruchtet wurde und sich die Zygote (befruchtete Eizelle) im frühen Entwicklungsstadium (bis zum 12. Tag) teilt. MZ besitzen daher theoretisch einen identischen Genotyp. Werden hingegen gleichzeitig zwei Eizellen von zwei verschiedenen Spermien befruchtet, entstehen DZ, die sich in ihrem Genotyp wie blutsverwandte Geschwister verhalten und sich somit 50% ihrer Allele teilen [179]. Die unterschiedlichen Genotypen beider Zwillingsformen macht man sich in Zwillingsstudien zu Nutze, um Eigenschaften oder Krankheiten dahingehend zu klassifizieren, ob sie auf genetische Faktoren oder Umwelteinflüsse zurückzuführen sind. Weisen MZ im Vergleich zu DZ in einem bestimmten Merkmal eine höhere Übereinstimmung (Konkordanz) auf, spricht dies für die Beteiligung einer genetischen Komponente. Verhält sich diese Konkordanz hingegen bei MZ und DZ ähnlich, so gibt dies Hinweis auf den Einfluss umweltbedingter Faktoren. So spielen beispielsweise bei Autismus mit einer Konkordanzrate von 92% bei MZ im Vergleich von 10% bei DZ genetische Faktoren eine wesentliche Rolle [180], während Typ-2-Diabetes mit einer Konkordanzrate von 50% bei MZ und 23% bei DZ [181] wesentlich in der individuellen Umwelt begründet zu liegen scheint. Das beste Beispiel

Einleitung

für den Einfluss von Umweltfaktoren auf das Epigenom liefern dabei MZ. Hierbei zeigen MZ, die theoretisch als „genetisch identisch“ gelten, Diskordanzen für einige Krankheiten, die sich in verschiedenen Phänotypen widerspiegeln (Abbildung 5).

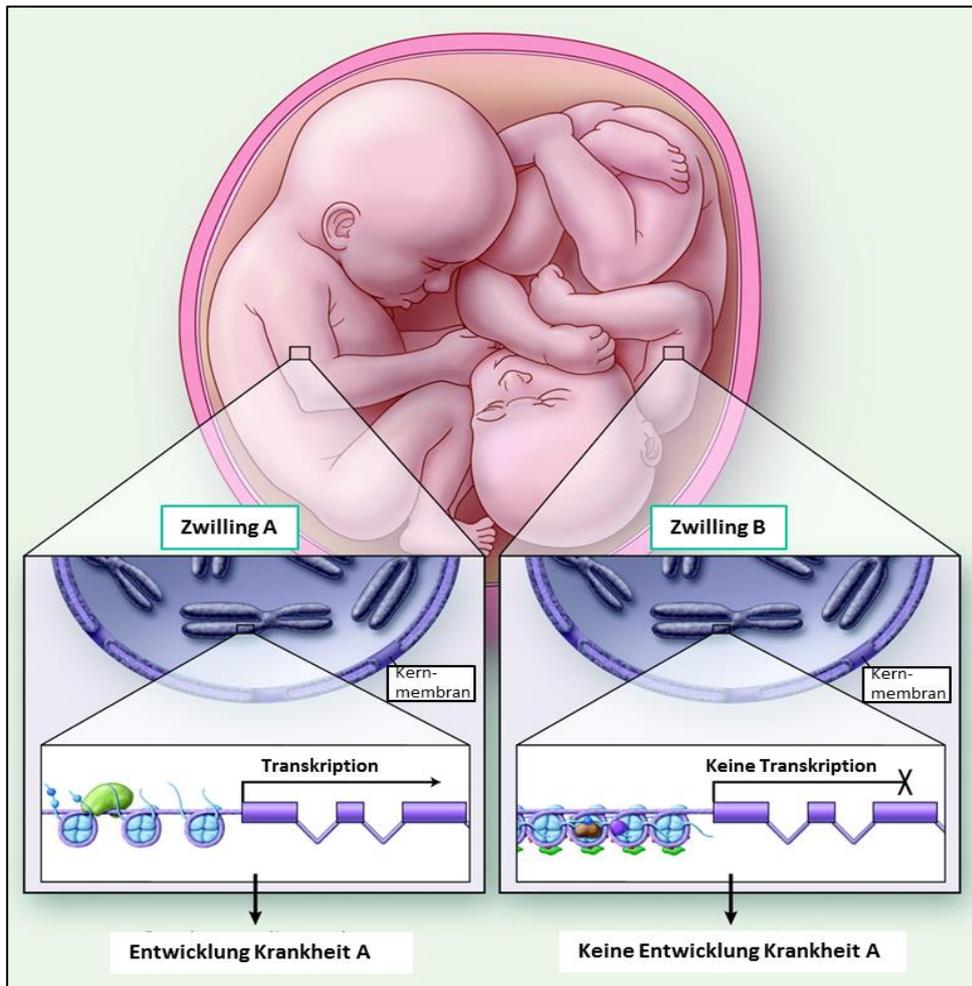


Abbildung 5: Eine epigenetische Erklärung für die Diskordanz von monozygoten Zwillingen (MZ) in gängigen Erkrankungen, die nicht dem Mendelschen Erbgang folgen, basiert auf unterschiedlichen Umwelteinflüssen.

In diesem Beispiel ist die 5'-regulatorische Region eines für eine bestimmte Krankheit anfälligen Gens in Endothelzellen des MZ A und MZ B dargestellt. In Zwilling A nimmt der Promotor eine "offene" Chromatinstruktur an. Diese ist durch aktivierende posttranslationale Histonmodifikationen, einer verringerten Nucleosomen-Dichte und das Fehlen einer DNA-Methylierung gekennzeichnet. Die RNA-Polymerase II kann zum Promotor rekrutiert werden und die Transkription initiieren. In Zwilling B nimmt der genetisch identische Promotor eine "geschlossene" Chromatinkonformation an, welche durch reprimierende posttranslationale Histonmodifikationen, eine erhöhte Nucleosomen-Dichte und einer DNA-Methylierung gekennzeichnet ist. Die RNA-Polymerase II kann nicht zum Promotor rekrutiert werden und folglich findet keine Transkription statt. Verändert nach [9].

Beiden Zwillingenformen sind sowohl das Alter als auch die intrauterinen und gewöhnlich auch die postnatalen Einflüsse gemeinsam. Durch den Vergleich des Grades der epigenetischen intra-Paar Zwillingen-Konkordanz von MZ- und DZ-Paaren ist es ferner möglich den Beitrag von erblichen und umweltbedingten Faktoren auf die Varianz der DNA-

Methylierung zu bestimmen. Während Fraga *et al.* bei MZ zeigen konnten, dass Umweltfaktoren das DNA-Methylierungsmuster möglicherweise beeinflussen [145], zeigten Kaminsky *et al.*, dass es signifikant mehr epigenetische intra-Paar Unterschiede bei DZ-Paaren gibt als bei MZ-Paaren [182]. Durch Mäuse-Studien konnten Kaminsky *et al.* die Hypothese, dass die größeren Unterschiede bei DZ-Paaren nicht allein durch Unterschiede in der DNA-Sequenz, sondern durch epigenetische Unterschiede in den Zygoten resultierten, bekräftigen.

MZ eignen sich demnach optimal, um die potentielle Rolle von Umweltfaktoren als bestimmende Faktoren von komplexen Krankheiten und Phänotypen zu untersuchen, da der genetische Hintergrund innerhalb der Paare identisch ist [183]. Hierfür wird das diskordante Zwillings-Design verwendet, in dem MZ, die für einen bestimmten Phänotyp eine Diskordanz zeigen, auf intra-Paar genomweite DNA-Methylierungsunterschiede analysiert werden. Da der Einfluss genetischer Faktoren auf das epigenetische Muster ausgeschlossen werden kann, müssen Unterschiede in der DNA-Methylierung somit in umweltbedingten Faktoren begründet liegen [184]. Mithilfe dieses Designs sollte in der vorliegenden Arbeit der Einfluss von Schichtarbeit auf das Methylohm von für Schichtarbeit diskordanten MZ-Paaren untersucht und in analog für Schichtarbeit diskordanten DZ-Paaren bestätigt werden.

1.5 Zielsetzung

Umweltfaktoren haben das Potenzial über epigenetische Mechanismen die Regulation der Genexpression zu beeinflussen und können somit zur Entwicklung oder dem Fortschreiten von Erkrankungen beitragen. Im Fokus der vorliegenden Arbeit war der epigenetische Prozess der DNA-Methylierung mit dem Ziel, umweltinduzierte Auswirkungen auf das genomweite DNA-Methylierungsprofil von monozygoten und dizygoten Zwillingen mittels arraybasierter Technologie zu identifizieren.

Für die durchzuführenden Methylierungsanalysen sollte ① der Infinium HumanMethylation 450K BeadChip (Illumina, San Diego, US) zunächst an einem Brustkrebskollektiv etabliert werden. Nach erfolgreicher Etablierung sollte ② der Array an einem gesunden Kontrollkollektiv bestehend aus monozygoten und dizygoten Zwillingen zur Analyse spezifischer Einflussfaktoren angewandt werden. Die individuellen Fragestellungen wurden dabei nach auf Fragebögen und anthropologischen Messungen basierender Verfügbarkeit

Einleitung

verschiedener Parameter generiert, so dass schließlich ③ die Untersuchung der epigenetischen Auswirkungen von chronologischem Alter und *Body Mass Index* (BMI) wie auch der Lebensstil- bzw. Umweltfaktoren Rauchen sowie Schichtarbeit mit Nachtdienst und Nachtarbeit im Vordergrund standen. Hierbei sollte deren Wirkung auf das globale DNA-Methylierungsmuster analysiert und in einen biologischen Kontext gebracht werden.

2 Material

2.1 Probenmaterial

Die Etablierung des Arrays erfolgte an Brustgewebe von gesunden und an Brustkrebs erkrankten Patientinnen (Institut für Pathologie, Medizinische Hochschule Hannover, Hannover, DE) sowie an den humanen Brustkrebs-Zelllinien HCC1937, MDA-MB-134 und PMC42 (ATCC, Wesel, DE) und IPH-926 (Institut für Pathologie, Medizinische Hochschule Hannover, Hannover, DE) [185].

Die Untersuchung der Einflussfaktoren Alter, *Body Mass Index*, Rauchen und Schicht- bzw. Nachtarbeit erfolgte an genomischer DNA von Zwillingen, die aus Zellen des peripheren Blutes extrahiert wurde. Die Blutentnahme sowie die DNA-Extraktion nach der Aussalzmethode [186] wurden von Kooperationspartnern (Institut für öffentliche Gesundheit, Süddänische Universität Odense, DK; Institut für regionale Gesundheitsforschung, Süddänische Universität Odense, DK) durchgeführt und in der angeforderten Gesamt-Konzentration auf Trockeneis versendet.

Die Prüfung der im Rahmen dieser Arbeit identifizierten mit dem Rauchen assoziierten Methylierungssignatur in einem Lungengewebekollektiv erfolgte an Lungen-Biopsien gesunder und an Lungenkrebs erkrankter Patienten (Forschungszentrum Borstel, Leibniz-Zentrum für Medizin und Biowissenschaften, Borstel, DE). Hierfür wurde auf die im Institut für Humangenetik bereits generierten Methylierungsdaten zugegriffen.

Eine Übersicht über die zur jeweiligen Fragestellung korrespondierende Anzahl der einbezogenen Proben wird in Kapitel 3.1 gegeben.

2.2 Kits

Bezeichnung	Hersteller/Lieferant
EpiTect Bisulfite Kit	Qiagen, Hilden, DE
EZ DNA Methylation™ Kit D5001/D5004	Zymo Research, Irvine, US
Infinium HumanMethylation450 BeadChip Kit	Illumina, San Diego, US
PyroMark PCR Kit	Qiagen, Hilden, DE

2.3 Chemikalien

Bezeichnung	Hersteller/Lieferant
10x Tris-Borat-EDTA (TBE)	Invitrogen, Darmstadt, DE
Biozym LE Agarose	Biozym Scientific GmbH, Hessisch Oldendorf, DE
Bromphenolblau	Serva, Heidelberg, DE
Desoxyribonukleosidtriphosphate (dNTPs)	Roche Diagnostics GmbH, Mannheim, DE
EDTA	Merck KGaA, Darmstadt, DE
Ethanol	Merck KGaA, Darmstadt, DE
Ethidiumbromid	Sigma-Aldrich-Chemie GmbH, Steinheim, DE
Formamid	Sigma-Aldrich-Chemie GmbH, Steinheim, DE
Isopropanol (2-Propanol)	Sigma-Aldrich-Chemie GmbH, Steinheim, DE
Methylierte DNA	Millipore, Darmstadt, DE
Natronlauge (NaOH); c(NaOH) = 1mol/l (1 N)	Merck KGaA, Darmstadt, DE
PyroMark Gold Q96 Reagents	Qiagen, Hilden, DE
PyroMark™ Annealing Buffer	Qiagen, Hilden, DE
PyroMark™ Binding Buffer	Qiagen, Hilden, DE
Streptavidin coated Sepharose Beads	GE Healthcare, München, DE
Tri Dye 100bp	New England Biolabs GmbH, Frankfurt am Main, DE

2.4 Verbrauchsmaterialien

Bezeichnung	Hersteller/Lieferant
0,2 ml bzw. 1,5 ml Reaktionsgefäße	Sarstedt AG&Co, Nürnberg, DE
0,8 ml 96-well storage plate	ThermoScientific, Erlangen, DE
Sealing mats round cap	ThermoScientific, Erlangen, DE
Solution Basin, 55-mL	Heathrow Scientific®LLC, Vernon Hills, US

2.5 Instrumente und Geräte

Bezeichnung	Hersteller/Lieferant
BioDoc Analyzer	Biometra, Göttingen, DE
DNA/RNA UV-Cleaner box UVC/T-M-AR	Kisker GbR, Steinfurt, DE

Material

Bezeichnung	Hersteller/Lieferant
Gelkammer	peqLab Biotechnologie, Erlangen, DE
Heizplatte Hotplate	VWR International GmbH, Darmstadt, DE
Hybex® Microsample Incubator	SciGene, Sunnyvale, CA, US
Hybridisierungssofen Hybridization Oven	Illumina, San Diego, US
iScan	Illumina, San Diego, US
Mikrowelle M500	Philips, Hamburg, DE
Plattenschüttler High-Speed Microplate Shaker	Illumina, San Diego, US
Plattenschüttler MS3 basic	IKA®Works, Inc., Wilmington, NC, US
Programmable temperature controller	VWR International GmbH, Darmstadt, DE
PyroMark Q96 Cartridge	Qiagen, Hilden, DE
PyroMark Q96 Vacuum Prep Trough	Qiagen, Hilden, DE
PyroMark Vacuum Prep Filter Probe	Qiagen, Hilden, DE
PyroMark™ Vacuum Prep Workstation	Qiagen, Hilden, DE
Pyrosequenzierer PyroMark Q96 ID	Qiagen, Hilden, DE
Rack Chamber Te-Flow	Tecan Schweiz AG, Männedorf, CH
Thermblock Thermomixer comfort	Eppendorf, Hamburg, DE
Thermocycler T3000	Biometra, Göttingen, DE
Vakuumpumpe Vacuum Pump	ThermoScientific, Erlangen, DE
Plattenabdichter Thermo sealer ALPS™25	ThermoScientific, Erlangen, DE
Plattenzentrifuge Megafuge 40R	Heraeus Kulzer GmbH, Dormagen, DE
Omega 869C Thermometer	Omega Engineering, Inc, US

2.6 Software und Datenbanken

Bezeichnung	Nutzung	Quelle/URL
Gene	Gen-Analyse	http://www.ncbi.nlm.nih.gov/gene/
GraphPad Prism 5.00	Statistik-Software	GraphPad Software, Inc., San Diego, US
Illumina® GenomeStudio V2011.1 Methylation Module v1.9	Auswertung des 450K BeadChips	Illumina, San Diego, US
iScan Control Software 3.3.28	Scannen des 450K BeadChips	Illumina, San Diego, US
Pubmed	Literatur-Recherche	http://www.ncbi.nlm.nih.gov/pubmed/

Material

Bezeichnung	Nutzung	Quelle/URL
PyroMark Assay Design Software (Version 2.0.1.15)	Primerdesign für PyroAssay	Qiagen, Hilden, DE
PyroMark ID (Version 1.0)	Erstellen der Pyrosequenzierungs-Assays	Qiagen, Hilden, DE
PyroMark® CpG Software (Version 1.0.11)	Lauf und Auswertung der Pyrosequenzierungs-Assays	Qiagen, Hilden, DE
Qlucore Omics Explorer 2.3	Auswertung des 450K BeadChips	Qlucore, Lund, SE
R (V.3.0.1–3.1.0)	Statistik-Software	http://www.r-project.org/
RnBeads (V.0.99.10)	Hintergrundkorrektur und Normalisierung der HM450K Methylierungsdaten	http://rnbeads.mpi-inf.mpg.de/
RStudio (Version 0.98.507)	Benutzeroberfläche für R	http://www.rstudio.com/
UCSC	Genomes, Blat, PCR	http://genome.ucsc.edu/
Venny	Darstellung von Venn-Diagrammen	http://bioinfogp.cnb.csic.es/tools/venny/
WebGestalt	Genontologie-Analyse	http://bioinfo.vanderbilt.edu/webgestalt/option.php

2.7 Primersequenzen für die Bisulfit-Pyrosequenzierung

Die in dieser Arbeit verwendeten Oligonukleotide wurden von der Firma biomers.net GmbH (Ulm, DE) bezogen und für die Bisulfit-Pyrosequenzierung zur Validierung der 450K-Methylierungsanalyse verwendet. Die Primersequenzen sind in Kapitel 3.4.1 in Tabelle 4 hinterlegt. Biotinylierte Oligonukleotide wurden in HPLC-gereinigter Qualität, die übrigen Oligonukleotide in Kartuschen-gereinigter Qualität eingesetzt.

3 Studienkollektiv und Methoden

3.1 Studienkollektiv

Das primäre Studienkollektiv der vorliegenden Arbeit wurde im Rahmen des von der Europäischen Union geförderten Interreg 4A Syddanmark-Schleswig-K.E.R.N. Projektes "Schlaf, Arbeit und deren Konsequenzen für menschliche Stoffwechselkrankheiten (SAME)" rekrutiert. Hierfür wurde auf Material und Daten des Dänischen Zwillingsregisters zurückgegriffen, das Daten von mehr als 88.000 Zwillingspaaren beinhaltet. Von 1512 gesunden Zwillingen, die zuvor an der GEMINAKAR-Zwillings-Studie teilgenommen hatten, stimmten insgesamt 1127 einer Teilnahme an der SAME-Studie zu. Insgesamt konnten 1119/1127 Zwillinge von den dänischen Kooperationspartnern besucht und medizinisch untersucht werden, wobei verschiedene Parameter, z.B. Blutdruck, Körperhöhe und Gewicht, erfasst wurden. Ebenso wurden für medizinische Analysen Blut-, Speichel- und Urin-Proben abgenommen. 1127 Zwillinge füllten spezifische Fragebögen aus, die Informationen z.B. über Gesundheit, Ernährung, Arbeitsmuster und Lebensstil beinhalteten.

Die Studie wurde von der entsprechenden dänischen Ethikkommission (GEMINAKAR, S-VF-19970271; SAME, S-20090065) und der dänischen Datenschutzbehörde (GEMINAKAR, 1999-1200-441; SAME, 2009-41-2990) positiv votiert. Von allen Studienteilnehmern lag eine schriftliche Einwilligungserklärung vor.

Da von der eigenen Arbeitsgruppe primär der Einfluss von Schichtarbeit auf das genomweite DNA-Methylierungsmuster untersucht werden sollte, wurde das Kollektiv spezifisch nach Zwillingspaaren gefiltert, die eine Diskordanz hinsichtlich ihrer Arbeitszeit aufwiesen sowie nach Zwillingen, deren Arbeitszeiten mit Nachtarbeit assoziiert waren. Gemäß dieser Fragestellung konnten in der Gesamtkohorte insgesamt 77/1127 Zwillinge im Alter von 30–63 Jahre identifiziert werden. Dabei handelte es sich um 37 Zwillingspaare sowie um drei einzelne Zwillinge ohne korrespondierenden Co-Zwilling (siehe Tabelle 1). Eine vollständige Tabelle mit allen verfügbaren Angaben der eingeschlossenen Probanden ist in Tabelle A1 im Anhang dargestellt. Über die von den Kooperationspartnern angegebene Zygote konnten die Zwillinge 18 monozygoten und 19 dizygoten Zwillingspaaren sowie drei dizygoten Zwillingen zugeordnet werden. Insgesamt wurden die DNA-Proben von 30 männlichen und 47 weiblichen Zwillingen untersucht.

Methoden

Tabelle 1: Studienkollektiv bestehend aus insgesamt 77 Zwillingen.

Nr	Zwillingspaar ID	Zygotie	Geschlecht Z1	Geschlecht Z2	Alter [Jahre]	GEMINAKAR	SAME
1	ZP11	MZ	M	M	58	ja	ja
2	ZP12	MZ	M	M	53	ja	ja
3	ZP32	MZ	M	M	46	ja	ja
4	ZP37	MZ	W	W	39	ja	ja
5	ZP10	MZ	W	W	55	ja	ja
6	ZP31	MZ	W	W	47	nein	ja
7	ZP34	MZ	W	W	33	ja	ja
8	ZP36	MZ	M	M	49	ja	ja
9	ZP17	MZ	W	W	48	ja	ja
10	ZP16	MZ	W	W	47	ja	ja
11	ZP20	MZ	W	W	38	ja	ja
12	ZP1	MZ	M	M	46	ja	ja
13	ZP9	MZ	M	M	40	ja	ja
14	ZP3	MZ	M	M	36	ja	ja
15	ZP21	MZ	W	W	57	ja	ja
16	ZP23	MZ	W	W	32	ja	ja
17	ZP39	MZ	W	W	52	nein	ja
18	ZP38	MZ	W	W	47	nein	ja
19	Z18	DZ	M	---	56	ja	ja
20	ZP6	DZ	M	M	56	ja	ja
21	ZP8	DZ	W	W	55	ja	ja
22	ZP35	DZ	M	M	52	ja	ja
23	Z4	DZ	M	---	44	ja	ja
24	ZP27	DZ	W	W	34	ja	ja
25	ZP28	DZ	W	W	63	ja	ja
26	ZP14	DZ	W	W	52	ja	ja
27	ZP13	DZ	M	M	50	ja	ja
28	ZP30	DZ	W	W	47	ja	ja
29	ZP2	DZ	W	W	41	ja	ja
30	ZP15	DZ	W	W	34	ja	ja
31	ZP5	DZ	W	W	40	ja	ja
32	ZP7	DZ	W	W	35	ja	ja
33	ZP24	DZ	M	M	31	ja	ja
34	ZP26	DZ	M	M	49	ja	ja
35	ZP25	DZ	W	W	42	nein	ja
36	ZP29	DZ	W	M	47	ja	ja

Methoden

Nr	Zwillingspaar ID	Zygotie	Geschlecht Z1	Geschlecht Z2	Alter [Jahre]	GEMINAKAR	SAME
37	ZP33	DZ	W	M	40	ja	ja
38	ZP19	DZ	W	M	39	ja	ja
39	ZP22	DZ	W	M	34	ja	ja
40	Z40	DZ	W	---	30	ja	ja

Z1: Zwilling 1. Z2: Zwilling 2. Alter [Jahre]: Alter in Jahren zum Zeitpunkt der Blutentnahme der SAME Studie. GEMINAKAR: Studie 1997–2000. SAME: Studie 2010–2013. ZP: Zwillingspaar. Z: Zwilling. MZ: monozygoter Zwilling. DZ: dizygoter Zwilling. M: männlich. W: weiblich.

Für 69/77 Zwillinge konnte aufgrund der vorausgegangenen Baselinestudie (GEMINAKAR) zusätzlich etwa 10 Jahre altes DNA-Material angefordert werden. Dies sollte dem Vergleich bestimmter Parameter, z.B. eines veränderten Rauchverhaltens oder des fortschreitenden Alters, dienen.

Nachfolgend ist die jeweilige Anzahl der einbezogenen Proben für die Fragestellungen ①: Etablierung des Infinium HumanMethylation450 BeadChips und ②: Anwendung des Arrays (siehe Tabelle 2) sowie ③ über die ausgewählten Einflussfaktoren (siehe Tabelle 3) zusammengefasst.

Tabelle 2: Übersicht der Probenanzahl und des entsprechenden Kollektivs für die jeweiligen Fragestellungen.

Fragestellung	Probenanzahl (n)	Kollektiv
① Etablierung Array	4	humane Brustkrebs-Zelllinien
	4	humanes Brustgewebe
	18	humanes Brustkrebsgewebe
② Anwendung Array	146	monozygote und dizygote Zwillinge (GEMINAKAR, SAME)

Tabelle 3: Übersicht der zu untersuchenden Einflussfaktoren mit Probenanzahl und der dazugehörigen Geschlechts- und Altersverteilung.

Ausgewählte Einflussfaktoren	Probenanzahl (n)	Geschlecht M/W	Altersbereich [Jahre]
Alter	116	52/64	20–63
BMI (adipös vs. normalgewichtig)	34 (5 vs. 29)	13/21	32–63
Rauchen (Raucher vs. Nichtraucher)	37 (12 vs. 25)	14/23	31–63
Schichtarbeit (Schicht- vs. Tagarbeiter)	28 (14 vs. 14)	14/14	31–63
Nachtarbeit (Nacht- vs. Tagarbeiter)	12 (6 vs. 6)	4/8	38–52

M: männlich. W: weiblich. BMI: *Body Mass Index*: gibt das Verhältnis von Körpergewicht zum Quadrat der Körperhöhe in kg/m² an. Die Zwillinge wurden gemäß der internationalen Klassifikation der Weltgesundheitsorganisation (WHO) in folgende BMI Gruppen eingeteilt: normalgewichtig (18,5–24,9 kg/m²),

übergewichtig (≥ 25 – $29,9 \text{ kg/m}^2$) und adipös ($\geq 30 \text{ kg/m}^2$). Die Klassifikation für Übergewicht und Adipositas ist geschlechtsunabhängig und gilt ferner für alle Altersklassen der Erwachsenen (<http://www.who.int/mediacentre/factsheets/fs311/en/>).

3.2 Bisulfitkonvertierung genomischer DNA

Damit die Information des DNA-Methylierungsmusters bei einer genomweiten Amplifikation (WGA) erhalten bleibt, wird die DNA im Vorfeld einer Bisulfitbehandlung unterzogen [187], wobei das Methylierungsmuster in eine Sequenzinformation überschrieben wird. Unmethylierte Cytosine werden dabei hydrolytisch zu Uracil desaminiert, während 5-Methylcytosine begünstigt durch eine kurze Reaktionsgeschwindigkeit unverändert bleiben. Die Bisulfitreaktion besteht aus insgesamt drei Schritten unter der Bildung von zwei Zwischenprodukten (siehe Abbildung 6). Im ersten Schritt erfolgt die Sulfonierung von Cytosin zu Cytosin-Sulfonat durch Hinzufügen von Bisulfit-Ionen (HSO_3^-) an die 5-6 Doppelbindung des Cytosins. Im nächsten Schritt wird das entstandene Cytosinsulfonat-Derivat hydrolytisch zu einem Uracilsulfonat-Derivat desaminiert bevor aus diesem im letzten Schritt durch Abspaltung der Sulfonyl-Gruppe infolge einer Alkalibehandlung Uracil gebildet wird.

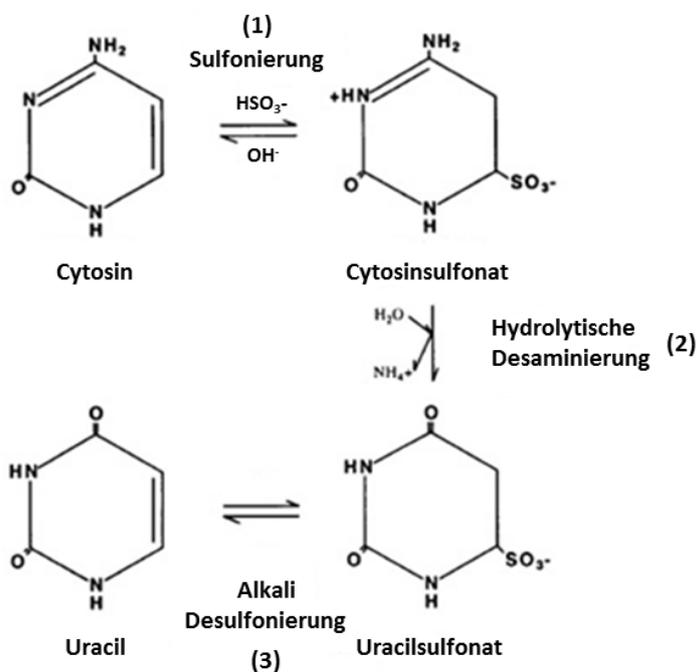


Abbildung 6: Chemischer Ablauf der Bisulfitkonvertierungsreaktion. Die Desaminierung von Cytosin durch Bisulfit beinhaltet folgende Schritte: Durch Addition von Bisulfit-Ionen an die 5-6 Doppelbindung der Cytosinbase wird Cytosin zu Cytosinsulfonat sulfoniert (1). Die hydrolytische Desaminierung des entstandenen Cytosinsulfonat-Derivats zu einem Uracilsulfonat-Derivat (2). Die Abspaltung der Sulfonyl-Gruppe durch Alkalibehandlung unter Bildung der Base Uracil (3). Verändert nach [187].

Methoden

In einer anschließenden PCR kommt es aufgrund des gleichen Basenpaarungsverhaltens zum Austausch von Uracil gegen Thymin, während methyliertes Cytosin als Cytosin-Base erhalten bleibt (siehe Abbildung 7). Ein Thymin entspricht demnach einem nicht methylierten Cytosin, ein Cytosin hingegen einem methylierten Cytosin. Somit kann der Methylierungszustand der jeweiligen Cytosine in einem nachfolgenden methylierungsspezifischen Assay, wie einer Pyrosequenzierung oder einem Methylierungsarray, analysiert werden.

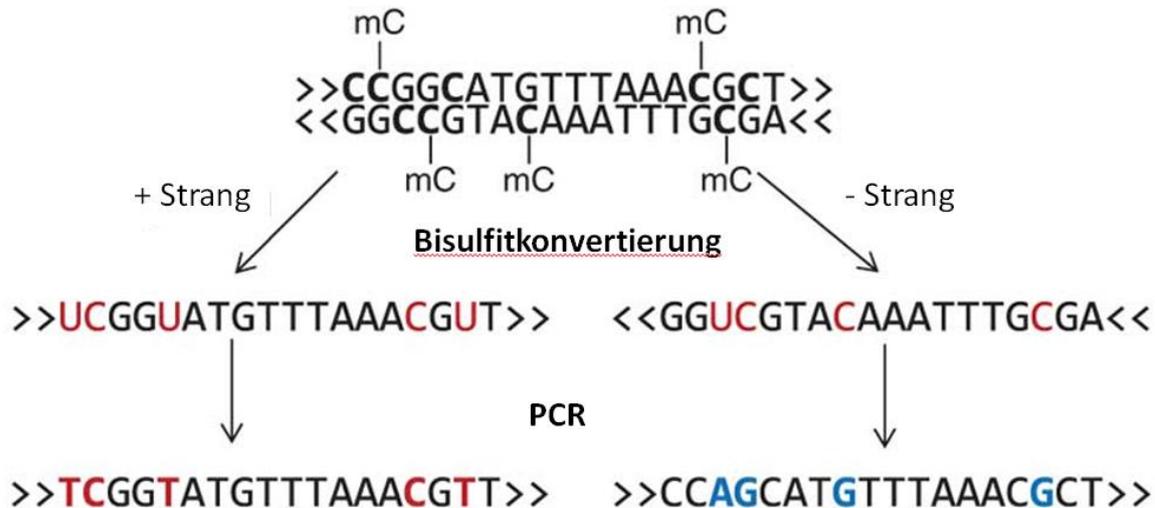


Abbildung 7: Bisulfitkonvertierung von genomischer DNA. Durch die Bisulfitbehandlung der DNA werden Cytosine zu Uracile konvertiert und in einer nachfolgenden PCR gegen Thymine ausgetauscht. Methylcytosine (mC) sind vor der Reaktion geschützt und bleiben somit erhalten. Abbildung verändert nach [188].

3.3 Infinium HumanMethylation450 BeadChip

Zur Untersuchung des globalen DNA-Methylierungsmusters wurde der Infinium HumanMethylation450 BeadChip (Illumina, San Diego, US), der im Folgenden als "HM450K BC" bezeichnet wird, verwendet. Dieser Array erlaubt die parallele Abfrage des Methylierungsstatus von >480000 Cytosin-Positionen, die in funktionellen Regionen von 21231 Genen liegen und somit etwa 99% der UCSC RefSeq-Gene des humanen Genoms (n=21474) ausmachen. Pro Genregion werden durchschnittlich etwa 17 CpG-Dinukleotid-Positionen (CpGs) abgedeckt. Diese verteilen sich über Regionen innerhalb von 200bp oder 1500bp vor dem Transkriptionsstartpunkt (TSS), 5'-untranslatierte Regionen (5'-UTRs), erste Exone, Genkörper (von TSS+500bp bis zum 3'-Ende des Transkriptes) und 3'-untranslatierte Bereiche (3'-UTRs). Des Weiteren sind auf dem Array CpG-Dinukleotide in Promotorregionen stärker abgedeckt als im Genkörper (siehe Abbildung 8).

Methoden

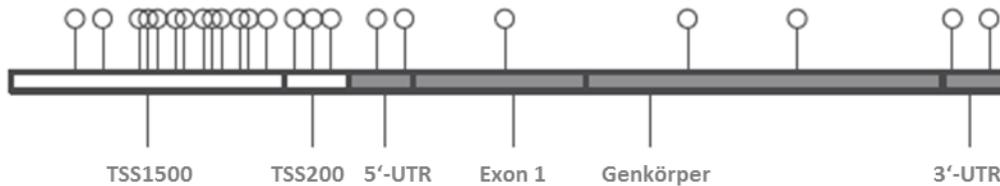


Abbildung 8: Beispielhafte CpG-Verteilung einer Genregion. Jedes Transkript ist unterteilt in sechs funktionelle Regionen. Die TSS1500 erstreckt sich von 200bp–1500bp vor dem Transkriptionsstartpunkt (TSS), die TSS200 liegt bis zu 200bp vor dem TSS, die 5'-UTR bezeichnet die Region vom TSS zum Translationsstartpunkt, Exon 1 bezeichnet das erste Exon eines Gens, der Genkörper ist definiert als das gesamte Gen von 500bp nach dem TSS bis zum 3'-Ende des Transkriptes, die 3'-UTR befindet sich hinter dem Translationsstopp-Codon. Ein Lollipop-Symbol stellt exemplarisch ein CpG-Dinukleotid dar, wobei die größte CpG-Dichte in der Promotorregion eines Gens anzutreffen ist. Verändert nach [189].

482421/485577 Positionen (99%) liegen innerhalb von CpGs und 3091 Positionen in nicht-CpG-Stellen (CNG), während 65 Positionen willkürliche Einzelnukleotid-Polymorphismen (SNPs) darstellen [51]. Ca. 31% der CpGs sind in CpG-Inseln lokalisiert, 23% in *Shores* (2kb Regionen, die die CpG-Inseln umgeben) und 10% in *Shelves* (2kb Regionen, die die *Shores* umgeben) (siehe Abbildung 9). Die restlichen 36% sind über das Genom verteilt als isolierte CpGs in der *Open Sea* vorzufinden [52]. Die Annotations-Terminologie wurde von Illumina übernommen und nicht eingedeutscht.

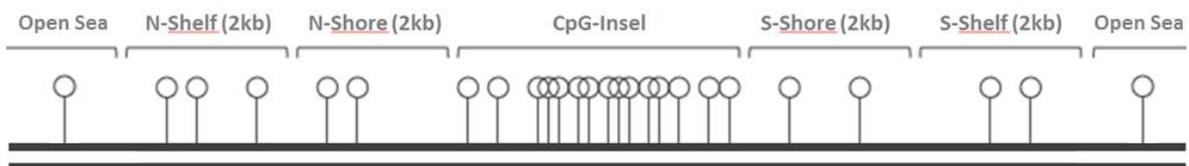


Abbildung 9: Unterteilung der CpG-Inseln und benachbarter Regionen. Shores (N-Shore und S-Shore) sind definiert als Bereiche, die innerhalb einer Entfernung von 2kb von CpG-Inseln liegen. Die Shores werden wiederum von entsprechenden Shelves (N-Shelf bzw. S-Shelf) umgeben, die ebenfalls eine Spanne von etwa 2kb aufweisen. Charakteristisch für CpG-Inseln ist der höhere CpG-Gehalt im Gegensatz zu den Shores und Shelves. Ein Lollipop-Symbol stellt exemplarisch ein CpG-Dinukleotid dar, wobei die größte Dichte in der CpG-Insel anzutreffen ist. Verändert [189].

Etwa 41% der CpGs (200339/482421) sind in proximalen Promotorbereichen lokalisiert, die innerhalb von 200bp oder 1500bp vor dem TSS, in der 5'-UTR sowie in Exon 1 liegen. Ca. 31% der CpGs liegen in Genkörpern, 25% in der sogenannten *Open Sea* und 3% in den 3'-UTRs.

Im Folgenden wird der theoretische Hintergrund zum Prinzip der auf dem HM450K BC verwendeten BeadChip Technologie erläutert.

Methoden

Die Oberfläche des Chips ist mit Beads beschichtet, die an ihrer spezifischen Adress-Sequenz jeweils eine Linker-DNA aus 50 Nukleotiden tragen, mit welcher die zu untersuchende DNA-Probe gebunden wird. Der HM450K BC kombiniert zwei technisch verschiedene Assay Designs, den Infinium I und den Infinium II Assay, die beide den C/T-Polymorphismus analysieren, der durch Bisulfitkonvertierung unmethylierter Cytosine entsteht [190]. Während etwa 28% der Cytosin-Positionen vom Infinium I Assay (n=135501 Sonden) detektiert werden, entfallen ca. 72% auf den Infinium II Assay (n=350076 Sonden). Der Infinium I Assay (siehe Abbildung 10A) kam schon auf dem Vorgängermodell des HM450K BC, dem HM27K BC, zur Anwendung. Dieses Verfahren unterscheidet zwei Bead-Typen. Für jeden CpG-Locus gibt es einen für das unmethylierte (U) und einen für das methylierte (M) Allel korrespondierenden Bead-Typ. Die U-Beads tragen an ihrer 3'-Position ein Adenin. Hybridisieren diese Sonden mit der entsprechenden Sequenz, die den unmethylierten Locus (=Thymin) beinhaltet, werden durch eine Polymerase markierte Nukleotide komplementär zur Ziel-Sequenz inkorporiert. Entsprechend ist nur ein methylierter CpG-Locus (=Cytosin) für die Extension an einem M-Bead verantwortlich, da dieser an seinem 3'-Ende ein Guanin trägt. Die enzymatische Einzelbasenverlängerung erfolgt somit allel-spezifisch. Unabhängig vom Methylierungszustand der CpG-Stelle wird im Falle einer Verlängerung bei beiden Bead-Typen das gleiche Nukleotid integriert, was somit zu einem gleichen Farbsignal führt. Aufgrund der jeweils spezifischen Adress-Sequenz des jeweiligen Beads, kann während des Array-Scans das Farbsignal dem entsprechenden Bead zugeordnet werden. Hingegen kommt beim Infinium II Assay (siehe Abbildung 10B) für jeden CpG-Locus nur ein Bead-Typ zum Einsatz. Hier wird die zu analysierende CpG-Stelle direkt eingebaut. Im Falle eines methylierten CpG-Locus wird die Base Guanin inkorporiert, während ein unmethylierter Locus den Einbau von Adenin bewirkt. Aufgrund der unterschiedlich markierten Nukleotide kann somit der Methylierungsstatus anhand des spezifischen Farbsignals ermittelt werden. Zur Berechnung des Methylierungsgrades der jeweiligen Cytosine wird das Verhältnis aus der Intensität des methylierten Allels zu der Summe der Intensitäten des methylierten und unmethylierten Allels herangezogen und als β -Wert definiert:

$$\beta = \frac{\text{Intensität des methylierten Allels}}{(\text{Intensitäten des methylierten + unmethylierten Allels} + 100)}$$

Der β -Wert liegt zwischen 0 und 1, wobei durch den Wert 0 keine Methylierung (0%) und durch den Wert 1 eine vollständige Methylierung (100%) angegeben wird.

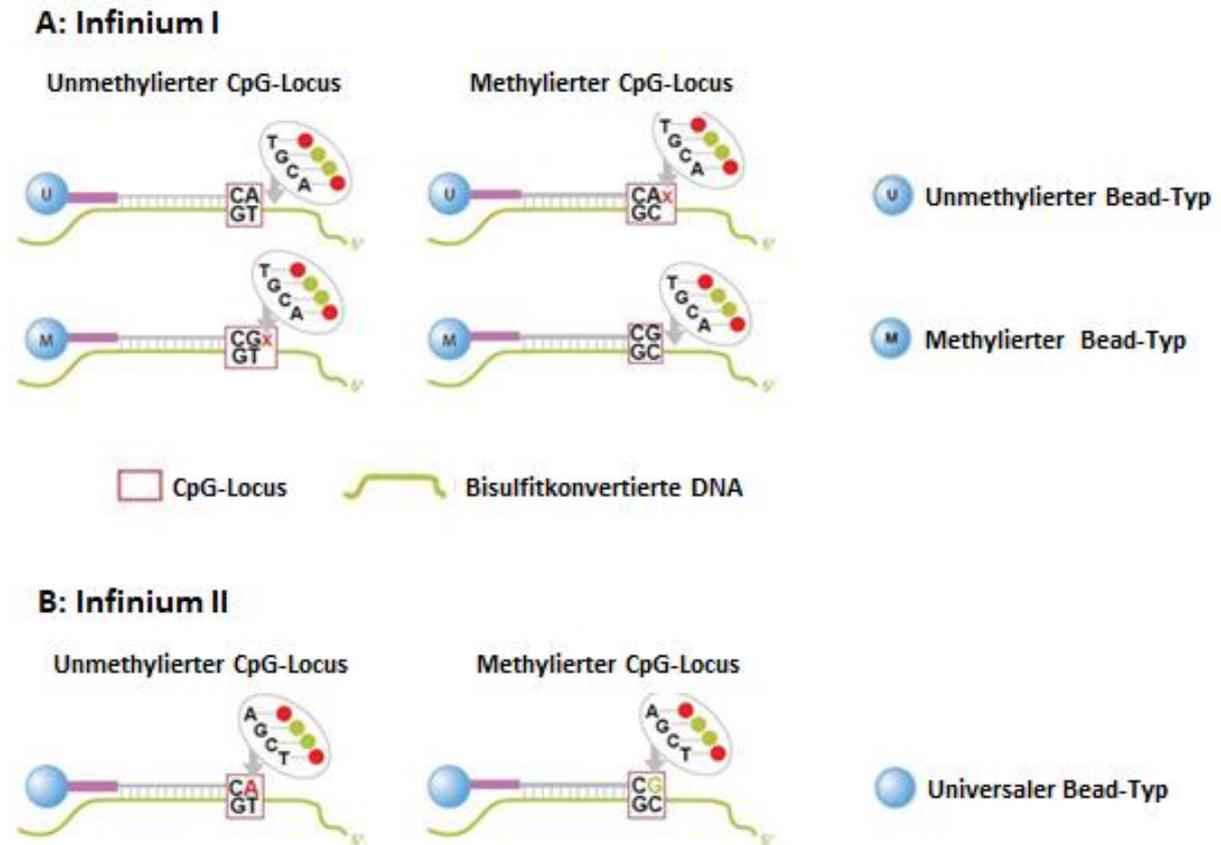


Abbildung 10: Unterschiede in den beiden Infinium-Technologien auf dem HM450K BC. **A:** Das Infinium I Assay Design beinhaltet pro CpG-Locus zwei Bead Typen. Bead Typ U korrespondiert für den unmethylierten Zustand (T) des abzufragenden CpG-Locus und Bead Typ M für den methylierten Zustand (C). **B:** Das Infinium II Assay Design beinhaltet einen universalen Bead Typ, der beide Zustände des Locus (unmethyliert und methyliert) detektiert. Verändert nach [187].

3.3.1 Methodischer Ablauf

Im Folgenden werden nun die Einzelschritte des methodischen Ablaufs für den HM450K BC näher beschrieben.

Bisulfitkonvertierung der genomischen DNA

Wie bereits unter 2.2.2. erwähnt, müssen die zu untersuchenden DNA-Proben zunächst einer Bisulfidbehandlung unterzogen werden, bevor mit dem eigentlichen Protokoll für den HM450K BC begonnen werden kann. Hierfür wurde bei weniger als 48 Proben das EZ DNA Methylation™ Kit D5001 (Zymo Research, Irvine, US) verwendet. Die einzelnen Zentrifugationsschritte wurden bei Raumtemperatur für 30 Sekunden bei 14000xg durchgeführt. Gemäß den Angaben des Herstellers wurde jeweils 1µg genomische DNA mit ddH₂O auf ein Volumen von 45µl gebracht. Nach Zugabe von 5µl M-Dilutions-Puffer wurde

Methoden

die DNA-Probe für 15 Minuten bei 37°C denaturiert, was Grundvoraussetzung für die nachfolgende einzelstrangspezifische Reaktion ist. Die Bisulfitreaktion erfolgte in einem Gesamtvolumen von 150µl (45µl DNA-Probe, 5µl M-Dilutions-Puffer und 100µl frisch vorbereitetes CT Konversions-Reagenz) in 200µl PCR-Reaktionsgefäßen durch Inkubation im Thermocycler unter folgenden Bedingungen:

Zyklus	Zeit [min]	Temperatur [°C]	Anzahl Zyklen
Denaturierung	0,5	95	16
Inkubation	60	50	
Lagerung	10	4	1

Es wurden Zymo-Spin™ IC Säulen in entsprechende Reaktionsgefäße überführt und mit 400µl M-Binde-Puffer befüllt. Nach Zugabe des inkubierten Bisulfitreaktionsgemisches wurden diese invertiert, zentrifugiert (14000xg, 30 Sekunden) und der Durchfluss verworfen. Es wurden weitere 100µl M-Wasch-Puffer auf die Säule gegeben und erneut zentrifugiert. Danach erfolgte die Desulfonierung der DNA, indem 200µl M-Desulfonierungs-Puffer zur Säule gegeben und für 15–20 Minuten bei Raumtemperatur inkubiert wurde. Anschließend wurde der Puffer abzentrifugiert. Zur Aufreinigung der DNA aus dem Reaktionsgemisch folgten zwei Waschschrte, indem jeweils zweimal mit 200µl M-Wasch-Puffer gewaschen und zentrifugiert wurde. Nach dem letzten Waschschrte wurde die Säule in ein 1,5ml Reaktionsgefäß überführt und die DNA mit 10µl ddH₂O durch Zentrifugation (14000xg, 30 Sekunden) eluiert. Anschließend wurde das Eluat in ein 200µl PCR-Reaktionsgefäß überführt und zur direkten Weiterbehandlung für den HM450K BC eingesetzt oder gegebenenfalls bis zur Nutzung bei -20°C aufbewahrt.

Bei mehr als 48 Proben gleichzeitig wurde das EZ DNA Methylation™ Kit D5004 (Zymo Research, Irvine, US) verwendet, wobei das Bisulfitreaktionsgemisch sowohl die gleiche Zusammensetzung als auch die gleichen Inkubationsbedingungen aufwies wie das des EZ DNA Methylation™ Kit D5001. Nach der Konvertierung wurde die Aufreinigung der Proben im 96er Plattenformat durchgeführt. Hierbei unterschied sich das entsprechende Protokoll in Zentrifugationsdauer und -geschwindigkeit sowie in den einzusetzenden Volumina. Statt den zuvor beschriebenen Zentrifugationsschritten bei 14000xg für 30 Sekunden, wurde bei der Aufreinigung der Proben stets bei 5000xg für 5 Minuten zentrifugiert. Weiterhin wurden die

Methoden

Proben vor und nach Desulfonierung mit jeweils 500µl statt 200µl Waschpuffer gewaschen. Das Elieren erfolgte außerdem in 15µl statt 10µl ddH₂O.

Die nachfolgenden Schritte wurden mit dem Infinium HumanMethylation450 BeadChip Kit (Illumina, San Diego, US) nach Angaben des Herstellers durchgeführt. Für die Durchführung des Protokolls wurden 4 Tage benötigt, wobei der erste Tag auf die bereits beschriebene Bisulfidkonvertierung genomischer DNA fiel. Eine Übersicht der für den HM450K BC weiteren relevanten Prozesse ist in Abbildung 11 dargestellt und wird im Folgenden näher beschrieben.

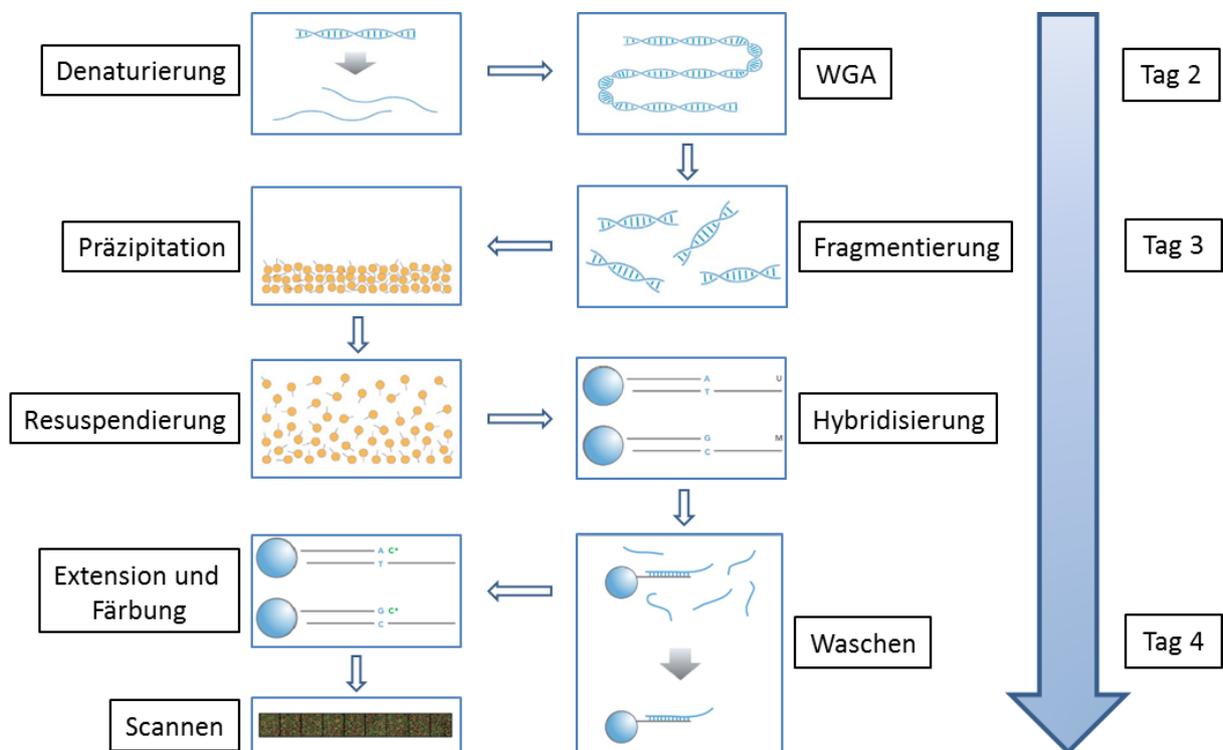


Abbildung 11: Übersicht der relevanten Arbeitsschritte mit Zeitverlauf für den HM450K BC. Nach Denaturierung der bisulfidkonvertierten DNA wird diese genomweit amplifiziert. Es folgen Schritte der Fragmentierung, Präzipitation und Resuspendierung, bevor die DNA-Proben auf den HM450K BC übertragen werden können. Nach Hybridisierung wird der Array gewaschen und die 50er-Nukleotide anschließend mit markierten Nukleotiden entsprechend verlängert und mittels Sandwich-Assay fluoreszenzmarkiert. Schließlich kann der Array gescannt werden. Abbildung adaptiert und verändert nach dem Infinium HD Methylation Assay Protokoll (Part # 15019522 Rev. B).

Genomweite Amplifikation der bisulfitkonvertierten DNA

Für die genomweite Amplifikation der bisulfitkonvertierten DNA wurden 4µl der konvertierten DNA in eine 96 Well Mikroplatte transferiert, in der bereits 20µl MA1 (*Multi-Sample Amplification 1 Mix*) vorgelegt waren. Zur Denaturierung der DNA-Proben wurden diese mit 4µl NaOH (0,1N) versehen. Die Platte wurde mit einer entsprechenden Matte abgedeckt, für 1 Minute bei 1600 rpm gevortext und nach 1-minütiger Zentrifugation bei 300xg für 10 Minuten bei Raumtemperatur inkubiert. Anschließend wurde zur Neutralisierung zu jeder Probe 68µl MA2 (*Multi-Sample Amplification 2 Mix*) gegeben sowie 75µl des Amplifikation-Mastermixes MSM hinzugefügt. Die versiegelte Platte wurde erneut für jeweils 1 Minute bei 1600 rpm gevortext und bei 300xg abzentrifugiert. Anschließend erfolgte die Inkubation für 20–24 Stunden bei 37°C im Hybridisierungssofen (Illumina), um eine optimale isothermische Amplifikation der denaturierten DNA zu gewährleisten.

Enzymatische Fragmentierung der amplifizierten DNA-Produkte

Am darauf folgenden Tag wurde die Platte mit den amplifizierten DNA-Produkten für 1 Minute bei 100xg abzentrifugiert und die DNA mithilfe von jeweils 50µl FMS (*Fragmentation Solution*) enzymatisch fragmentiert. Die versiegelte Platte wurde erneut gevortext (1600 rpm, 1 Minute), bei 100xg für 1 Minute abzentrifugiert und für 1 Stunde bei 37°C im Heizblock inkubiert.

Präzipitation der DNA-Proben

Im nachfolgenden Schritt wurden die DNA-Proben durch Zusatz von PM1 (*Precipitation Solution*) und 2-Propanol präzipitiert. Hierfür wurden die Proben mit 100µl PM1 versetzt und die geschlossene Mikroplatte bei 1600 rpm gevortext, auf dem Heizblock für 5 Minuten bei 37°C inkubiert und für 1 Minute bei 100xg abzentrifugiert. Im Anschluss daran wurden die Proben mit jeweils 300µl 2-Propanol (100%) versetzt und die Platte mit einer neuen Deckelmatte verschlossen. Die Platte wurde erneut für 1 Minute bei 1600 rpm gevortext und für 30 Minuten bei 4°C inkubiert. Unmittelbar nach Zentrifugation bei 4000xg für 20 Minuten bei 4°C wurde der Überstand durch schnelles Umdrehen der Platte abgegossen. Die Platte wurde anschließend solange ausgeklopft bis die einzelnen Wells frei von restlicher Flüssigkeit waren. Zum Trocken der Pellets wurde die Platte mit der Öffnung nach unten für 1 Stunde auf ein Reagenzglasständer gelegt.

Resuspension der DNA

Anschließend wurden die DNA-Pellets in jeweils 42µl RA1 (*Resuspension, hybridization and wash solution*) resuspendiert. Die Platte wurde mit einer Folie abgedichtet und im Hybridisierungssofen für 1 Stunde bei 48°C inkubiert. Nach anschließendem 1-minütigen Vortexen (1800 rpm) und Zentrifugieren (300xg) wurden die DNA-Proben auf dem Heizblock bei 95°C für 20 Minuten denaturiert. Während einer sich etwa 30-minütigen anschließenden Abkühlungsphase der DNA-Proben bei Raumtemperatur wurden zum einen die benötigten BeadChips aus dem Kühlschrank geholt und zum anderen vorgesehene Hybridisierungskammern nach Anweisung zusammengebaut und vorbereitet. Hierfür wurden in entsprechende Reservoirs jeweils 200µl PB2-Puffer (*Humidifying buffer*) gegeben, um in der Kammer eine gewisse Luftfeuchtigkeit aufrechtzuerhalten und die Chips während der Hybridisierungskubation so vor dem Austrocknen zu schützen. Nach der Abkühlungsphase wurden die Proben bei 300xg abzentrifugiert.

Hybridisierung der DNA-Proben auf den HM450K BC

Anschließend konnten 12µl von jeweils zwölf fragmentiert und resuspendierten WGA-Proben mit einer Mehrkanalpipette in die entsprechenden Positionen eines HM450K BC überführt werden. Eine spezielle Folie auf dem Chip unterstütze diesen Prozess, indem die einzelnen Arrays auf dem HM450K BC nicht nur sichtbar getrennt, sondern auch schützend voneinander abgegrenzt vorlagen, wodurch die Pipettenspitzen sicher angelegt werden konnten. Die Folie verhinderte durch ihre Abgrenzung weiterhin eine Vermischung der einzelnen Proben im Falle eines Überlaufens und bestimmte gleichzeitig die Schichtdicke und das Volumen der zu hybridisierenden DNA-Proben.

Die beladenen Arrays wurden in die vorbereiteten Hybridisierungskammern gelegt. Diese wurden luftdicht verschlossen und über Nacht im Hybridisierungssofen für 16–24 Stunden bei 48°C inkubiert, um eine Hybridisierung der Proben durch Anlagerung an die Bead-Typ gebunden Locus-spezifischen Nukleotide (Linker-DNA) zu gewährleisten.

Waschen des HM450K BC

Nach erfolgter Inkubationszeit und einer etwa 25-minütigen Abkühlungsphase wurden die Chips durch mehrmaliges Schwenken in jeweils 200 ml der Waschlösungen WB1 und PB1, welche der Vorbereitung für die Färbung dienen, gewaschen. Die Chips wurden gemäß

Methoden

Protokoll für den Kammer-Ständer, der über ein Wasser-Zirkulationssystem auf eine Temperatur von 44°C gebracht wurde, präpariert und eingesetzt.

Enzymatische Einzelstrangverlängerung der hybridisierten DNA-Fragmente

Zuerst wurden nicht hybridisierte und nicht spezifisch gebundene DNA-Fragmente vom HM450K BC gewaschen, indem die Chips mit jeweils 150µl RA1 beladen und für 30 Sekunden inkubiert wurden. Dieser Vorgang wurde insgesamt 6 Mal durchgeführt. Dann wurde zuerst mit 450µl XStain-Lösung XC1 beladen und für 10 Minuten inkubiert, danach wiederholte sich der Ablauf für die XStain-Lösung XC2. Im darauffolgenden Schritt wurden durch Zugabe von 200µl TEM-Lösung (*Two Color Extension Master Mix*) und 15-minütiger Inkubation die 3'-Enden der jeweiligen Linker-DNAs an den entsprechenden Beads mittels markierter Nukleotide verlängert, wobei die Nukleotide C und G mit Biotin und die Nukleotide A und T mit Dinitrophenyl (DNP) markiert waren. Nach der spezifischen Extension wurde die hybridisierte DNA durch Zugabe von 450µl 95% Formamid/1 mM EDTA-Lösung wieder entfernt, mit 2x 450µl XStain-Lösung XC3 und einer einminütigen Inkubation neutralisiert und die Temperatur auf 32°C gebracht, um optimale Bedingungen für einen mehrschichtigen Färbungsprozess zu gewährleisten.

Färbeprozess zum Nachweis des Methylierungsgrads

Um die jeweiligen Methylierungsgrade nachweisen zu können, wurde im Folgenden ein spezifischer Färbeprozess angewandt. Für diesen Prozess wurde ein Sandwich-Assay verwendet, der auf der Bindung von zwei verschiedenen Antikörpern basiert. Hierfür wurden jeweils 250µl des *Labeling Two Color Master Mix* LTM auf die BeadChips gegeben und für 10 Minuten inkubiert. Die mit Biotin und DNP markierten Nukleotide werden dabei mit den in der Lösung enthaltenen fluoreszenzmarkierten primären Antikörpern angefärbt. Anschließend wurden nicht gebundene Antikörper mit jeweils 450µl XC3, 1-minütiger Inkubation, erneuter Zugabe von 450µl XC3 und einer Inkubationszeit von 6 Minuten wieder entfernt. Schließlich wurde mit Sekundärantikörpern in jeweils 250µl *Anti-Stain Two Color Master Mix* ATM und einer Inkubationszeit von 10 Minuten der Färbeprozess abgeschlossen. Anschließend wurde durch Zugabe von 450µl XC3, 1-minütiger Inkubation, erneuter Zugabe von 450µl XC3 und einer 6-minütigen Inkubation ungebundene Antikörper vom BeadChip gewaschen. Der beschriebene Sandwich-Assay wurde ein weiteres Mal wiederholt, wobei mit einem zusätzlichen Färbeschritt mit LTM und XC3, dessen Menge und Inkubationszeiten

bereits beschrieben wurden, abgeschlossen wurde. Zusätzlich wurde die letzte Inkubationsphase noch um 5 Minuten verlängert. Die BeadChips wurden aus der Apparatur entfernt und in einer mit etwa 310ml PB1 gefüllten Waschschale durch vorsichtiges Schwenken und 5-minütiges Ruhenlassen gewaschen. Dieser Schritt wurde danach mit dem Reagenz XC4 wiederholt, welches sich als dünne Schicht um den Chip legt und somit als UV-, Bleich- und Austrocknungs-Schutz dient. Anschließend wurden die Chips in einem Vakuumexsikkator bei etwa 680 mbar für etwa 1 Stunde getrocknet. Um eine ebene Fläche für das Scannen der Arrays zu gewährleisten, wurden XC4-Rückstände von der Rückseite der Chips vorsichtig mit 2-Propanol entfernt.

Auslesen der Fluoreszenzintensitäten mittels Illumina iScan

Anschließend wurden die Chips auf dem Illumina iScan bei zwei verschiedenen Anregungswellenlängen gescannt. Hierbei wird für jeden Kanal ein idat-File generiert (zwei Files pro Array und entsprechend 24 Files pro BeadChip), auf dem die einzelnen Signalintensitäten der Beads gespeichert sind. Mit der GenomeStudio Software (V2011.1) werden die Intensitäten in das Methylation Module (v1.9) eingelesen und können in Verbindung mit dem HM450K BC-Manifest (HumanMethylation450_15017482_v.1.1) in Form von avg.β-(Methylierungs-)Werten extrahiert werden.

3.3.2 Datenvorbereitung für Auswertung der Methylierungsdaten

Eine präzise Auswertung der Methylierungsdaten hinsichtlich der biologischen Varianz wurde durch eine Qualitätsanalyse mit anschließender Normalisierung des Rohdatensatzes gewährleistet.

3.3.2.1 Normalisierung mit GenomeStudio

Für die Initial-Analyse wurde die von Illumina vorgeschlagene Normalisierungsmethode innerhalb der GenomeStudio Software (v2011.1) genutzt. Die Normalisierung erfolgte gegen interne Kontrollen, die ebenfalls auf dem HM450K BC implementiert sind. Insgesamt werden über 90 gepaarte Normalisierungskontrollen verwendet, die alle innerhalb von Haushaltsgenen, also nicht über die DNA-Methylierung regulierten Genen, liegen und in ihrer 50er-Sequenz keinen CpG-Locus enthalten. Eine Sonde wird dabei mit A oder T verlängert, die andere Sonde entsprechend mit C oder G. Die Normalisierungswerte werden berechnet

und separat für die beiden Kanäle eingesetzt. Für den grünen Kanal werden die CG Kontroll-Werte genutzt, für den roten Kanal die AT Kontroll-Werte.

Die Intensitäten werden dann für alle Proben durch Multiplikation jeder Sonden-Intensität mit einem konstanten Normalisierungsfaktor normiert und durch den Mittelwert der Normalisierungskontrollen des entsprechenden Kanals geteilt. Der Normalisierungsfaktor setzt sich dabei aus der mittleren Intensität der AT und CG Normalisierungskontrollen in einer willkürlich ausgewählten Referenz-Probe zusammen (GenomeStudio Methylation Module v1.8 User Guide).

3.3.2.2 Ausschluss von Loci mit nicht quantifizierbaren Signalintensitäten

Um zu überprüfen, ob einzelne Loci ausgewertet werden können, wird der *detection.p-value* herangezogen. Dieser wird parallel mit den avg.β-Werten von der GenomeStudio Software ermittelt und beurteilt die einzelnen Signalintensitäten jeder Probe, indem er für jeden avg.β-Wert die Wahrscheinlichkeit angibt, dass das detektierte Signal vom Hintergrundsignal unterschiedlich ist. Für die durchgeführten Analysen wurde ein *detection.p-value* <0,01 gewählt, mit dem die Wahrscheinlichkeit für eine Falsch-Positiv-Rate auf <1% gesetzt wurde. Somit wurden alle CpG- und CpN-Positionen, die einen korrespondieren *detection.p-value* ≥0,01 aufwiesen, unabhängig von der Anzahl der betroffenen Fälle, aus der Analyse entfernt.

3.3.2.3 Reproduzierbarkeit der hybridisierten DNA-Proben

Eine erste Qualitätskontrolle der hybridisierten DNA-Proben lieferte die GenomeStudio Software durch die Anzahl der Targets unterhalb eines Schwellenwertes (*detection.p-value* =0,01 bzw. 0,05). Da die durchgeführten Analysen für einen *detection.p-value* <0,01 abgestimmt waren (siehe 3.3.2.2), wurde die prozentuale *loci detection rate* mit den hierfür entsprechenden Werten berechnet. Diese ist als das Verhältnis der Anzahl der Targets unterhalb eines Schwellenwertes zu der Gesamtanzahl der auf dem Array abgedeckten Loci (n=485577) definiert:

$$\text{Loci detection rate [\%]} = \frac{\text{Anzahl der Targets unterhalb eines Schwellenwertes}}{\text{Gesamtanzahl der auf dem Array abgedeckten Loci}} \times 100.$$

Für die Selektion der Proben in die Analyse wurde eine *loci detection rate* von mindestens 98% gewählt.

Als nächstes wurden die Proben hinsichtlich einer Korrelation untereinander überprüft. Hierfür wurde eine Korrelationsanalyse nach Pearson in der Statistik-Software R durchgeführt, wobei das globale Methylierungsmuster aller Proben miteinander verglichen wurde. Eine Gesamtkorrelation für jede Probe ergab sich demnach aus dem Mittelwert der Summen der einzelnen Korrelationen. Proben, die eine Gesamtkorrelation $R > 0,95$ aufwiesen, wurden in die Analyse aufgenommen.

Da es sich bei den DNA-Proben um gepaarte Proben handelte, konnte eine Genotypisierung von Zwillingspaaren und Zwillingen mit zwei verschiedenen Zeitpunkten unter Bezug der auf dem HM450K BC willkürlich ausgewählten 65 SNP-Loci durchgeführt werden. Mittels SNP-Korrelationsmatrix konnten somit DNA-Proben des Co- bzw. desselben Zwillings verifiziert bzw. ein potentieller Probenvertausch detektiert werden. Proben, deren Verwandtschaftsgrad mittels SNP-Analyse nachgewiesen werden konnte, wurden in die Analyse einbezogen. Im Gegensatz dazu wurden Proben mit zweifelhaftem Verwandtschaftsnachweis ausgeschlossen.

Im letzten Qualitätskontrollschritt wurden die Daten hinsichtlich eines potentiellen Batch-Effektes überprüft, da die einzelnen DNA-Proben sowohl zu verschiedenen Zeitpunkten prozessiert als auch auf unterschiedliche Arrays hybridisiert wurden. Nach Importieren des Datensatzes in den QluCore Omics Explorer v.2.3 (Lund, SE) wurde von diesem eine Hauptkomponentenanalyse (PCA) generiert. Nach graphischer Darstellung ließen sich vom *Batch*-Effekt betroffene Proben visuell identifizieren und wurden noch vor der anschließenden Normalisierung aus der Analyse entfernt. Die verwendeten Analyseparameter der PCA werden in den jeweiligen Abbildungen aufgeführt.

3.3.2.4 Hintergrundkorrektur und Normalisierung mit RnBeads

Zum Minimieren der technischen Varianz und potentieller *Batch*-Effekte wurde der finale Datensatz mit dem Paket „RnBeads“ (V.0.99.10) innerhalb des Statistik-Programms R (V.3.0.1) prozessiert [191]. Die Fluoreszenzintensitäten der qualifizierten Proben wurden hierfür in die R-Umgebung importiert. In einem ersten Schritt wurde die Hintergrundfluoreszenz des Arrays durch eine Hintergrundsubtraktion mittels NOOB-Methode [192] korrigiert. Weiterhin wurde zum Reduzieren der beobachteten Unterschiede in der β -Verteilung zwischen Infinium I und II Assay [51,193] die Normalisierungsmethode

SWAN (*subset-quantile within array normalisation*) angewendet [190]. Die Normalisierung nach SWAN basiert auf der Idee, dass sich die Gesamt-Intensitätsverteilung der Proben mit gleichem CpG-Gehalt unabhängig vom Assay-Typ ähnlich sein sollte. Daher wird eine Quantil-Normalisierung der 50bp-Proben, die die gleiche Anzahl an CpGs beinhalten, durchgeführt. Dabei werden die Infinium I- und II-Proben innerhalb eines einzelnen Arrays zusammen normalisiert, jedoch getrennt für die methylierten und unmethylierten Kanäle. Der Quantil-Normalisierung innerhalb des Arrays folgt dann die Zwischen-Array-Normalisierung [190].

3.3.2.5 Ausschlussbedingungen für die finale Analyse

Um zu verhindern, dass sich geschlechtsspezifische Unterschiede im Methylierungsmuster der Proben in der finalen Analyse hervorheben, wurden alle auf den beiden Geschlechtschromosomen X (n=11232) und Y (n=416) lokalisierten Loci ausgeschlossen.

Die bereits erwähnten 65 SNPs wurden ebenfalls von der finalen Analyse ausgeschlossen. Ferner wurden alle Sonden (n=17043), die in unmittelbarer Nähe (bis zu 3bp) ihrer Bindestellen einen bekannten SNP mit einer minimalen Allelfrequenz (MAF) von mindestens 5% beinhalteten, entfernt. Die SNP-Liste wurde von Illumina zur Verfügung gestellt (HumanMethylation450 v1.2 SNP Table (Stand: Februar 2013); www.illumina.com) und entsprechend der in dieser Arbeit gesetzten Parameter gefiltert.

3.3.3 Auswertung der Infinium HumanMethylation450 BeadChip-Daten

Die Auswertung der Methylierungsdaten gliederte sich in die Identifikation differentiell methylierter Genorte, wobei unterschiedliche Strategien verfolgt wurden, sowie deren bioinformatische Auswertung.

3.3.3.1 Identifizierung differentiell methylierter Genorte über die *False Discovery Rate* (FDR)

Der erste und wohl bedeutendste Ansatz in der Identifizierung differentiell methylierter Loci erfolgte über den von Qlucore Omics Explorer berechneten q-Wert. Dieser entspricht der *False Discovery Rate* (FDR), die den für multiples Testen korrigierten p-Wert angibt [194]. Mittels FDR wird die zu erwartende Fraktion von falsch-positiven Ergebnissen unter allen signifikanten Test-Ergebnissen vorgegeben. Für die Identifizierung differentiell methylierter

Loci wurde die FDR im Rahmen dieser Arbeit auf $q < 0,05$ gesetzt. Dies entspricht einem für multiples Testen angepassten p-Wert, der ein Falsch-Positiv-Ergebnis von $< 5\%$ beinhaltet. Mittels FDR konnten somit Genorte identifiziert werden, die in zuvor definierten Gruppen signifikante Unterschiede in ihrem Methylierungsmuster aufwiesen.

3.3.3.2 Identifizierung differentiell methylierter Genorte über p-Wert und $\Delta\beta$

Konnten über die FDR nur wenige signifikante Methylierungsveränderungen ermittelt werden, wurde nach differentiell methylierten Genorten in einem zweiten Ansatz mittels nicht korrigiertem p-Wert ($p < 0,05$; t-Test für unabhängige Stichproben) und einem $\Delta\beta > 0,05$ bzw. $< -0,05$ gefiltert.

3.3.3.3 Identifizierung differentiell methylierter Genorte im diskordanten Zwillingsmodell

Da es sich bei den Studienteilnehmern um Zwillinge handelte, konnte für die monozygoten Zwillinge ein zwillingsbasierter Ansatz angewendet werden, indem die Zwillinge eines jeden Paares, das sich in einem spezifischen Merkmal unterschied, als verbundene Stichproben behandelt wurden. Um möglichst reale, biologisch relevante Methylierungsunterschiede sowohl innerhalb der diskordanten Zwillingspaare als auch in der Gesamtheit der zu vergleichenden Phänotypen zu identifizieren, wurde ein stringentes Analyseverfahren gewählt. Dieses beinhaltete sowohl die Signifikanz als auch eine zwischen den zu vergleichenden Gruppen gleichgerichtete Methylierungsdifferenz ($\Delta\beta$ einheitlich < 0 oder > 0). Mittels Wilcoxon-Test wurden Genorte ermittelt, die sich in ihrer Methylierungsdifferenz signifikant und in die gleiche Richtung gehend in für Schichtarbeit diskordanten MZ-Paaren unterschieden ($p < 0,05$). Anschließend wurden diese mit den ermittelten Intra-Paar-Methylierungsdifferenzen der für Schichtarbeit diskordanten DZ-Paare (Schichtarbeiter–Tagarbeiter) hinsichtlich eines gleichen Vorzeichens überprüft.

War die Anzahl der diskordanten MZ-Paare für eine statistische Auswertung wie im Falle der Nachtarbeit zu gering ($n=3$), wurden die entsprechenden Intra-Paar-Methylierungsdifferenzen (Nachtarbeiter–Tagarbeiter) gebildet. Zur Identifizierung differentiell methylierter Loci, die mit einer anderen Methode validierbar sein würden, wurde für diese ein absolutes $\Delta\beta \geq 0,05$ festgelegt, was einer Methylierungsdifferenz von

mindestens 5% entspricht. Die Differenzen der ermittelten Loci innerhalb der für Nacharbeit diskordanten MZ-Paare sollten weiterhin die gleiche Methylierungsrichtung aufweisen. Somit qualifizierten sich nur differentiell methylierte Loci, die entweder in allen monozygoten Nacharbeitern eine einheitliche Hyper- oder Hypomethylierung im Vergleich zu den monozygoten Tagarbeitern aufwiesen. Die ermittelten Loci wurden anschließend in den diskordanten DZ-Paaren überprüft und nur diejenigen Loci ausgewertet, die ebenfalls in ihrer Intra-Paar-Differenz ein $\Delta\beta \geq 0,05$ oder $\Delta\beta \leq -0,05$ aufwiesen und in ihrer Methylierungsrichtung (Hyper- oder Hypomethylierung) mit der der MZ-Paare übereinstimmten.

3.3.4 Bestimmung des epigenetischen Alters nach Steve Horvath

Das von Steve Horvath entwickelte Modell zur Berechnung des epigenetischen Alters [152] kann auf einer frei zugänglichen Website (<https://dnamage.genetics.ucla.edu/>) genutzt werden. Zur Berechnung wurden daher die über die HM450K BC Methylierungsanalyse gewonnenen Methylierungsdaten aller Loci (n=485577) von allen Proben, die die Qualitätskriterien erfüllten, hochgeladen. Anschließend wurde das auf den DNA-Methylierungswerten der 353 CpGs basierende geschätzte epigenetische Alter der einzelnen Proben in einer Excel-Datei mitgeteilt. Aus chronologischem und epigenetischem Alter konnte daraufhin die Gesamtkorrelation R (Pearson) mit der Statistik-Software GraphPad Prism ermittelt werden.

3.3.5 Biologische Charakteristika differentiell methylierter Genorte

Nach Identifizierung differentiell methylierter Genorte wurden diese im nächsten Schritt nach biologischen Charakteristika untersucht.

3.3.5.1 Genontologie-Analyse

Mittels einer Genontologie- (GO-) Analyse können Gensignaturen, die beispielsweise aus einer Arrayanalyse resultieren, auf Anreicherungen innerhalb eines zuvor definierten Gensets überprüft und Informationen über zugrunde liegende biologische Prozesse ermittelt werden. Eine funktionelle Anreicherungsanalyse von GO-Kategorien innerhalb der differentiell methylierten Gene wurde mittels der web-basierten Software WebGestalt

(*WEB-based GENE SeT AnaLysis Toolkit*) durchgeführt [195]. Hierbei werden den abzufragenden Genen Unterkategorien der funktionellen Hauptgruppen „*biological process*“, „*molecular function*“ und „*cellular component*“ zugewiesen und diese gegenüber den GO-Annotationen der Referenzgen-Liste, die aus den verbliebenen auf dem HM450K BC unauffälligen Genen hervorging, auf Anreicherung überprüft.

3.3.5.2 Analyse von angereicherten Signalwegen

Mittels WebGestalt konnten weiterhin *KEGG- (Kyoto Encyclopedia of Genes and Genomes)* Signalwege, krankheitsassoziierte Gensets und Gene, die Bindungsstellen für bestimmte Transkriptionsfaktoren bieten, analysiert werden.

3.3.5.3 Weitere Anreicherungsanalysen

Die jeweiligen Gruppen der differentiell methylierten Loci wurden hinsichtlich einer Anreicherung von ausgewählten regulatorischen Elementen, Verteilungen im Gen-Zusammenhang, assoziierten DNA-Regionen und CpG-Lokalisationen analysiert. Über das Quotenverhältnis *Odds Ratio* (OR) und dem Chi-Quadrat-Test (χ^2 Test) wurde die Wahrscheinlichkeit eines bestimmten Merkmals für zwei Gruppen verglichen und somit die Häufigkeit eines bestimmten Merkmals in den jeweiligen differentiell methylierten Loci der Häufigkeit in der Referenzliste des HM450K BC (n=436136 Loci) gegenüber gestellt. Eine $OR > 1$ gab dabei eine Anreicherung, eine $OR < 1$ eine Abreicherung der untersuchten Merkmale in den differentiell methylierten Loci an.

Innerhalb der Gruppe der regulatorischen Elemente wurden drei regulatorische Regionen analysiert: Die von Illumina bioinformatisch ermittelten Enhancer-Bereiche sowie die experimentell ermittelten DNase I-hypersensitiven Stellen (DHS) und differentiell methylierten Regionen (DMR). Enhancer sind DNA-Sequenzbereiche, die die Transkription eines Genes verstärken und dabei orientierungsunabhängig wirken [196]. DHS sind nicht von Nukleosomen bedeckte Chromatinregionen, die aufgrund ihrer Zugänglichkeit für die DNase I mit aktiven regulatorischen Elementen assoziiert werden [197]. DMR sind Regionen im Genom, die an mehreren benachbarten CpGs eine differentielle DNA-Methylierung aufweisen [198]. Entsprechend dem Illumina Manifest sind auf dem Array die Informationen

zu cDMRs („krebs-spezifische“ DMRs) und rDMRs („neuprogrammierungs-spezifische“ DMRs) abgelegt.

Hinsichtlich der Verteilung im Gen-Zusammenhang wurden die Bereiche 5'-UTR, TSS1500, TSS200, Exon1, Genkörper und 3'-UTR analysiert.

Für die assoziierten DNA-Regionen wurden Gen-, Nicht-Gen- und Promotor-assoziierte Regionen berücksichtigt.

Bezüglich der CpG-Lokalisation wurde die Anreicherungsanalyse für N- und S-Shelf-, Island- sowie N- und S-Shore-Regionen durchgeführt.

3.4 Bisulfit-Pyrosequenzierung

Die Pyrosequenzierung von bisulfit-behandelter DNA dient zur Quantifizierung des Methylierungsgrades von individuellen CpG-Stellen. Bei dieser Methode wird ein bisulfit-behandelter DNA-Einzelstrang durch die Synthese seines Komplementärstrangs in Echtzeit sequenziert und das während des Strangaufbaus freigesetzte Pyrophosphat detektiert.

Hierfür wird die bisulfitkonvertierte DNA, die den abzufragenden CpG-Locus enthält, zunächst in einer PCR amplifiziert. Da für die sich anschließende Pyrosequenzierung ein DNA-Einzelstrang vorliegen muss, wird jeweils einer der beiden PCR-Primer mit Biotin markiert (siehe Abbildung 12). Der biotinylierte PCR-Strang wird daraufhin mittels Streptavidin-beschichteter Sepharose-Beads immobilisiert und mit einem Sequenzierungsprimer hybridisiert (siehe Abbildung 12A).

Nach Zugabe eines Enzym-Cocktails (DNA Polymerase, ATP-Sulfurylase, Luciferase und Apyrase) und Substrat-Gemisches (Adenosin-Phosphosulfat (APS) und Luciferin) werden von der Polymerase die zum Template-Strang komplementären Nukleotide eingebaut, die zuvor separat hinzu pipettiert wurden. Mit dem Nukleotideinbau wird Pyrophosphat (PPi) freigesetzt, woraufhin eine enzymatische Kaskade gestartet wird. Zunächst wird das freigesetzte PPi durch die ATP-Sulfurylase in Gegenwart des Substrates APS zu Adenosin-Triphosphat (ATP) umgesetzt. Luciferin wird daraufhin zusammen mit dem gebildeten ATP durch die Luciferase zu Oxiluciferin oxidiert, wobei Energie in Form von Licht freigesetzt wird. Die Intensität des Lichtsignals ist dabei proportional zu der freigesetzten Menge an PPi und somit zur Anzahl der eingebauten Nukleotide und wird als Signal-Peak in einem Pyrogramm

dargestellt. Nicht eingebaute Nucleotide und überschüssiges ATP werden durch die Apyrase abgebaut. Nach vollständiger Degradierung wird das nächste Nucleotid dem Reaktionsansatz hinzugefügt (Abbildung 12B).

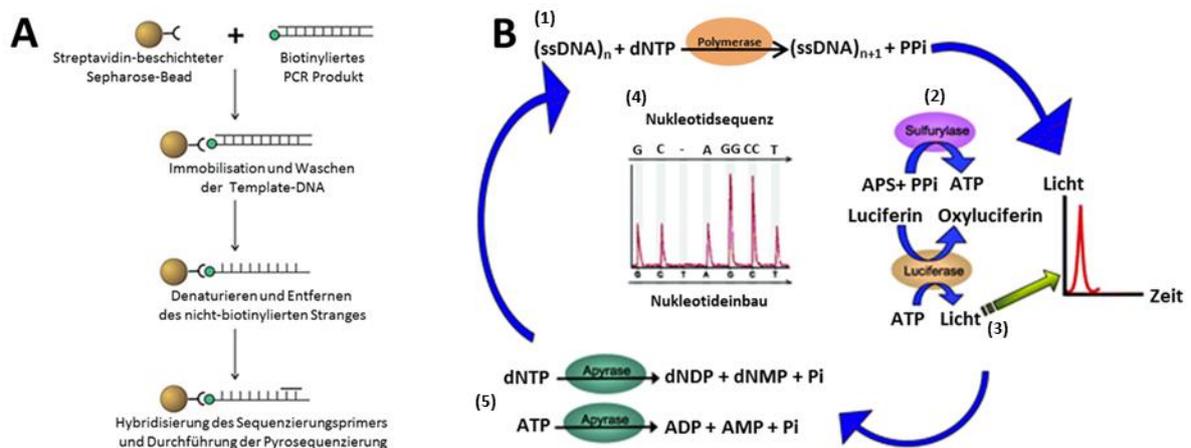


Abbildung 12: Prinzip der Bisulfit-Pyrosequenzierung. A. Separation des biotinylierten DNA-Strangs. Die zu untersuchende DNA-Region wird mittels PCR amplifiziert, wobei einer der Primer mit 5'-Biotin modifiziert vorliegt. Danach werden die Amplikons auf die Oberfläche der Beads immobilisiert und der nicht biotinylierte Strang nach Denaturierung entfernt. Der Sequenzierungsprimer wird hinzugefügt und hybridisieren folglich an das einzelsträngige DNA Template. Verändert nach <http://www.lifetechnologies.com/>. **B. Übersicht der enzymatischen Reaktionen während der Bisulfit-Pyrosequenzierung.** Nach Hybridisierung des Sequenzierungsprimers an das DNA-Template werden die Enzyme und Substrate zusammen mit einem der vier Nucleotide hinzugefügt. Ist das Nucleotid komplementär zur Base des Template-Strangs, wird es von der Polymerase unter Freisetzen einer äquimolaren Menge an Pyrophosphat eingebaut (1). In Anwesenheit von Adenosin-5'-Phosphosulfat (APS) wird das freigesetzte Pyrophosphat (PPi) von der Sulfurylase zu Adenosin-Triphosphat (ATP) konvertiert (2). Nach ATP-vermittelter Umwandlung von Luciferin zu Oxyluciferin durch die Luciferase wird ein Lichtblitz generiert, dessen Intensität proportional zur Anzahl der eingebauten Nucleotide ist (3). Das Lichtsignal wird als Peak in einem Pyrogramm dargestellt (4). Bevor das nächste Nucleotid hinzugefügt wird und die Signalkaskade somit erneut beginnen kann, wird jedes nicht eingebaute Nucleotid und überschüssige ATP von der Apyrase degradiert (5). Verändert nach [199].

Die Technologie der Bisulfit-Pyrosequenzierung diente in dieser Arbeit zur Validierung der mittels HM450K BC detektierten Methylierungsdifferenzen in Zwillingen, die sich innerhalb eines Zwillingspaares für Nacharbeit diskordant verhielten.

3.4.1 Methodischer Ablauf der Bisulfit-Pyrosequenzierung

Im Folgenden wird der technische Ablauf der Bisulfit-Pyrosequenzierung näher erläutert.

Primerdesign

Das Primerdesign für die jeweils zu analysierende DNA-Region wurde mit der PyroMark Assay Design Software erstellt. Hierfür wurde die Nucleotidsequenz der abzufragenden CpG-

Methoden

bzw. CpH-Stelle +/-1000bp aus dem Genome Browser (<http://genome.ucsc.edu/>) exportiert, in die Software eingespeist und *in-silico* bisulfit-konvertiert. Nach Festlegung der Zielregion, in der der zu analysierende Locus lokalisiert war, wurden von der Software automatisch Primerkombinationen ermittelt. Diese beinhalteten die beiden PCR-Primer (Forward und Reverse), von denen einer 5'-biotinyliert wird, sowie den Sequenzierungsprimer, der komplementär zu dem nach Amplifikation biotinylierten DNA-Strang ist. Die vorgeschlagenen Kombinationen wurden daraufhin manuell auf spezifische Qualitätsparameter überprüft. Das PCR Produkt sollte hierbei eine Größe von 100–300bp aufweisen und die verwendeten Primer keine Sekundärstrukturen wie Haarnadeln ausbilden. Um ein homogenes Hybridisieren der beiden Primer zu gewährleisten, wurde auf eine ähnliche Schmelztemperatur geachtet, die eine vorgegebene Differenz von 10°C nicht überschreiten sollte. Basierend auf den Vorschlägen der Software wurden die möglichen Primersequenzen manuell solange auf der Template-DNA-Sequenz verschoben und gegebenenfalls verlängert bzw. verkürzt bis eine optimale Kombination aller drei Primer vorlag. Anschließend wurde die Spezifität der ausgewählten PCR-Primerpaare für die zu analysierende DNA-Region durch Simulation einer *in-silico* PCR (<http://genome.ucsc.edu/>) überprüft. Weiterhin wurde jede Primersequenz auf SNPs mittels Sequenz-Abgleich im Genome Browser überprüft. Primer, die keinen bekannten SNP beinhalteten und spezifisch für das gewünschte Transkript waren, konnten daraufhin von der Firma Biomers (Ulm, DE) bezogen werden und sind in nachstehender Tabelle dargestellt.

Tabelle 4: Primersequenzen und PCR-Bedingungen für die Bisulfite-Pyrosequenzierung.

Gen	Position (hg19)	TargetID	PCR				T _A [°C]	Bisulfite-Pyrosequenzierung Sequenzierungsprimer (5'→3')
			Vorwärtsprimer (5'→3')	Rückwärtsprimer (5'→3')	Amplikonlänge [bp]	Amplikonlänge [bp]		
<i>SH2D1B</i>	chr1:162382848	cg01062020	GAATTTTATTTTGTA TTTGGTATAG	[Btm]TCAATTATTTCC TCTCTATATCA	246	55	TTATTTTGTATTTGGTATAGTG	
<i>SLC6A1</i>	chr3:11033246	cg01078332	TGAATGGGTTAAGAA AGATATTA	[Btm]ATCCCCAAAAAT AACCTTTA	186	55	AGAAAAGATATTAAGATTTTTTG	
<i>CLSTN2</i>	chr3:139678075	cg02760495	GGATGGTTGTTAAAG TTTTAGGT	[Btm]AAATACCCCTCC ATATCTATTTA	277	55	AAAGTTTTAGGTTTAT	
<i>TENM3</i>	chr4:183167912	cg04638710	[Btm]AAATGGGTTAT GTTTGTTAGA	AAAAATAAAATTCATC CCTACA	173	55	TTCATCCCTACACTACTA	
<i>CREB5</i>	chr7:28452066 chr7:28452093	cg27448433 cg19638435	TTTTTTTTTAGTATG GAGATTG	[Btm]CAATAATCTTAA ACTCCCTAAAT	228	55	TTTTAGTATGGAGATTGTAG	
N/A	chr12:108784905	cg12993543	GATTTAGTAGGGAG ATTTATTATAG	[Btm]TCATCAACCCCT TCTTAATA	163	55	TATAGAAAATTTTGTAGGAG	
<i>TMEM132D</i>	chr12:130357212	cg07230440	AGTTGTGGTATTATT TTTTAGTAT	[Btm]CCCCAAAACCTTA AAACTTAC	303	55	GTTTAGAAGAGTTATGTT	
<i>KANSL2</i>	chr12:49046495	ch.12.973812R	GTTTAGTTAGGGTTG AGGTATTTAT	[Btm]AAAAAATATCACT AAAAAACCCCAAT	240	55	GAATTTGAAATGGGTATA	
<i>MED1</i>	chr17:37560925	cg23679344	AGAAATGTATGGGG GGTAG	[Btm]ACCAATAAATCA TACAATATCCA	130	55	GAAGGAAAAAAAATGAG	

Position (hg19) gibt die genomische Position nach UCSC Genome Browser des zu analysierenden CpG-Locus (TargetID) an. TargetID: CpG-Locus Identifier, wobei cg# jeweils einen CpG-Locus und ch.# einen nicht-CpG-Locus repräsentieren. T_A: Annealing-Temperatur. Gen „N/A“: Nach UCSC Genome Browser konnte kein annotiertes Transkript mit dieser CpG-Position assoziiert werden. [Btm]: Biotinylierter Primer.

Methoden

Bisulfitkonvertierung der genomischen DNA

Für die Bisulfitkonvertierung der DNA wurde im Rahmen der Bisulfit-Pyrosequenzierung statt des EZ DNA Methylation™ Kits (Zymo Research) das EpiTect Bisulfite Kit (Qiagen, Hilden, DE) nach Herstellerangaben verwendet. Hierfür wurde jeweils 1µg genomische DNA mit RNase-freiem ddH₂O auf ein Volumen von 20µl gebracht. Nach Zugabe von 85µl Bisulfit-Mix und 35µl DNA-Schutz-Puffer wurde das Reaktionsgemisch im Thermocycler unter folgenden Bedingungen inkubiert:

Zyklus	Zeit [min]	Temperatur [°C]
Denaturierung	5	95
Inkubation	25	60
Denaturierung	5	95
Inkubation	85	60
Denaturierung	5	95
Inkubation	175	60
Lagerung	∞	20

Nach Abschluss der Konvertierung erfolgte die Aufreinigung der bisulfitkonvertierten DNA über die EpiTect Spin-Säulen. Hierfür wurden die Bisulfitreaktionsgemische kurz zentrifugiert und zusammen mit 560µl frischem BL-Puffer in die Spin-Säulen überführt und zum Binden der DNA an die Membran bei 14000xg für 1 Minute zentrifugiert. Der Überstand wurde verworfen und 500µl BW-Puffer zum Waschen der Membran auf die Säulen gegeben. Nach erneuter Zentrifugation (14000xg, 1 Minute) und Verwerfen des Überstandes wurden 500µl BD-Puffer auf die Säulen gegeben und zur alkalischen Desulfonierung für 15 Minuten bei RT inkubiert. Die Säulen wurden erneut bei 14000xg für 1 Minute zentrifugiert und der Überstand verworfen. Danach wurde jeweils zweimal mit 500µl BW-Puffer und erneuter Zentrifugation bei 14000xg für 1 Minute gewaschen. Die Säulen wurden in neue Reaktionsgefäße überführt und für 1 Minute bei 14000xg zentrifugiert, um jegliche Flüssigkeitsreste zu entfernen. Anschließend wurden die Säulen mit geöffnetem Deckel in saubere Reaktionsgefäße überführt und bei 56°C für 5 Minuten im Heizblock inkubiert. Anschließend wurden die Säulen in neue Reaktionsgefäße überführt und durch Zugabe von 20µl EB-Puffer bei 1-minütiger Zentrifugation (14000xg) eluiert. Die eluierte DNA wurde

Methoden

entweder direkt für die Amplifikation eingesetzt oder bis zur weiteren Verwendung bei -20°C gelagert.

Amplifikation der bisulfitkonvertierten DNA

Für die Amplifikation wurde das PyroMark-PCR Kit (Qiagen, Hilden, DE) entsprechend den Herstellerangaben verwendet und folgender Reaktionsansatz pipettiert:

Komponente (Stammlösung)	Volumen [μ l]	finale Konzentration
2x PyroMark® PCR Master Mix	12,5	1x
10x CoralLoad™ Konzentrat	2,5	1x
Forward Primer (20pmol/ μ l)	0,5	10pmol/ μ l
Reverse Primer (20pmol/ μ l)	0,5	10pmol/ μ l
ddH ₂ O	8	

Zunächst erfolgten die Primeraustestung und die Optimierung der Reaktionsbedingungen bei drei verschiedenen Annealing-Temperaturen (55°C, 60°C, 65°C) für jedes Primerpaar sowohl an zwei Positivkontrollen als auch an einer Negativkontrolle. Bei den Positivkontrollen handelte es sich um eine kommerziell bisulfitkonvertierte vollständig methylierte DNA-Probe (Millipore, Darmstadt, DE) und um eine bisulfitkonvertierte gepoolte DNA-Probe. Als Negativkontrolle wurde zum Ausschluss von nicht methylierungsspezifischen Primern eine genomische nicht bisulfit-behandelte DNA mitgeführt. Mittels Gelelektrophorese wurden die PCR-Produkte mit den optimalsten Reaktionsbedingungen visuell ermittelt und entsprechend für die zu analysierenden Amplifikate übernommen.

Für jeden Assay-spezifischen PCR-Ansatz wurden neben der zu untersuchenden DNA-Probe die bereits erwähnten Positivkontrollen sowie eine Negativkontrolle (ddH₂O) mitgeführt, um die Funktionalität und Qualität des jeweiligen Assays zu beurteilen.

Der Master Mix wurde entsprechend mit jeweils 1 μ l der beiden Positivkontrollen, der Wasserkontrolle und der zu untersuchenden bisulfit-behandelten DNA-Probe versetzt und im Thermocycler nach nachstehendem Programm inkubiert.

Methoden

Zyklus	Zeit [min]	Temperatur [°C]	Anzahl Zyklen
Initiale Denaturierung	15	95	1
Denaturierung	0,5	94	
Annealing	0,5	55	45
Elongation	0,5	72	
Elongation	10	72	1
Lagerung	∞	4	1

Hierbei wurde die Anzahl der Zyklen für die Amplifikation auf 45 Zyklen gesetzt, um einen möglichst vollständigen Verbrauch der biotinylierten Primer zu gewährleisten, die in der Pyrosequenzierung sonst zu fehlerhaften Ergebnissen führen könnten.

Erfolgreich amplifizierte Produkte wurden entweder direkt für die Pyrosequenzierung vorbereitet oder bis zur weiteren Verwendung bei -20°C gelagert.

Separation und Aufreinigung der biotinylierten PCR-Amplifikate

Für die Pre-Prozessierung wurde das biotinylierte PCR-Produkt über das Vakuum-gesteuerte Filtersystem PyroMark™ Vakuum Prep Workstation (Qiagen, Hilden, DE) aufgereinigt und denaturiert. Hierfür wurde das Amplifikat in einer 96 Well Mikroplatte mit 40µl Binding Puffer (Qiagen, Hilden, DE) und 5µl Streptavidin-beschichteten Sepharose-Beads (GE Healthcare, München, DE) versetzt. Durch eine 10-minütige Inkubation auf dem Plattenschüttler bei 1400rpm erfolgte eine Anlagerung der biotinylierten Amplifikate an die Beads. Anschließend wurde das PCR-Gemisch über die magnetischen Stifte des PyroMark™ Vakuum Prep Tools angesaugt, wobei die Sepharose-Beads mit den angelagerten PCR-Produkten an den integrierten Filtern haften blieben. Zum Entfernen noch vorhandener PCR-Reagenzien wurden die biotinylierten DNA-beladenen Beads mit 70% Ethanol gewaschen und danach in einer 0,2M Natronlauge denaturiert. Die nicht biotinylierten DNA-Stränge wurden durch einen weiteren Waschvorgang mit Waschpuffer entfernt. Die aufgereinigten biotinylierten Einzelstränge wurden anschließend in eine mit jeweils 0,3µl Sequenzierungsprimer und 25µl Annealing Puffer pro Well vorbereitete Mikroplatte überführt und für 2 Minuten bei 85°C inkubiert. Das Hybridisieren der Sequenzierungsprimer an die Amplikons erfolgte während der sich anschließenden Abkühlungsphase auf RT.

Pyrosequenzierungsreaktion

Die abgekühlte Platte wurde zusammen mit einer mit den vier Nukleotiden (dATP α S, dCTP, dGTP, dTTP), Enzym-Cocktail und Substrat-Gemisch befüllten Kartusche in den Pyrosequenzierer gestellt. Für das Befüllen der Kartusche wurden die von der PyroMark ID Software berechneten Volumina der benötigten Reagenzien herangezogen. Das Nukleotid dATP wurde hierbei durch das Basenalogon dATP α S ersetzt, was in dieser Form von der Luciferase nicht genutzt und gleichzeitig von der DNA-Polymerase effizient eingebaut werden kann. Nach Eingabe der für jeden Sequenzierungsprimer spezifischen Dispensions-Reihenfolge der Nukleotide wurde der Sequenzierlauf gestartet. Dabei werden automatisiert zuerst Enzym- und Substrat-Lösung und anschließend die Nukleotide in der vorgegebenen Reihenfolge aus der Kartusche in die jeweiligen Wells der Platte injiziert.

Auswertung der Pyrogramme

Die Auswertung der Pyrosequenzierung erfolgte mit der PyroMark CpG Software. Da dem Sequenzierungsansatz bei jedem CpG-Motiv sequentiell ein dCTP und dTTP injiziert wird, kann die Software den prozentualen Methylierungsgrad des betreffenden Cytosins umgehend berechnen. Dieser wird aus dem Verhältnis der jeweiligen Peakhöhen berechnet:

$$\text{Methylierung [\%]} = \frac{\text{Peakhöhe Cytosin}}{\text{Peakhöhe Cytosin} + \text{Peakhöhe Thymin}} \times 100.$$

Zur Bewertung der Qualität der jeweiligen Sequenzierung sind im Pyrogramm als interne Kontrollen sogenannte Referenz-Peaks hinterlegt, die von der Software erstellt wurden. Dabei handelt es sich um fixe Peaks, die in keiner variablen Position, wie in einem SNP, auftreten. Ferner dienen sie als Referenz zur Berechnung der einzelnen Peakhöhen (PyroMark[®] Q24 Software User Guide, 07/2009).

3.5 Statistische Verfahren

Für die grafische Aufbereitung der Daten wurden die Software-Programme R (V.3.0.1–3.1.0) und GraphPad Prism 5 (GraphPad Software Inc., San Diego, USA) verwendet.

- Mittels zweiseitigem Signifikanztest für unabhängige Stichproben wurden über das Quotenverhältnis *Odds Ratio* (OR) und dem Chi-Quadrat-Test (χ^2 Test) die Häufigkeitsverteilungen für spezifische Merkmale in der Gruppe der differentiell

Methoden

methylierten Loci und der Referenzliste des HM450K BC mit der Software GraphPad Prism bestimmt. Das Signifikanzniveau war mit $p < 0,05$ definiert.

- Eine Anreicherung von GO-Kategorien wurde mit dem hypergeometrischen Test in der web-basierten Software WebGestalt ermittelt und der Signifikanzwert p mit der FDR-Methode nach Benjamini und Hochberg [194] korrigiert. Vorgeschlagene funktionelle Kategorien mit $q < 0,05$ wurden als statistisch signifikant gewertet.
- Für Korrelationsanalysen wurde der Korrelationskoeffizient nach Pearson bzw. Spearman sowohl mit der Software R als auch mit GraphPad Prism berechnet.
- Für die Identifizierung von differentiell methylierten Loci zwischen zwei unabhängigen Gruppen wurde innerhalb des Qlucore Omics Explorers 2.3 (Qlucore, Lund, SE) jeweils ein Vergleich der Mittelwerte mittels t-Test für unabhängige Stichproben durchgeführt. CpG-Positionen zwischen zwei zu vergleichenden Gruppen wurden als differentiell methyliert identifiziert, wenn $q < 0,05$ (FDR) bzw. $p < 0,05$ (t-Test ohne Korrektur für multiples Testen) bei gleichzeitigem $|\Delta\beta| > 0,05$.
- Für die Identifizierung von differentiell methylierten Loci zwischen zwei verbundenen Stichproben wurde der Wilcoxon-Vorzeichen-Rang-Test (Wilcoxon-Test) verwendet. Mit diesem Test wurden DNA-Methylierungswerte bei monozygoten Zwillingspaaren über die Rangmittelwerte in der Software R verglichen. Das Signifikanzniveau war mit $p < 0,05$ definiert und wurde nicht für multiples Testen korrigiert.
- Für den Vergleich der Ergebnisse des HM450K BC und der Bisulfit-Pyrosequenzierung wurden im Rahmen der Etablierung des Arrays Bland-Altman-Diagramme erstellt. Dabei werden die Differenzen zwischen den zu vergleichenden Methoden für jede einzelne Messung gegen den Mittelwert der beiden Methoden aufgetragen. Ferner wird im Diagramm das 95% Konfidenzintervall (Mittelwert der Differenzen $\pm 2 \times \text{SD}$) angegeben und durch gepunktete Linien dargestellt. Für die Darstellung wurden die β -Werte (HM450K BC-Analyse) mit 100 multipliziert, um eine gleiche Größenordnung zu gewährleisten.

4 Ergebnisse

Ziel der vorliegenden Arbeit war die Identifizierung umweltinduzierter Effekte auf das Methylo monozygoter und dizygoter Zwillinge. Der paarweise Vergleich von MZ ermöglicht dabei die Auswirkungen bestimmter Umwelteinflüsse unabhängig vom individuellen genetischen Hintergrund zu ermitteln. Mittels BeadChip-Technologie sollte somit der Einfluss von Einflussfaktoren und typischer Lebensstil-Parameter von MZ und DZ auf Unterschiede im DNA-Methylierungsmuster untersucht werden.

Zur Einführung wird in diesem Kapitel zuerst die Etablierung der HM450K BC basierten DNA-Methylierungsanalyse präsentiert, welche als Bestandteil der Publikation von Roessler, Ammerpohl, Gutwein *et al.* ausführlich beschrieben ist [185]. Danach erfolgt die technische Interpretation der Methylierungsdaten, die Voraussetzung für alle weiterführenden Analysen war. Anschließend werden die Ergebnisse der einzelnen DNA-Methylierungsanalysen aufgeführt, wobei multiple Fragestellungen insbesondere zum Lebensstil beantwortet werden sollten. Aufgrund des Informationsgehalts an verschiedenen Umwelt- und Lebensstilfaktoren wurde im Speziellen zunächst der möglicherweise damit zusammenhängende Einfluss des Alters auf die DNA-Methylierung überprüft. Anschließend wurden die Parameter *Body Mass Index* (BMI) und Rauchen sowie Schichtarbeit und Nacharbeit hinsichtlich eines Einflusses auf das DNA-Methylo analysiert.

4.1 Etablierung des Infinium HumanMethylation450 BeadChips

Um die Eignung und Validität des HM450K BC zu überprüfen, wurden zunächst 18 invasive Brustkrebs-Proben, vier Brustkrebs-Zelllinien und vier unauffällige Brustgewebe-Proben mit dem HM450K BC analysiert [185]. Für die Etablierung dieser Arraytechnik eigneten sich diese Proben besonders gut, da es sich um einen Vergleich gesunden Gewebes gegen Krebs-Gewebe handelte und somit starke Effekte in Form einer differentiellen Methylierung zu erwarten waren [200,201]. Die anschließende Validierung der Ergebnisse erfolgte mittels Bisulfit-Pyrosequenzierung (BSPS). Hierbei konnten wir sowohl in humanen Zelllinien (n=4) als auch in humanen Gewebe-Proben (n=18) eine hohe Korrelation (Spearman) zwischen den Ergebnissen des HM450K BC und der BSPS zeigen (Abbildung 13A: R=0,88 und Abbildung 13B: R=0,85). Die Bland-Altman-Diagramme zeigten, dass die Korrelationskoeffizienten insbesondere bei niedrigen und hohen Methylierungswerten hoch waren, wohingegen die

Ergebnisse

Korrelationen zwischen HM450K BC und BSPS im mittleren Methylierungsbereich (25%–75%) schwächer erschienen (Abbildung 13C und D).

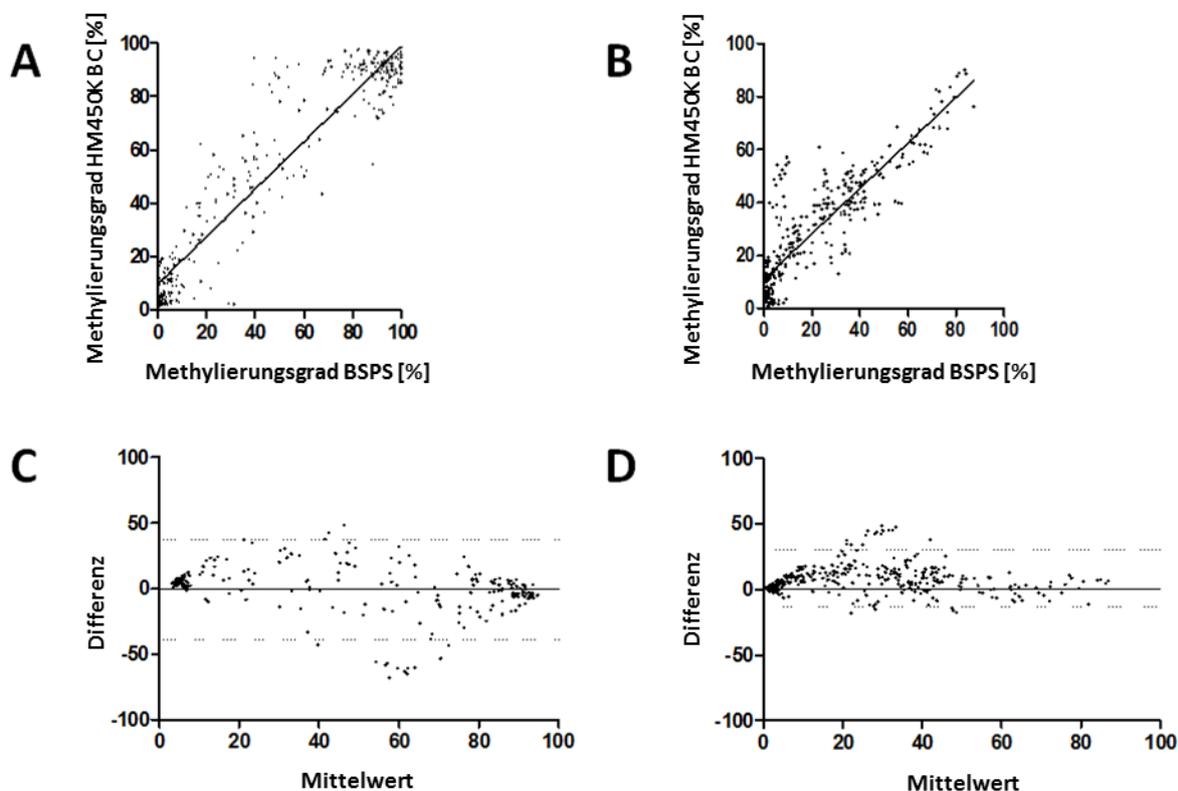


Abbildung 13: Validierung des HumanMethylation450 BeadChips (HM450K BC) mit der Bisulfit-Pyrosequenzierung (BSPS). Dargestellt sind Punktdiagramme für die Methylierungswerte identischer CpGs aus der BSPS (x-Achse) und den β -Werten aus der HM450K BC Methylierungsanalyse. **(A)** 340 CpGs in vier Zelllinien ($R=0,88$; $R^2=0,89$), **(B)** 352 CpGs in 18 humanen Krebs-Geweben ($R=0,85$; $R^2=0,77$). Jeder Punkt stellt den Vergleich der Methylierung (%) von Pyrosequenzierung gegenüber dem β -Wert für ein einzelnes CpG in einer Probe dar. **(C)** und **(D)** sind die zu **(A)** und **(B)** korrespondierenden Bland-Altman-Diagramme. R: Rangkorrelationskoeffizient nach Spearman. R^2 : Bestimmtheitsmaß. Abbildung übernommen aus [185].

Demnach konnten wir zeigen, dass die mittels HM450K BC erhobenen Methylierungsdaten mit einer unabhängigen Methode (hier: BSPS) verifizierbar waren, wenngleich mit gewissen Einschränkungen im mittleren Methylierungsbereich.

Mit der in unserer Arbeit nachgewiesenen Belastbarkeit der HM450K BC-Technologie (erfolgreiche genomweit quantitative DNA-Methylierungsanalyse sowie insgesamt hohe Korrelation zwischen HM450K BC und BSPS) wurde somit die Eignung des Arrays gezeigt. Dieser sollte im Folgenden für DNA-Methylierungs-Studien an einem gesunden Probanden-Kollektiv für die Beantwortung von Fragen insbesondere zum Lebensstil angewendet werden.

Ferner wurde aufgrund der insgesamt guten Übereinstimmung der mittels HM450K BC generierten Methylierungsdaten mit denen der BSPS, diese für die Validierung der Methylierungsdaten der für Nacharbeit diskordanten Zwillinge gewählt. Die gute Übereinstimmung der beiden Methoden resultierte hierbei aus dem Zusammenspiel einer hohen Gesamtkorrelation und einer mit dem HM450K BC gemessenen tendenziell höheren Methylierung.

4.2 Anwendung des Infinium HumanMethylation450 BeadChips zur Analyse ausgewählter Einflussfaktoren

Nach der erfolgreich durchgeführten Etablierung des Arrays stehen in den nächsten Unterkapiteln zunächst die Resultate der Qualitätskontrolle (siehe 4.2.1) und Normalisierung (siehe 4.2.2) im Vordergrund, da diese für die eigentlichen Analysen von essentieller Bedeutung waren.

Anschließend werden die jeweiligen Ergebnisse der auf verschiedene Fragestellungen basierenden durchgeführten DNA-Methylierungsanalysen an mono- und dizygoten Zwillingen präsentiert. Dabei wurde zuerst der Einfluss des kalendarischen (chronologischen) Alters auf das DNA-Methylierungsprofil analysiert (siehe 4.2.3). Zusätzlich wurde überprüft, ob sich im DNA-Methylierungsmuster ein kontinuierlicher Prozess des chronologischen Alterns widerspiegelt oder ob dieses möglicherweise vom jeweiligen Lebensstil beeinflusst sein könnte. Weiterhin wurden die epigenetischen Effekte aufgrund von Übergewicht (siehe 4.2.4), Zigarettenkonsum (siehe 4.2.5) und Schicht- (siehe 4.2.6) bzw. Nacharbeit (siehe 4.2.7) untersucht.

4.2.1 Qualitätskontrolle der Methylierungswerte

Zur Reduktion technischer Artefakte wurde vor der Beantwortung der biologischen Fragestellung eine mehrstufige Qualitätskontrolle durchgeführt.

4.2.1.1 Ausschluss von Loci mit nicht quantifizierbaren Signalintensitäten

Für die Detektion von Loci, deren Signalintensitäten sich mit einer Wahrscheinlichkeit $\geq 99\%$ vom Hintergrundsignal unterscheiden lassen, wurde für den von der GenomeStudio Software für jeden avg. β -Wert zugewiesene „*detection.p-value*“ der Schwellenwert auf $< 0,01$

gesetzt. Nach Ausschluss aller Loci mit *detection.p-value* $\geq 0,01$ gingen 472098 von 485577 Loci in die nachfolgende Qualitätsanalyse ein.

4.2.1.2 Reproduzierbarkeit der hybridisierten DNA-Proben

Für die Selektion von DNA-Proben, die verlässliche Methylierungswerte liefern würden, wurde die *loci detection rate* der einzelnen Proben überprüft, deren Schwellenwert zuvor auf 98% festgelegt wurde. Bei allen hybridisierten Proben ($n=146$; vgl. Tabelle 1) lag die *loci detection rate* im Median bei 99,91% (99,06%–100%). Dies deutete auf eine robuste technische Durchführung der Hybridisierung hin und somit gingen alle 146 Proben in die nächsten Schritte der Qualitätskontrolle ein.

4.2.1.3 Ausschlussbedingungen für die finale Analyse

Für die weitere Qualitätsanalyse wurden alle auf den Geschlechtschromosomen X ($n=11232$) und Y ($n=416$) lokalisierten Loci ausgeschlossen, um die geschlechtsspezifischen Unterschiede im DNA-Methylierungsmuster zu reduzieren. Des Weiteren wurde der Einfluss der genetischen Varianz auf die anschließende DNA-Methylierungs-Studie durch Ausschluss aller Loci, die SNPs mit einer MAF (*Minor Allele Frequency*) $>5\%$ bis zu 3 Basenpaare vor dem zu untersuchenden CpG-Locus beinhalteten ($n=17043$) sowie der von Illumina als SNP (rs) gekennzeichneten Loci ($n=65$) minimiert. Für die Qualitätskontrolle in Abschnitt 4.2.1.5 wurde hingegen auf die Daten der 65 exkludierten SNPs zugegriffen. Es verblieben 443823 analysierbare autosomale Loci.

4.2.1.4 Identifizierung technischer Artefakte mittels Korrelationsanalyse des genomweiten Methylierungsmusters

Da es sich bei dem Ausgangsmaterial um das gleiche Gewebe (peripheres Blut) gesunder Probanden handelte und daher weder gewebe- noch krankheitsspezifische Unterschiede im DNA-Methylierungsmuster zu erwarten waren, war somit ein im Mittel über alle Loci homogenes globales Methylierungsmuster zu erwarten. Daher wurde als weitere Qualitätskontrolle basierend auf dem Korrelationskoeffizienten nach Pearson (R) eine Matrix angefertigt, in der die jeweiligen Korrelationen der Proben zueinander hervorgingen. Um

Ergebnisse

Ausreißer bereits visuell ermitteln zu können, wurden die einzelnen Koeffizienten in einer Heatmap dargestellt (Abbildung 14).

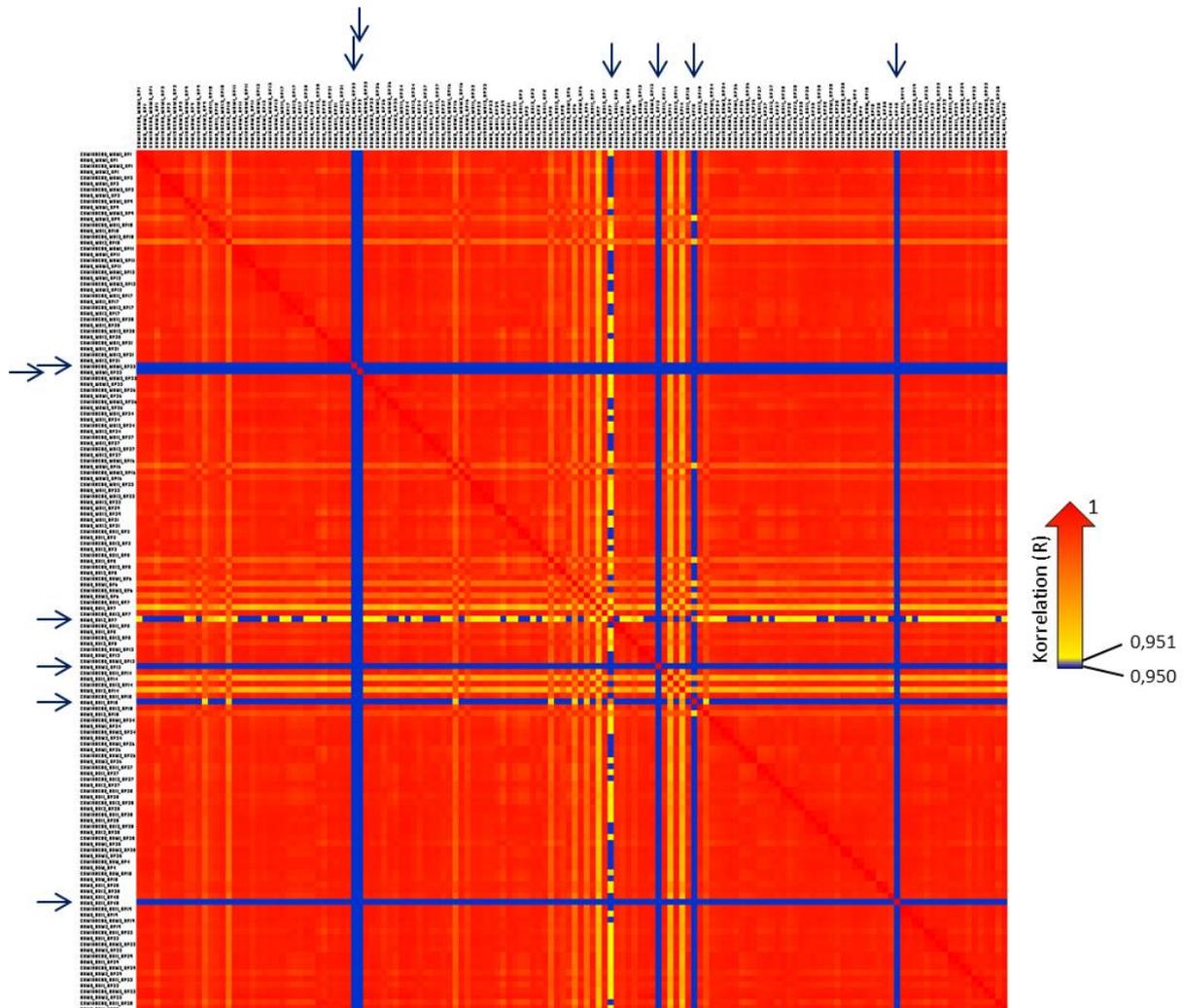


Abbildung 14: Korrelationsmatrix des globalen Methylierungsmusters von 146 DNA-Proben aus Zellen des peripheren Blutes. Alle Proben, deren mittleren Methylierungswerte eine Gesamtkorrelation $R \leq 0,95$ aufwiesen, hier durch blaue Pfeile markiert, wurden von der weiteren Analyse ausgeschlossen. Es verblieben 140 Proben. Die Hauptdiagonale enthält die jeweilige Eigenkorrelation der Proben, die somit immer $R=1$ ist. Aufgrund der Symmetrie spiegeln sich die Korrelationen zwischen den Proben an der Hauptdiagonale. R: Korrelationskoeffizient nach Pearson.

Es konnten insgesamt sechs Proben identifiziert werden, die aufgrund einer starken Abweichung der Korrelationskoeffizienten als technische Artefakte anzusehen waren. Die Gesamt-Korrelation dieser Proben belief sich dabei jeweils auf $R \leq 0,95$, weshalb diese Proben somit von der weiteren Analyse ausgeschlossen wurden, bevor die biologischen Fragestellungen untersucht wurden.

4.2.1.5 Identitätsüberprüfung mittels SNP-Analyse

Da es sich bei dem Studienkollektiv um Zwillinge handelte, die theoretisch 100% (monozygot) bzw. 50% (dizygot) ihrer genetischen Information teilen, wurden als weiterer Qualitätsparameter die auf dem Array hinterlegten SNPs ($n=65$) herangezogen. Anhand des individuellen SNP-Musters ließen sich die Zuordnung der einzelnen Zwillinge zu Zwillingspaaren auf ihre Richtigkeit überprüfen und ein möglicher Probenvertausch ausschließen. Demnach sollten monozygote Zwillinge ebenso wie identische Individuen (GEMINAKAR und SAME) eine sehr hohe Korrelation ($R \approx 1$) in ihrem SNP-Muster aufweisen. Für diese Qualitätskontrolle wurde die erstellte Korrelationsmatrix zur besseren Visualisierung ebenfalls in einer Heatmap dargestellt (Abbildung 15).

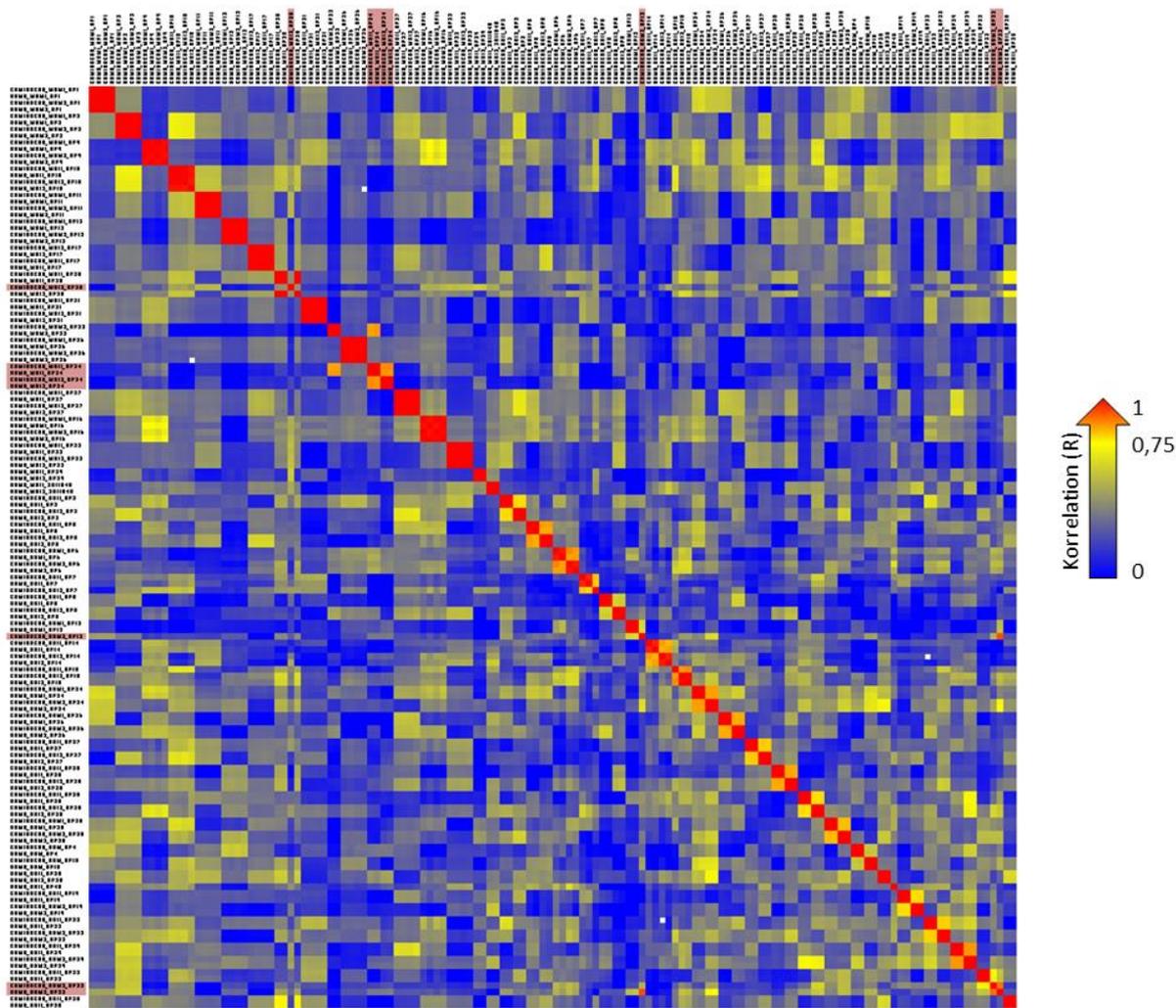


Abbildung 15: SNP-Korrelationsmatrix von 140 Zwillingenproben. Proben mit schwachen Korrelationen, hier in altrosa hinterlegt, wurden aus der weiteren Analyse entfernt (siehe Text). Die Hauptdiagonale enthält die jeweilige Eigenkorrelation der Proben, die somit immer $R=1$ ist. Aufgrund der Symmetrie spiegeln sich die Korrelationen zwischen den Proben an der Hauptdiagonale. R: Korrelationskoeffizient nach Pearson.

Aufgrund ihrer Korrelationskoeffizienten mussten von den 140 Proben insgesamt weitere acht Proben ausgeschlossen werden. Dabei handelte es sich um einen monozygoten Zwilling zum Zeitpunkt 1 (GEMINAKAR_MZF2_ZP20), der weder mit sich zu Zeitpunkt 2 ($R=0,29$) noch mit seiner Zwillingsschwester zu beiden Zeitpunkten (jeweils $R=0,28$) korrelierte. Weiterhin wurden die Proben eines monozygoten Zwillingspaars (MZF_ZP7) aus beiden Studien (GEMINAKAR und SAME) ausgeschlossen, da sich die Korrelation der monozygoten Zwillinge ($R=0,73$) wie eine vergleichbare Korrelation von dizygoten Zwillingen verhielt. Ferner wurde eine dizygoten Zwillingssprobe (DZM2_ZP13) mit ihren zwei Zeitpunkten ausgeschlossen, da sie mit sich selbst schlecht korrelierte ($R=0,78$) sowie zur Sicherheit die Probe eines dizygoten Zwillingss (GEMINAKAR_DZM2_ZP33), der mit seinem Zwillingssbruder eine ungewöhnlich geringe Korrelation ($R=0,42$) aufwies. Die ausgeschlossenen Proben wurden in Abbildung 15 in altrosa farblich hinterlegt.

4.2.1.6 Identifizierung potentieller *Batch*-Effekte mittels *Principal Component Analyse* (PCA)

Die verbliebenen 132 Proben wurden erneut dem unter den Abschnitten 4.2.1.1–4.2.1.3 beschriebenen Filterungsprozessen unterzogen. Es wurden 448136 Loci identifiziert und von diesen in einem letzten Qualitätskontrollschritt mittels Qluore Omics Explorer eine unsupervidierte PCA generiert, um potentielle *Batch*-Effekte der Hybridisierung zu identifizieren. Die Verteilung der Proben auf die jeweiligen Arrays erfolgte zufallsbedingt. Die unsupervidierte PCA ist für die verschiedenen Arrays ($n=14$) in Abbildung 16A dargestellt. Da die DNA-Proben aufgrund der durch unsere Kooperationspartner fortbestehenden Proben-Rekrutierung folglich nur chargenweise angefordert werden konnten, erfolgte die Hybridisierung der DNA-Proben chargenabhängig. Die einzelnen Chargen ($n=6$) wurden ebenfalls in der PCA dargestellt (siehe Abbildung 16B), um somit potentielle Chargeneffekte ohne biologischen Hintergrund identifizieren zu können.

Ergebnisse

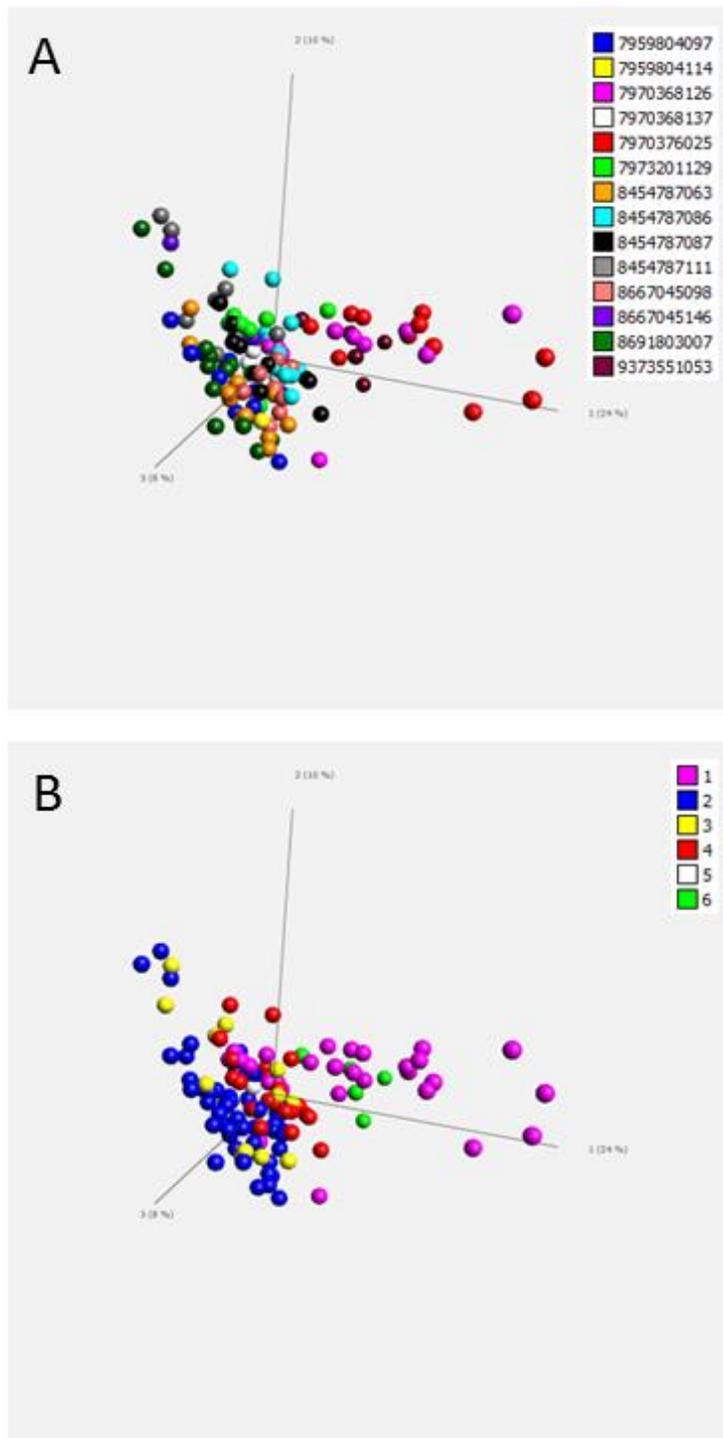


Abbildung 16: Unsupervidierte *Principal Component Analyse* (PCA) der 132 Proben. In (A) sind die einzelnen Arrays (n=14) entsprechend des Array-Barcodes farblich dargestellt, während in (B) die einzelnen Chargen (1–6) farblich hinterlegt wurden.

Nachdem in der PCA (132 Proben und 448136 Loci) weder ein deutlicher Array- noch Chargen-*Batch*-Effekt identifiziert werden konnte, wurde im nächsten Schritt die Varianz innerhalb der unsupervidierten PCA angepasst. Über den Varianzfilter lassen sich diejenigen Loci ausschließen, die innerhalb aller Proben nur sehr geringe Methylierungsveränderungen aufweisen. Bei einem $\sigma/\sigma_{\max} > 0,3$ wurden folglich nur CpG-Loci berücksichtigt, die eine

Ergebnisse

Varianz von mehr als 30% der Varianz des CpG-Locus mit der größten Variabilität über allen Proben aufwiesen. Bei der genannten Einstellung umfasste die erste PC 16% der Gesamtvarianz, die auf Proben innerhalb der 1. Charge (siehe Abbildung 17A, pink) fiel und diese deutlich von den übrigen Chargen abgrenzte.

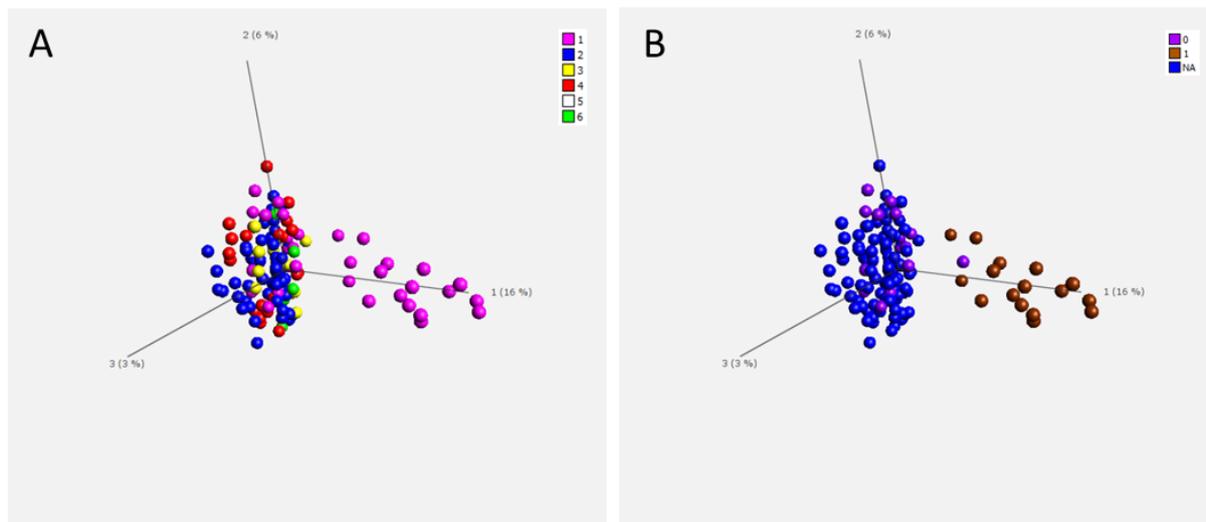


Abbildung 17: Unsupervidierte *Principal Component Analyse* (PCA) ($\sigma/\sigma_{\max}>0,3$) der 132 Proben. A. Nach Setzen des Varianzfilters innerhalb der unsupervidierten PCA spaltete sich eine separate Gruppe von Proben deutlich ab, welche Proben aus der Charge 1 (pink) zuzuordnen war. **B.** Nach Aufspaltung der Charge 1 in den GEMINAKAR- (Code 0) und SAME- (Code 1) Datensatz konnten 16/18 Proben aus dem beobachteten separaten Cluster dem SAME Datensatz zugeordnet werden und wurden daher aufgrund eines technischen Artefakts in Form eines Chargeneffektes aus der Analyse ausgeschlossen.

Nachdem die Proben aus der 1. Charge (siehe Abbildung 17B) in das etwa 10 Jahre alte DNA-Material der GEMINAKAR-Studie (0) und das frisch extrahierte DNA-Material der SAME-Studie (1) unterteilt wurden, konnten 16/18 Proben aus der separierten Gruppe den SAME-Proben zugeordnet werden. Aufgrund dieses identifizierten Chargeneffekts wurde der gesamte Teil der SAME-Proben aus der 1. Charge (n=16 Proben) von der finalen Analyse ausgeschlossen.

4.2.1.7 Zusammenfassung des Qualitätskontrollprozesses

Durch die gesetzten Qualitätsparameter mussten insgesamt 30 von 146 Proben von der finalen Analyse ausgeschlossen werden. Ausschließlich der Rohdatensatz der verbliebenen 116 Proben wurde daraufhin in GenomeStudio importiert und nach den oben beschriebenen Ausschlusskriterien erneut gefiltert, was in 453136/485577 Loci (93,3%) resultierte, die in die weitere Analyse eingeschlossen wurden. Eine schematische Übersicht ist in Abbildung 18

dargestellt. Die ausgeschlossenen Loci beinhalteten 2900 Loci mit einem *detection.p-value* $\geq 0,01$ in mindestens einer der analysierten DNA-Proben, 11217 Loci auf den beiden Geschlechtschromosomen sowie die bereits erwähnten von Illumina deklarierten 65 SNPs und weitere 18259 SNPs mit einer MAF $>5\%$.

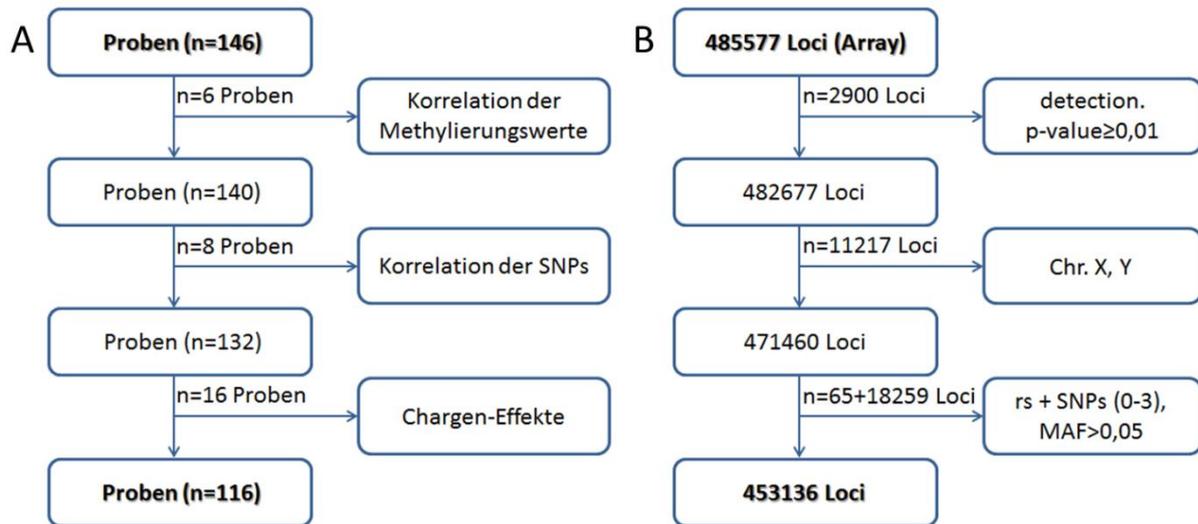


Abbildung 18: Schematische Darstellung der Qualitätsparameter und Ausschlusskriterien der initialen 146 Proben mit jeweils 485577 Loci. Nach dem Filterungsprozess verblieben (A) 116/146 Proben und (B) 453136/485577 analysierbare Loci.

4.2.2 Normalisierung der Signalintensitäten mittels RnBeads

Mittels der RnBeads-Pipeline [191] wurden die Signalintensitäten in der Statistik-Software R nach der SWAN-Methode normalisiert, um die technische Varianz zwischen individuellen Messungen zu minimieren. Der SWAN-Methode steht die von Illumina vorgeschlagene Normalisierungsmethode, die innerhalb der GenomeStudio Software durchgeführt wird, gegenüber. Um eine Einschätzung über die Leistung der SWAN-Methode zu erhalten, wurden beispielhaft an zehn monozygoten Zwillingspaaren der SAME-Studie die auf dem HM450K BC implementierten SNPs (n=65) sowohl nach der Illumina-Methode als auch nach der SWAN-Methode normalisiert und verglichen. Da die beiden Zwillinge eines monozygoten Zwillingspaares als genetisch identisch angesehen werden, sollte hier die Intra-Paar-Differenz aufgrund gleicher SNPs, also die Differenz der jeweiligen β -Werte beider Zwillinge eines Paares, theoretisch nahe bei 0 liegen. Die Ergebnisse der beiden Normalisierungsmethoden sind in Abbildung 19 dargestellt. Demnach lieferte die Normalisierung nach SWAN (siehe Abbildung 19B) im Vergleich zu dem von Illumina angewandten Normalisierungsmodell

Ergebnisse

(siehe Abbildung 19A) ein durchwegs konstanteres SNP-Muster. Ferner ordnete sich der Median der einzelnen Zwillingspaare bei der SWAN-Methode für die 65 SNPs näher um 0 herum an als bei der Illumina-Methode.

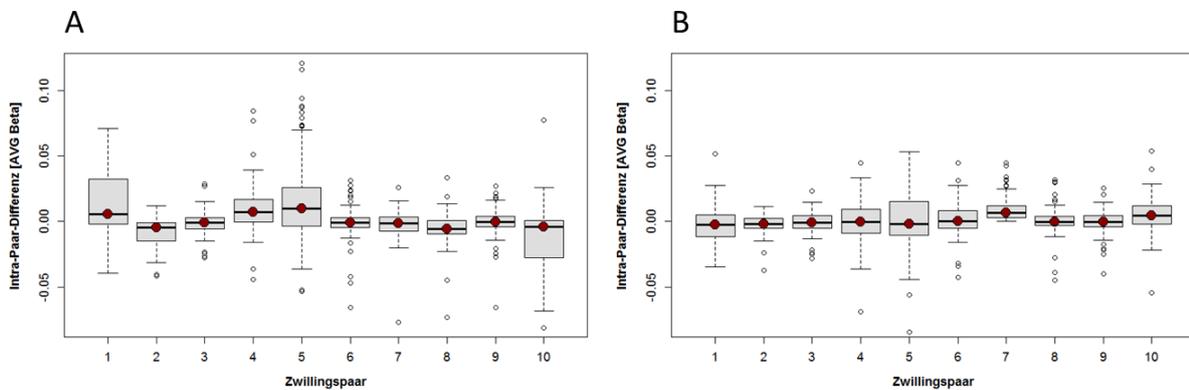


Abbildung 19: Boxplots der Intra-Paar-Differenzen (IPDs) von 10 monozygoten Zwillingspaaren (SAME-Studie) hinsichtlich der auf dem HM450K BC implementierten SNPs nach zwei verschiedenen Normalisierungsmethoden. (A) zeigt die IPDs entsprechend der Illumina-Normalisierung gegen interne Kontrollen. In (B) sind die IPDs nach der SWAN-Methode dargestellt.

Aufgrund der dargestellten Resultate wurde für diese Studie nach erfolgreichem Durchlaufen der Qualitätskontrolle die Normalisierung der β -Methylierungswerte mit der SWAN-Methode vorgenommen.

Nach Abschluss der Untersuchung technischer Aspekte werden im Folgenden die einzelnen DNA-Methylierungsanalysen aufgrund verschiedener Umweltfaktoren dargestellt.

4.2.3 Auswirkung des chronologischen Alters auf die genomweite DNA-Methylierung

Zur Überprüfung des Einflusses vom Lebensalter in Jahren auf die genomweite DNA-Methylierung wurden alle hybridisierten GEMINAKAR (n=63) und SAME (n=53) Proben, welche die Qualitätskriterien erfüllten, in die Analyse einbezogen (vgl. Tabelle 3). Die Altersspanne der Probanden bei Probenentnahme erstreckte sich hierbei von 20–63 Jahre.

4.2.3.1 Identifizierung altersabhängig methylierter CpG-Loci über die FDR

Für die Identifizierung einer altersabhängigen DNA-Methylierung wurde für die 436136 autosomalen Loci eine lineare Regression innerhalb des Qlucore Omics Explorers durchgeführt. Dadurch wurden Loci identifiziert, deren Methylierungswerte eine signifikante Änderung mit dem chronologischen Altern zeigten, sich also mit zunehmendem Alter progressiv veränderten. Bei einer FDR $q < 1 \times 10^{-11}$ und Varianz $\sigma/\sigma_{\max} > 0,1$ konnte eine altersabhängige DNA-Methylierung an 111 CpG-Positionen identifiziert werden, die innerhalb von 63 Genorten lokalisiert waren und in Abbildung 20 in einer Heatmap (A) und korrespondierenden PCA (B) dargestellt sind. In 91 der 111 Loci (82%) konnte mit zunehmendem Alter eine Zunahme der Methylierung beobachtet werden, wohingegen sich in 20 Loci (18%) simultan eine Abnahme der Methylierung manifestierte (siehe Abbildung 20A). Die Zuordnung der 111 altersassoziierten CpG-Loci ist in Tabelle A2 im Anhang hinterlegt. Aufgrund der Tatsache, dass die verschiedenen Blutzelltypen verschiedene DNA-Methylierungsprofile aufweisen [202,203] und sich die Zellzusammensetzung des Blutes mit dem Alter ändert [204,205], sollte die Möglichkeit in Betracht gezogen werden, dass die zelltyp-spezifischen Unterschiede die detektierte Variabilität der 111 identifizierten Loci beeinflussen könnte, da in diesem Ansatz nicht für die Blutzellzusammensetzung korrigiert wurde.

Ergebnisse

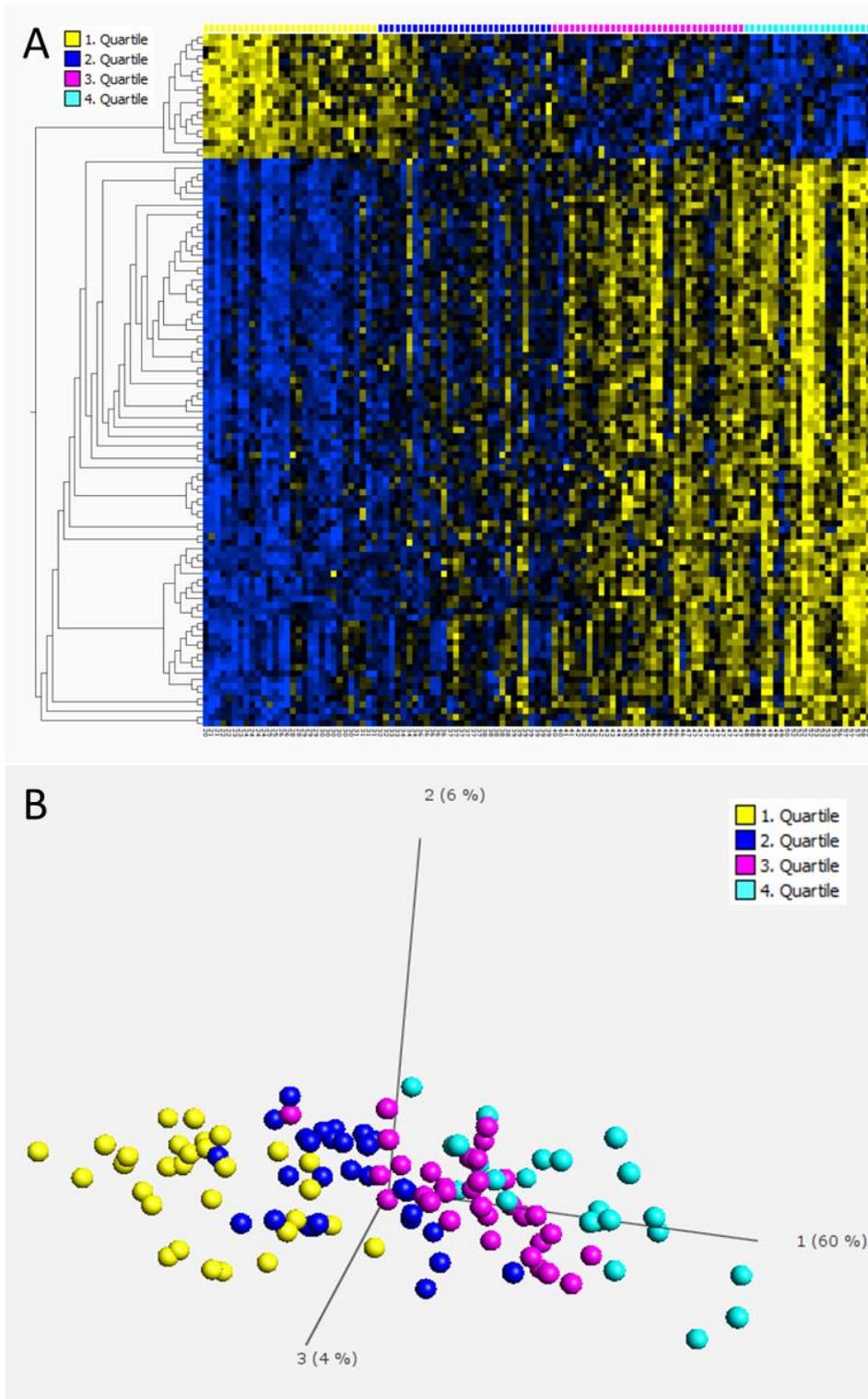


Abbildung 20: Differentielle Methylierungsanalyse von 116 Zwillingsproben im Alter von 20 und 63 Jahren. Eine lineare Regressionsanalyse resultierte in 111 differentiell methylierten CpGs ($q < 1 \times 10^{-11}$, $\sigma / \sigma_{\max} > 0,1$). **(A)** zeigt die entsprechende Heatmap (*average linkage*), in der 91/111 CpGs mit einer Zunahme der Methylierung und entsprechend 20/111 CpGs mit einer Abnahme der Methylierung während des Alterungsprozesses assoziiert waren. Spalten repräsentieren Proben, Reihen hingegen CpG-Loci. Innerhalb der Heatmap gibt ein blauer Farbcode eine geringe Methylierung an, während gelb einen höheren Methylierungsgrad widerspiegelt. Auf der x-Achse der Heatmap ist die Probenzuordnung über das chronologische Alter in Jahren angegeben. Zur Visualisierung wurden die Mittelwerte sowie die Varianz für jeden CpG-Locus gegen 0 bzw. 1 normalisiert ($\text{mean}=0$, $\text{var}=1$). In **(B)** sind die ersten drei PCAs der korrespondierenden PCA dargestellt. Zur Visualisierung wurde das Lebensalter in die entsprechenden Quartilen eingeteilt.

Ergebnisse

Dieser Analyse konnte weiterhin entnommen werden, dass das chronologische Altern nicht zu einer kontinuierlichen Änderung des DNA-Methylierungsmusters führte (Abbildung 20). Im untersuchten Zwillingskollektiv fanden sich beispielsweise Individuen im jüngeren Lebensjahr (1. Quartile: 20–31 Jahre), die aufgrund ihres DNA-Methylierungsmusters der 111 CpG-Positionen der nächstälteren Gruppe (2. Quartile: 32–39 Jahre) zugeordnet werden konnten. Analog fanden sich auch Individuen aus der 4. (48–63 Jahre) und 3. Quartile (40–47 Jahre), die sich in den nächstjüngeren Alters-Quartilen einordneten. Demnach ließe sich das biologische Alter auch im DNA-Methylierungsmuster der besagten CpG-Loci erkennen, welches durch den Lebensstil geprägt sein könnte. So konnte beispielsweise das DNA-Methylierungsmuster eines 34-jährigen rauchenden (SAME_DZM2_ZP22) und eines 38-jährigen nichtrauchenden Schichtarbeiters (SAME_MZF1_ZP20) mit einem älteren biologischen Alter assoziiert werden. Hingegen konnte aber auch ein 47-jähriger nichtrauchender Nachtarbeiter (SAME_DZM2_ZP29) sowie ein 49-jähriger gelegentliches rauchender Schichtarbeiter mit Nachtdienst (SAME_MZM1_ZP36) mit einem jüngeren biologischen Alter assoziiert werden.

Die Bestimmung des epigenetischen Alters wurde, wie von Steve Horvath beschrieben [152], an den 116 Proben durchgeführt. Die daraus ermittelte Korrelation zwischen chronologischem und epigenetischem Alter ist in nachstehender Abbildung dargestellt.

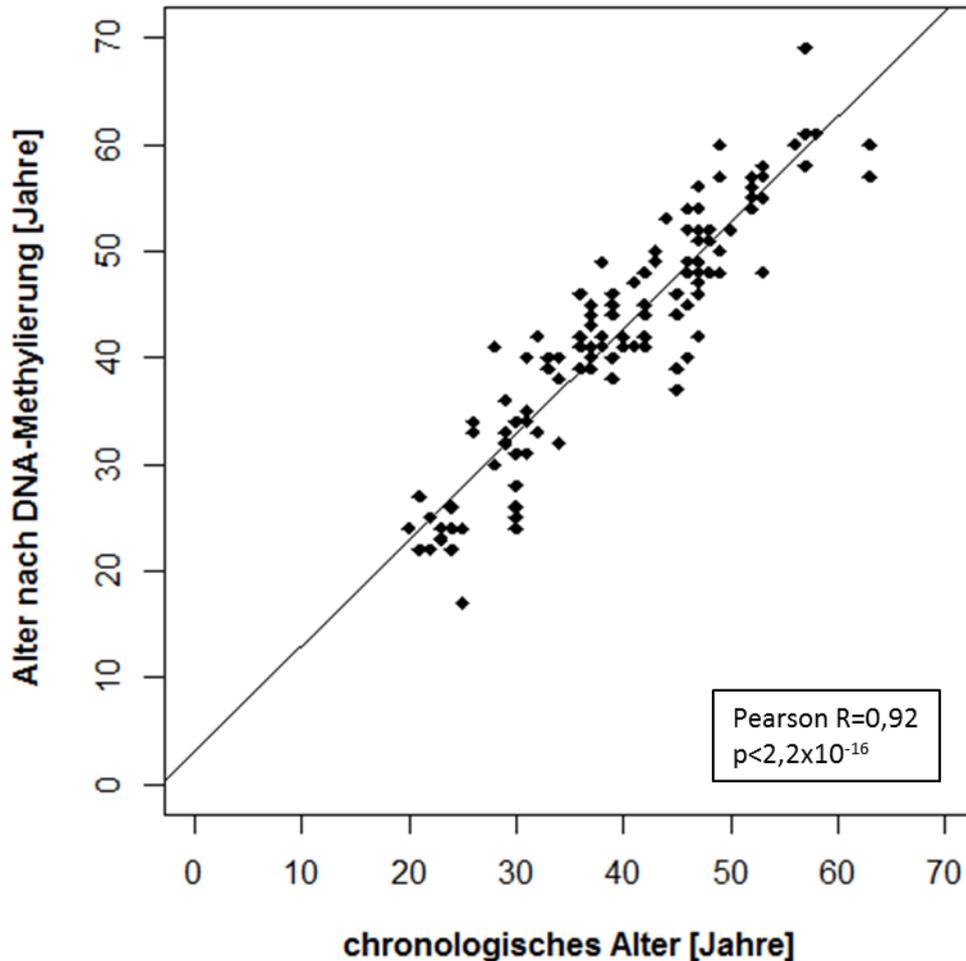


Abbildung 21: Korrelation zwischen dem chronologischen und epigenetischen Alter basierend auf dem Algorithmus von Horvath [152]. Jeder Datenpunkt stellt eine analysierte DNA-Probe dar. Auf der x-Achse ist das chronologische Alter, auf der y-Achse das Alter aufgrund der 353 CpG-Signatur nach Horvath jeweils in Jahren angegeben. Die berechnete Korrelation betrug $R=0,92$ ($p<2,2 \times 10^{-16}$; t-Test) mit einem Standardfehler von 4,2 Jahren.

Die Korrelation des mittels Algorithmus berechneten und des tatsächlichen Alters lag bei 0,922 mit einem Standardfehler von 4,2 Jahren. Während für 26 Zwillinge (22%) das epigenetische Alter mit dem chronologischen Alter nahezu übereinstimmte (± 1 Jahr), zeigten 76 Zwillinge (66%) ein älteres und 14 Zwillinge (12%) ein jüngeres epigenetisches Alter. Die Tendenz, dass mehr Zwillinge ein älteres als ein jüngeres epigenetisches Alter aufwiesen, konnte in der in dieser Arbeit durchgeführten Analyse auch bestätigt werden (siehe Abbildung 20A).

Aufgrund des gleichen genetischen Hintergrundes der MZ-Paare (MZPs) wurde die Analyse im Folgenden auf diejenigen Paare beschränkt, deren DNA-Proben zu zwei verschiedenen Zeitpunkten (GEMINAKAR und SAME) analysiert werden konnten (siehe Abbildung 22). Dies schloss neun MZPs ein, deren chronologisches Alter mit dem epigenetischen Alter zu

Ergebnisse

Zeitpunkt 1 (1997–2000) mit $R=0,971$ und einem Standardfehler von 2,2 Jahren (Abbildung 22A) korrelierte. Etwa 10–15 Jahre später (Zeitpunkt 2) betrug diese Korrelation $R=0,901$ mit einem Standardfehler von 4,3 Jahren (Abbildung 22B). Dabei unterschied sich das epigenetische Alter in der Intra-Paar-Differenz der jeweiligen Zwillingspaare zu Zeitpunkt 1 signifikant von der ermittelten epigenetischen Altersdifferenz zu Zeitpunkt 2 ($p<0,0092$; t-Test für gepaarte Stichproben).

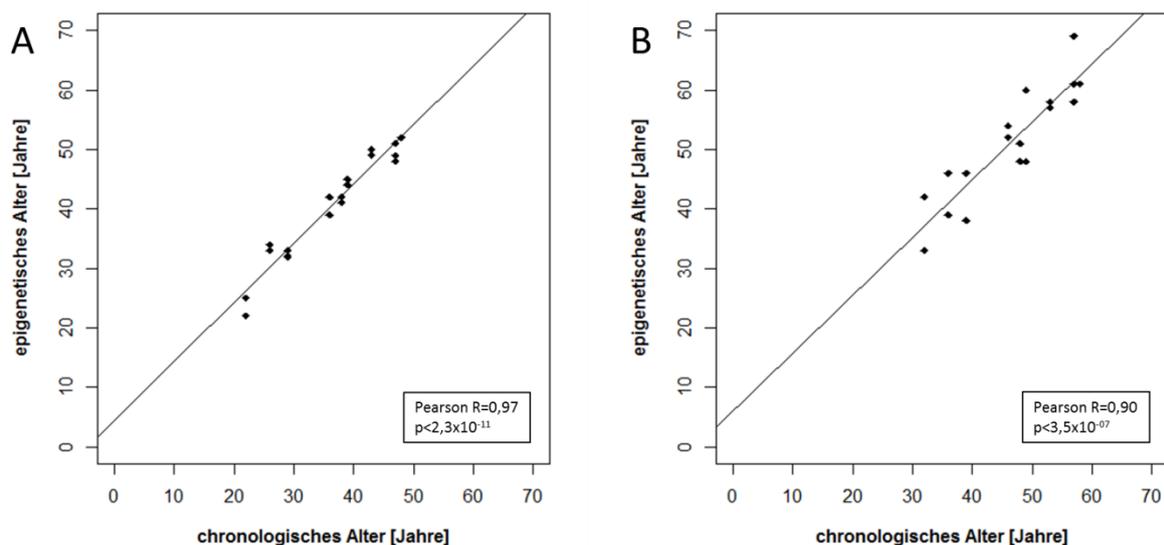


Abbildung 22: Korrelation zwischen dem chronologischen und epigenetischen Alter basierend auf dem Algorithmus von Horvath [150] in neun monozygoten Zwillingspaaren zu zwei verschiedenen Zeitpunkten. Jeder Datenpunkt stellt eine analysierte DNA-Probe dar. Auf der x-Achse ist das chronologische Alter, auf der y-Achse das Alter aufgrund der 353 CpG-Signatur nach Horvath jeweils in Jahren angegeben. **A.** Die berechnete Korrelation betrug für die Zwillinge zu Zeitpunkt 1 (GEMINAKAR) $R=0,97$ ($p<2,3 \times 10^{-11}$; t-Test) mit einem Standardfehler von 2,2 Jahren. **B.** Für die Zwillinge zu Zeitpunkt 2 (SAME) konnte eine Korrelation von $R=0,90$ ($p<3,5 \times 10^{-07}$; t-Test) mit einem Standardfehler von 4,3 Jahren ermittelt werden.

4.2.3.2 Eigenschaften der altersabhängig differentiell methylierten CpG-Regionen

Anreicherungen der altersassoziierten Loci in bestimmten regulatorischen Elementen

Die 111 altersassoziiert differentiell methylierten CpG-Loci aus der in dieser Arbeit ermittelten Signatur fanden sich in DHS ($OR=4,31$; $p<0,0001$; χ^2 Test), DMR ($OR=2,16$; $p<0,05$; χ^2 Test) und RDMR ($OR=3,36$; $p<0,001$; χ^2 Test) jeweils signifikant angereichert vor. Dies konnte auch bei ausschließlicher Berücksichtigung der mit zunehmendem Alter hypermethylierten Loci ($n=91$) festgestellt werden: DHS ($OR=5,08$; $p<0,0001$; χ^2 Test), DMR ($OR=2,09$; $p=0,05$; χ^2 Test) und RDMR ($OR=3,67$; $p<0,001$; χ^2 Test).

Anreicherung der altersassoziierten Loci im 1. Exon

Eine signifikante Anreicherung der altersabhängigen Loci konnte im 1. Exon (OR=2,22; $p < 0,01$; χ^2 Test) ermittelt werden, was auch nur für die mit zunehmendem Alter hypermethylierten CpGs gilt (OR=2,50; $p < 0,005$; χ^2 Test).

Abreicherung der altersassoziierten Loci in Promotor-assoziierten Regionen

Promotor-assoziierte Regionen waren in der Gruppe der altersabhängigen differentiell methylierten CpG-Positionen signifikant abgereichert (OR=0,32; $p = 0,001$; χ^2 Test), was auch nur für die hypermethylierten Positionen gilt (OR=0,29; $p < 0,005$; χ^2 Test).

Anreicherung der altersassoziierten Loci in CpG-Inseln bei simultaner Abreicherung in S-Shore Regionen

Eine signifikante Anreicherung hinsichtlich der 111 differentiell methylierten CpGs konnte in CpG-Inseln beobachtet werden (OR=5,64; $p < 0,0001$; χ^2 Test). Auffallend war hier, dass alle CpGs, die in CpG-Inseln lagen ($n = 80/111$), innerhalb der mit zunehmendem Alter hypermethylierten Loci ($n = 91$) lagen (OR=15,88; $p < 0,0001$; χ^2 Test). Weiterhin wurde eine signifikante Abreicherung in S-Shore Regionen (OR=0,42; $p < 0,05$; χ^2 Test) gefunden, die auch hier überwiegend den hypermethylierten Loci zuzuschreiben war (OR=0,30; $p < 0,05$; χ^2 Test).

4.2.3.3 GO-Analyse der 63 altersabhängig methylierten Genorte

Ziel war es, über eine GO-Abfrage der korrespondierenden altersabhängig methylierten Gene signifikante Anreicherungen definierter funktioneller GO-Kategorien innerhalb des abzufragenden Gensets gegenüber der Referenzgen-Liste zu ermitteln. Die Referenzgen-Liste bestand hierbei aus den auf dem HM450K BC abgebildeten nach Qualitätskontrolle gefilterten Genen ($n = 19708$). Eine GO-Analyse lieferte unter der Kategorie des „*biological process*“ für die 63 altersabhängig methylierten Gene signifikant angereicherte GO-Kategorien, die in Abbildung 23 graphisch dargestellt sind. Wurde diese Analyse separat für die stärker ($n = 52$) bzw. schwächer methylierten ($n = 11$) Gene durchgeführt, konnte beobachtet werden, dass die GO-Analyse hauptsächlich von den stärker methylierten Genen beeinflusst wurde. So brachte die Analyse für die mit zunehmendem Alter ausschließlich schwächer methylierten Gene keine angereicherten Prozesse hervor.

Ergebnisse

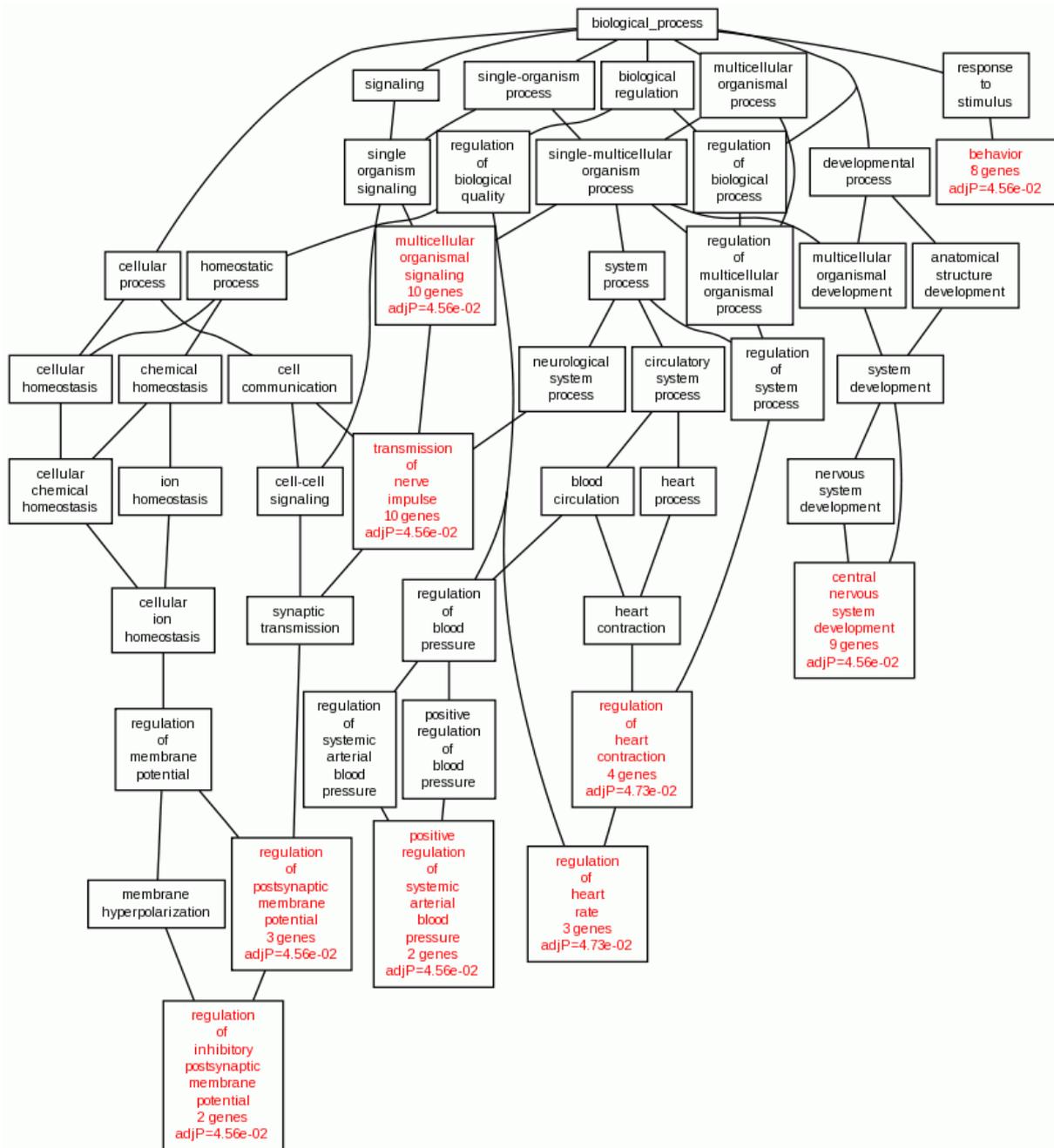


Abbildung 23: Anreicherungsanalyse für Genontologie-Kategorien. Für multiples Testen (Methode nach Benjamini und Hochberg) korrigierte signifikante GO-Kategorien sind in hellrot hinterlegt ($q < 0,05$).

Die signifikant angereicherten Prozesse waren hauptsächlich Unterkategorien von Zellkommunikation, Regulation des Blutdrucks und Herzkontraktion. Ferner konnten angereicherte Prozesse in der Entwicklung des Nervensystems sowie im Verhalten als Reaktion auf einen Stimulus beobachtet werden. Die angereicherten GO-Prozesse sind in nachstehender Tabelle aufgelistet.

Ergebnisse

Tabelle 5: GO-Analyse der altersabhängig differentiell methylierten Gene. Die identifizierten Prozesse sind entsprechend ihrer Signifikanz angeordnet.

GO-Kategorie	beobachtet	erwartet	OR	p-Wert	q-Wert
<i>biological process</i>					
Signalweiterleitung von Mehrzellern	10	2,74	3,65	$3,00 \times 10^{-04}$	0,046
Übertragung eines Nervenimpulses	10	2,68	3,73	$3,00 \times 10^{-04}$	0,046
Regulation des postsynaptischen Membranpotentials	3	0,14	20,79	$4,00 \times 10^{-04}$	0,046
Verhalten	8	1,83	4,37	$4,00 \times 10^{-04}$	0,046
Regulation des inhibitorischen postsynaptischen Membranpotentials	2	0,03	58,52	$5,00 \times 10^{-04}$	0,046
Positive Regulation des systemischen arteriellen Blutdrucks	2	0,04	52,66	$6,00 \times 10^{-04}$	0,046
ZNS Entwicklung	9	2,46	3,66	$6,00 \times 10^{-04}$	0,046
Regulation der Herzkontraktion	4	0,41	9,66	$8,00 \times 10^{-04}$	0,047
Regulation der Herzfrequenz	3	0,18	16,46	$8,00 \times 10^{-04}$	0,047
<i>molecular function</i>	-	-	-	-	-
<i>cellular component</i>	-	-	-	-	-

GO: Genontologie. OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

4.2.3.4 Erweiterte Analyse der 63 altersabhängig methylierten Genorte mittels WebGestalt

Für eine Anreicherungsanalyse krankheitsassoziiertes Genes wurde ebenfalls WebGestalt genutzt, das hierfür eine Abfrage über das GLAD4U Tool nutzt [206], welches für viele Krankheiten spezifische Genlisten aus in PubMed hinterlegten Publikationen erstellt. Somit konnten über die 63 altersabhängig methylierten Gene zahlreiche signifikante Anreicherungen für diverse Krankheitsbilder ermittelt werden (siehe Tabelle A3 im Anhang). Unter ihnen waren auch Erkrankungen, deren Auftreten mit einem erhöhten Risiko im höheren Lebensalter assoziiert sein könnte.

In einer anschließenden KEGG-Analyse, mit der Signal- bzw. Stoffwechselwege identifiziert werden können, wurden von altersabhängiger Methylierung betroffene KEGG-Signalwege mit einer neuroaktiven Ligand-Rezeptor-Interaktion (OR=7,38; $q < 0,005$), dem Alanin-, Aspartat- und Glutamat-Metabolismus (OR=20,07; $q < 0,05$), des „Altersdiabetes“ Diabetes

Ergebnisse

mellitus Typ II (OR=13,52; $q < 0,05$), dem Kalzium-Signalweg (OR=5,62; $q < 0,05$) sowie der dilatativen Kardiomyopathie (OR=7,23; $q < 0,05$) assoziiert vorgefunden.

Eine Analyse hinsichtlich angereicherter Transkriptionsfaktor-Zielgene ergab eine signifikante Anreicherung von Transkriptionsfaktorbindestellen für insgesamt vier verschiedene TFs, von denen drei TFs Funktionen im Nervensystem erfüllten (siehe Tabelle 6). Zu diesen gehörten der *neuron-restrictive silencer factor* NRSF, der für die Repression neuronaler Gene in nicht-neuronalen Zellen verantwortlich ist [207], der transkriptionelle Repressor RP58, der als Schlüssel-Regulator der neuronalen Wanderung dient [208], und das LIM homeobox 3 Gen *LHX3*, das für die Hypophysen-Entwicklung und für die Spezifikation der Motoneurone benötigt wird [209–211]. Alle drei TFs kontrollieren weiterhin Gene, unter deren Beteiligung bereits Erkrankungen des Gehirns beschrieben wurden [208,212,213].

Tabelle 6: Transkriptionsfaktorbindestellen-Analyse der identifizierten altersabhängig differentiell methylierten Gene. Die assoziierten Transkriptionsfaktoren sind entsprechend ihrer Signifikanz angeordnet.

Transkriptionsfaktor	beobachtet	erwartet	OR	p-Wert	q-Wert
hsa_V\$NRSF_01	6	0,32	18,79	$7,75 \times 10^{-07}$	$2,00 \times 10^{-04}$
hsa_V\$SP3_Q3	7	0,76	9,27	$1,01 \times 10^{-05}$	0,001
hsa_V\$RP58_01	6	0,64	9,35	$4,33 \times 10^{-05}$	0,004
hsa_YTAATTAA_V\$LHX3_01	5	0,59	8,42	$3,00 \times 10^{-04}$	0,018

OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

4.2.4 Auswirkung des *Body Mass Index* (BMI) auf die genomweite DNA-Methylierung

Für die Untersuchung des Effekts des BMIs auf die genomweite DNA-Methylierung in Zellen des peripheren Blutes konnten nur die Zwillinge der SAME-Studie ($n=53$) berücksichtigt werden, da für die Zwillinge der GEMINAKAR-Studie die Informationen zum individuellen BMI nicht vorlagen. Für die Analyse des Einflusses eines hohen BMIs auf die DNA-Methylierung wurden Zwillinge des äußeren BMI-Extrems gewählt, was dem Vergleich des Methylierungsmusters von normalgewichtigen Zwillingen (BMI von 18,5–24,9) mit adipösen Zwillingen (BMI ≥ 30) entsprach. Die allgemeinen Charakteristika dieser 34 Zwillinge (vgl. Tabelle 3) sind in nachstehender Tabelle detailliert aufgeführt.

Ergebnisse

Tabelle 7: Charakteristika der 34 in die BMI-Analyse eingeschlossenen Zwillinge.

Merkmal	adipöse Zwillinge (n=5)	normalgewichtige Zwillinge (n=29)	p-Wert
Geschlecht (männlich/weiblich)	2/3	11/18	
Alter [Jahre]	43,2±5,7 (36–48)	45,7±9,5 (32–63)	0,57
BMI [kg/m ²]	32,4±3,1 (30,24–37,50)	22,1±1,5 (19,13–24,51)	<0,0001

Die jeweiligen Parameter sind als Mittelwert ± SD (Minimum–Maximum) angegeben. p-Wert: t-Test für unabhängige Stichproben.

Weiterhin erlaubte diese Analyse eine visuelle Darstellung der dazwischen liegenden Gewichtsklasse, dem übergewichtigen Phänotyp, für die zwischen normalgewichtigen und adipösen Phänotyp identifizierten differentiell methylierten Loci, ohne Einfluss der Daten in die statistische Auswertung. Für die durchgeführten Analysen konnten fünf adipöse, 29 normalgewichtige sowie 19 übergewichtige Zwillinge berücksichtigt werden.

4.2.4.1 Identifizierung der zwischen dem adipösen und normalgewichtigen Phänotyp differentiell methylierter Genorte über die FDR

Über den Vergleich des globalen DNA-Methylierungsmusters von adipösen (n=5) und normalgewichtigen (n=29) Zwillingen konnten durch Setzen der FDR $q < 0,05$ und des Varianzfilters $\sigma/\sigma_{\max} > 0,02$ drei differentiell methylierte CpG-Positionen identifiziert werden (siehe Tabelle 8 und Abbildung 24), die im adipösen im Vergleich zum normalgewichtigen Phänotyp eine Hypomethylierung zeigten und insgesamt zwei Genen zugeordnet werden konnten.

Tabelle 8: Liste der mittels Infinium HumanMethylation450 BeadChip identifizierten differentiell methylierten CpG-Regionen ($q < 0,05$; t-Test; $\sigma / \sigma_{\max} > 0,02$) im Vergleich von adipösen ($n=5$) und normalgewichtigen ($n=29$) Zwillingen. Die Loci sind nach Stärke der mittleren Methylierungsdifferenz (adipöse – normalgewichtige Zwillinge) geordnet.

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (adipös – normal)	Mittlere Methylierung [β] adipös	Mittlere Methylierung [β] normal	q-Wert
cg03849834	6	41195891	TREML4	TSS200	N/A	-0,179	0,783	0,962	0,044
cg10351692	7	75053196	POM121C	Body	S Shore	-0,052	0,909	0,961	0,044
cg23218473	3	176921953	N/A	N/A	N/A	-0,024	0,944	0,968	0,044

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position des analysierten CpG-Locus nach UCSC Genome Browser. q-Wert: p-Wert korrigiert nach Benjamini-Hochberg. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

Ergebnisse

Die größte Differenz in der Methylierung wies die CpG-Position cg03849834 mit 17,9% auf und konnte der TSS200 im Gen *TREML4* (*triggering receptor expressed on myeloid cells-like 4*) zugeordnet werden. Eine Differenz von 5,2% im Methylierungsgrad wurde im Genkörper des *POM121C*-Gens (cg10351692) nachgewiesen. Auf den CpG-Locus cg23218473 fiel ein Methylierungsunterschied von 2,4%. Dieser Locus konnte keinem Gen zugeordnet werden und befand sich auf Chromosom 3 (3q26.32) in der *Open Sea* (intergene Region).

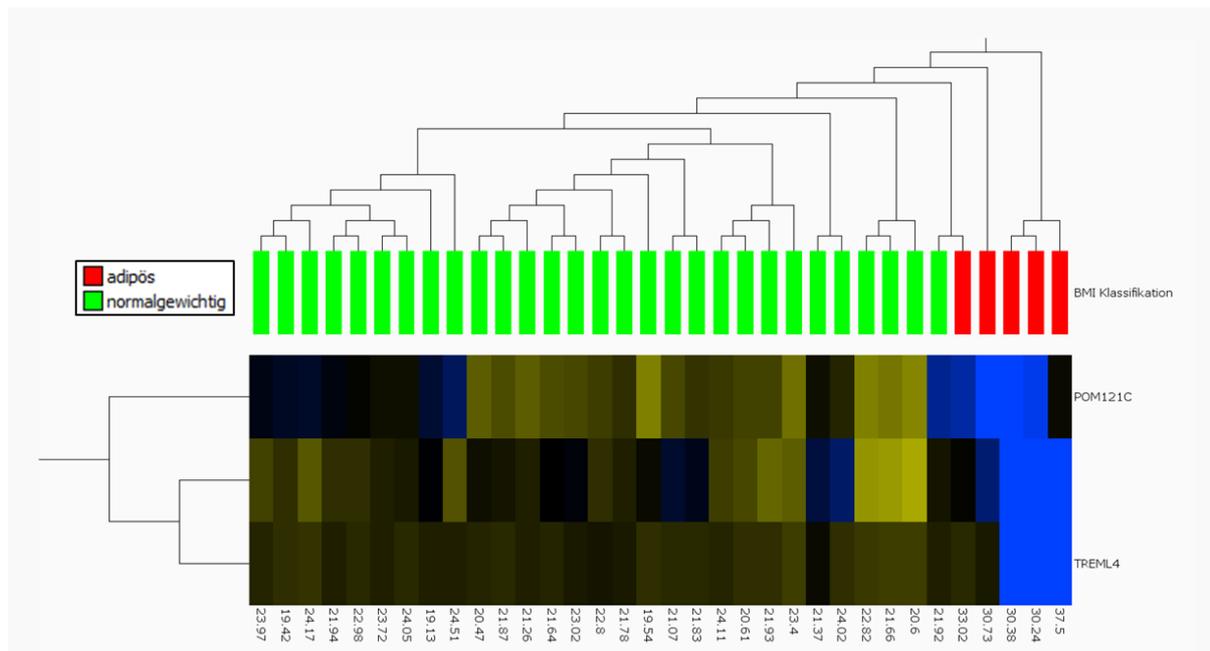


Abbildung 24: Differentielle Methylierungsanalyse von nach BMI definierten normalgewichtigen vs. adipösen Zwillingen. Eine hierarchische Clusteranalyse (*average linkage*) der drei differentiell methylierten CpGs (t-Test, $q < 0,05$; $\sigma/\sigma_{\max} > 0,02$) zeigte in den adipösen Zwillingen (rot) tendenziell eine Hypomethylierung im Vergleich zu den normalgewichtigen Zwillingen (grün). Spalten repräsentieren Proben, Reihen hingegen CpG-Loci. Innerhalb der Heatmap gibt ein blauer Farbcode eine geringe Methylierung an, während gelb eine hohe Methylierung widerspiegelt. Unterhalb der Heatmap ist der BMI der einzelnen Zwillinge angegeben, rechts neben der Heatmap ist, falls verfügbar, die Genzuordnung der jeweiligen CpGs zu finden. Zur Visualisierung wurden die Mittelwerte sowie die Varianz für jeden CpG-Locus gegen 0 bzw. 1 normalisiert (mean=0, var=1).

Um die BMI-Abhängigkeit des Methylierungsmusters der drei CpG-Loci zu unterstützen, wurde in der zur Heatmap (siehe Abbildung 24) korrespondierenden PCA (Abbildung 25A) das Methylierungsmuster der 19 übergewichtigen Zwillinge dargestellt, ohne dass die Daten in die statistische Analyse eingeflossen sind (Abbildung 25B).

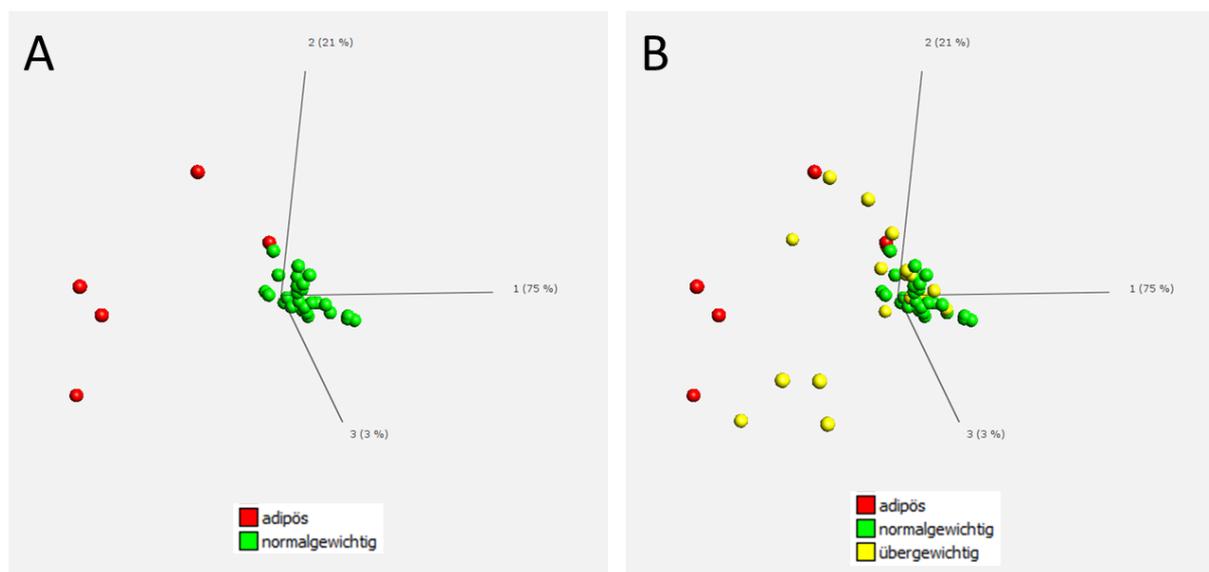


Abbildung 25: PCA der differentiellen Methylierungsanalyse von nach BMI definierten normalgewichtigen vs. adipösen Zwillingen. (A) zeigt die PCA der drei differentiell methylierten CpGs ($q < 0,05$; t-Test; $\sigma/\sigma_{\max} > 0,02$) zwischen den adipösen Zwillingen (rot) und den normalgewichtigen Zwillingen (grün). In **(B)** sind zusätzlich in der in (A) generierten PCA simultan die übergewichtigen Zwillinge (gelb) eingeblendet ohne in die PCA selbst einzugehen.

Die übergewichtigen Zwillinge gruppierten sich aufgrund ihres Methylierungsmusters in den drei CpG-Loci überwiegend zwischen den anderen beiden Phänotypen, was sich mit dem zwischen Normalgewichtigen und Adipösen liegenden BMI vereinbaren ließ. Dieses unterstützt eine hier identifizierte mögliche Assoziation zwischen BMI und DNA-Methylierung an den CpG-Loci cg03849834, cg10351692 und cg23218473.

4.2.4.2 Identifizierung differentiell methylierter Genorte über p-Wert und $\Delta\beta$

In einem weniger stringenten Ansatz, der auch in der Literatur beschrieben wurde [214,215], wurde die differentielle Methylierung mit dem nicht für multiples Testen korrigierten p-Wert ($p < 0,05$) und dem korrespondierenden $\Delta\beta$ -Wert, der der mittleren Methylierungsdifferenz des abzufragenden Locus zwei zu vergleichender Gruppen entspricht, berechnet. Unter der Berücksichtigung von $p < 0,0003$ und einem $\Delta\beta > 0,05$ bzw. $< -0,05$ lieferte dieser Ansatz 80 differentiell methylierte CpG-Stellen, die mit 43 Genorten assoziiert werden konnten. Die entsprechende hierarchische Clusteranalyse ist in Abbildung 26 dargestellt, wobei die zwei Hauptarme von den beiden zu vergleichenden Gruppen repräsentiert werden. 75/80 CpGs (94%) wiesen in den adipösen Zwillingen im Vergleich zu den Normalgewichtigen eine Hypomethylierung und entsprechend 5/80 Loci (6%) eine Hypermethylierung auf. Der p-Wert variierte in diesen 80 Positionen von $p < 0,0003$ bis hin zu $p < 3,47 \times 10^{-07}$. Eine Auflistung

Ergebnisse

der individuellen CpG-Loci befindet sich in Tabelle A4 im Anhang. Unter den 43 identifizierten Genen befanden sich auch die aus dem ersten Ansatz identifizierten Gene (*TREML4* und *POM121C*). Des Weiteren wurden zwei Gene identifiziert, die bereits mit Adipositas in Verbindung gebracht werden konnten: Der A1-Adenosinrezeptor *ADORA1* [216,217] und die Lysophosphatidylglycerol Acyltransferase 1 *LPGAT1* [218]. Ferner konnten den Genen *CCDC144NL*, *EGFLAM*, *TMEM232*, *DPCR1*, *PTPRN2* und *CNTNAP1* sowie den Regionen 2p25.3 und 10q26.3 jeweils mindestens zwei differentiell methylierte CpG-Loci zugeordnet werden.

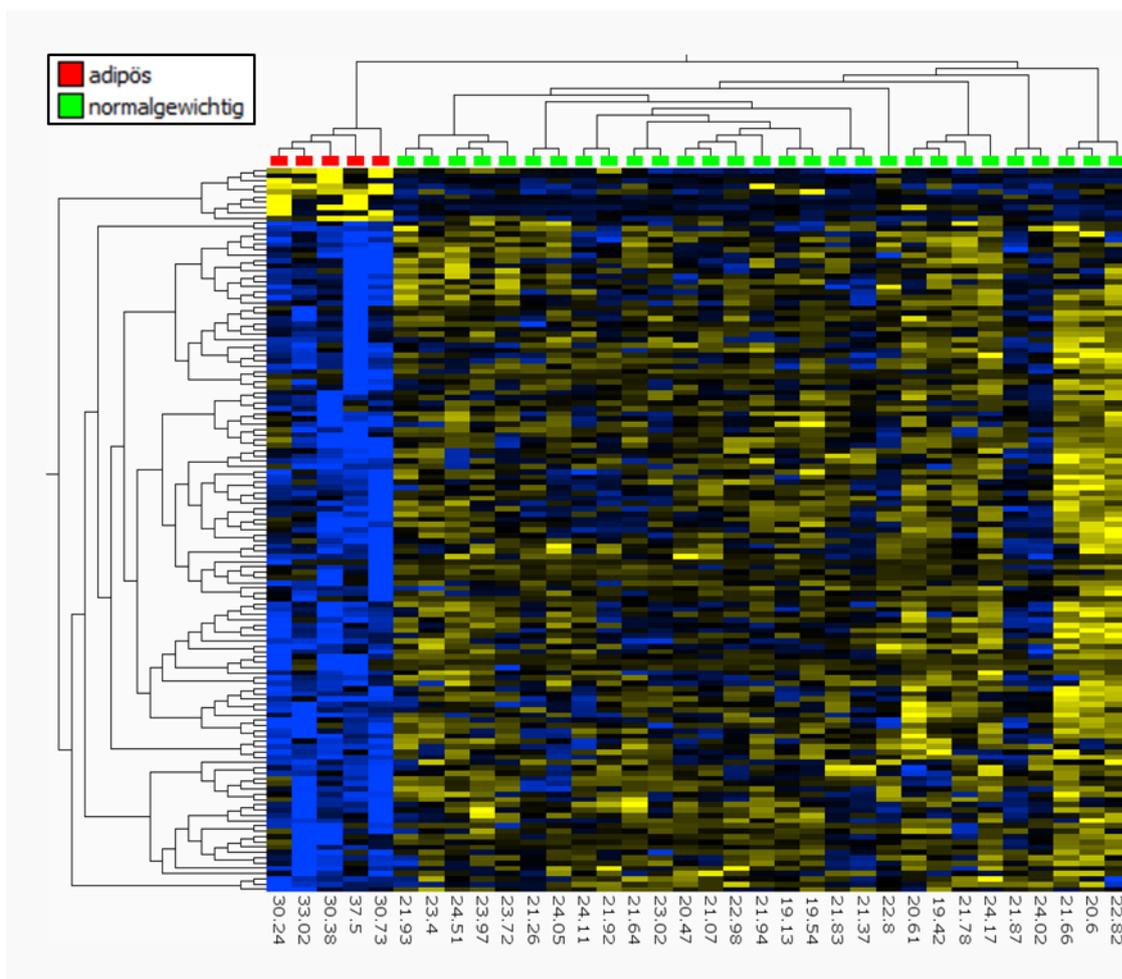


Abbildung 26: Differentielle Methylierungsanalyse von nach BMI definierten normalgewichtigen vs. adipösen Zwillingen. Eine hierarchische Clusteranalyse (*average linkage*) der 80 differentiell methylierten CpGs ($p < 0,0003$; t-Test; $|\Delta\beta| > 0,05$) zeigt, dass in den adipösen Zwillingen (rot) eine Hypomethylierung (75/80 CpGs) im Vergleich zu den normalgewichtigen Zwillingen (grün) dominiert. Spalten repräsentieren Proben, Reihen hingegen CpG-Loci. Innerhalb der Heatmap gibt ein blauer Farbcode eine geringe DNA-Methylierung an, während ein gelber Farbcode eine hohe DNA-Methylierung widerspiegelt. Unterhalb der Heatmap ist der BMI der einzelnen Zwillinge angegeben. Zur Visualisierung wurden die Mittelwerte sowie die Varianz für jeden CpG-Locus gegen 0 bzw. 1 normalisiert (mean=0, var=1).

Ergebnisse

Auch in diesem Ansatz sollte die Abhängigkeit zwischen dem DNA-Methylierungsmuster der 80 CpGs und dem BMI durch die mittlere Gruppe, die die dazwischenliegende Gewichtsklasse repräsentierte, unterstützt werden. In der zur Heatmap (Abbildung 26) korrespondierenden PCA (Abbildung 27A) wurde daher das Methylierungsmuster der 19 übergewichtigen Zwillinge für diese CpGs dargestellt, ohne die Daten in die statistische Analyse einzubeziehen (Abbildung 27B).

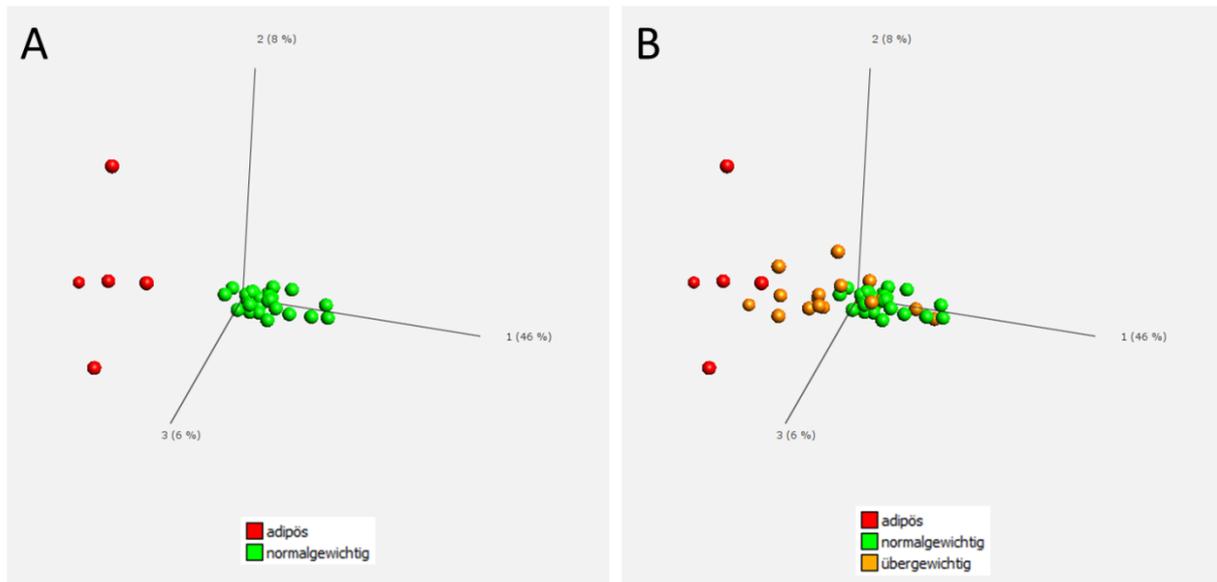


Abbildung 27: PCA der differentiellen Methylierungsanalyse von nach BMI definierten normalgewichtigen vs. adipösen Zwillingen. (A) zeigt die PCA der zwischen den adipösen Zwillingen (rot) und den normalgewichtigen Zwillingen (grün) 80 differentiell methylierten CpGs ($p < 0,0003$; t-Test; $|\Delta\beta| > 0,05$). In (B) sind zusätzlich in der in (A) generierten PCA simultan die übergewichtigen Zwillinge (orange) eingeblendet.

Die übergewichtigen Zwillinge ordneten sich basierend auf ihrem DNA-Methylierungsmuster auch in diesem zweiten Ansatz überwiegend zwischen der Gruppe der adipösen und der Gruppe der normalgewichtigen Zwillinge in der Hauptkomponente 1 der PCA an. Somit könnte auch hier ein Zusammenhang zwischen dem Grad der DNA-Methylierung an den 80 identifizierten Loci und dem BMI bestehen.

4.2.4.3 Eigenschaften der 80 zwischen den adipösen und normalgewichtigen Zwillingen differentiell methylierten CpG-Regionen

Hinsichtlich einer Anreicherungsanalyse bezüglich bestimmter Eigenschaften und Merkmale, wie regulatorische Elemente, Verteilung im Gen-Zusammenhang, assoziierte DNA-Regionen oder CpG-Lokalisation, konnten in den über die FDR ermittelten differentiell methylierten

Ergebnisse

CpG-Regionen aufgrund der geringen Ausbeute (n=3 Loci) keine Ergebnisse erzielt werden. Daher wurden im Folgenden nur die über dem Alternativ-Ansatz bestimmten Loci (n=80) analysiert.

Keine An-/Abreicherung der BMI-abhängigen Loci innerhalb bestimmter regulatorischer Elemente

Unter den mit hohem BMI assoziierten identifizierten CpG-Loci konnten keine signifikanten An- oder Abreicherungen der untersuchten regulatorischen Elemente (Enhancer, DHS und DMR) gefunden werden ($p > 0,05$; χ^2 Test).

Keine An-/Abreicherung der BMI-abhängigen Loci im Gen-Zusammenhang

Beim Vergleich der Verteilung der aufgrund des BMIs differentiell methylierten CpGs im Gen-Zusammenhang (5'-UTR, TSS1500/200, Exon1, Genkörper oder 3'-UTR) konnte ebenfalls keine signifikante An- bzw. Abreicherung festgestellt werden ($p > 0,05$; χ^2 Test).

Abreicherung der BMI-abhängigen Loci in Promotor-assoziierten Regionen

Beim Vergleich der CpG-Loci in assoziierten DNA-Regionen (Gen-, Nicht-Gen- oder Promotor-assoziiert) konnte eine signifikante Abreicherung der 80 CpGs in Promotor-assoziierten Regionen (OR=0,11; $p=0,0001$; χ^2 Test) identifiziert werden. Diese konnte ebenfalls beobachtet werden, nachdem die Analyse auf die im adipösen Phänotyp hypomethylierten CpGs (n=75) beschränkt wurde (OR=0,11; $p=0,0002$; χ^2 Test). Dies besagt, dass Promotor-assoziierte Regionen signifikant weniger von einer BMI-abhängigen differentiellen Methylierung betroffen sind als statistisch zu erwarten gewesen wäre.

Keine An-/Abreicherung der BMI-abhängigen Loci hinsichtlich CpG-Lokalisation

Bezüglich der CpG-Lokalisation (N- und S-Shelf, Insel oder N- und S-Shore) konnten keine signifikanten Ab- oder Anreicherungen hinsichtlich der 80 differentiell methylierten CpGs ermittelt werden ($p > 0,05$; χ^2 Test). Dies zeigt, dass die hier identifizierte BMI-assoziierte DNA-Methylierung nicht auf spezielle Bereiche in CpG-Inseln und assoziierten Regionen beschränkt ist.

4.2.4.4 GO-Analyse der 43 mit hohem BMI korrelierenden differentiell methylierten Genorte

Während keine signifikante Anreicherung der aufgrund eines hohen BMIs detektierten differentiell methylierten Gene für den „*biological process*“ im peripheren Blut ermittelt werden konnte, waren diese für die „*molecular function*“ „Dynein Bindung“ (OR=94,21; $q < 0,05$) sowie für die „*cellular component*“, die überwiegend Kompartimente der Nervenzelle betraf (Tabelle 9), signifikant angereichert ($q < 0,05$).

Tabelle 9: GO-Analyse der 43 mit hohem BMI assoziierten differentiell methylierten Gene. Die identifizierten Prozesse sind entsprechend ihrer Signifikanz angeordnet.

GO-Kategorie	beobachtet	erwartet	OR	p-Wert	q-Wert
<i>biological process</i>	-	-	-	-	-
<i>molecular function</i>					
Dynein Bindung	2	0,02	94,21	$2,00 \times 10^{-04}$	0,014
<i>cellular component</i>					
Bestandteil des Axons	4	0,29	14,01	$2,00 \times 10^{-04}$	0,015
Axon	5	0,62	8,12	$3,00 \times 10^{-04}$	0,015
Zytoplasmatische Mikrotubuli	2	0,07	30,67	0,002	0,046
Perinukleäre Region des Cytoplasmas	5	1,00	4,99	0,003	0,046
Synapse	5	1,03	4,85	0,003	0,046
Neuronaler Zellkörper	4	0,63	6,40	0,003	0,046
Endknöpfchen	2	0,09	22,24	0,004	0,046
Haupt-Axon	2	0,09	22,24	0,004	0,046
Zellkörper	4	0,67	5,97	0,004	0,049

GO: Genontologie. OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

An den Kompartimenten der Nervenzelle waren hierbei insgesamt 8/43 Gene involviert: *ADORA1*, *CNTNAP1*, *COBL*, *EGFLAM*, *FABP7*, *PSEN1*, *PTPRN2* und *SYN2*.

Die GO-Analyse wurde in einem weiteren Ansatz zusätzlich separat für die im adipösen Phänotyp schwächer ($n=39$) und stärker methylierten ($n=4$) Gene durchgeführt, wobei die oben durchgeführte GO-Analyse (Tabelle 9) den Resultaten der Analyse der mit hohem BMI schwächer methylierten Gene entsprach.

4.2.4.5 Erweiterte Analyse der zwischen normalgewichtigen und fettleibigen Zwillingen differentiell methylierten Genorte (n=43) mittels WebGestalt

Innerhalb der mit hohem BMI assoziierten differentiell methylierten Gene konnten in einer *KEGG*-Analyse unter der Beteiligung von jeweils zwei Genen Anreicherungen für „Purin-Stoffwechsel“ (OR=5,71; $q < 0,05$; χ^2 Test), „Alzheimer“ (OR=5,71; $q < 0,05$; χ^2 Test) sowie „dilatative Kardiomyopathie“ (OR=10,16; $q < 0,05$; χ^2 Test) ermittelt werden. Demnach konnte im Rahmen dieser Analyse ein hoher BMI mit Methylierungsveränderungen an Genen assoziiert werden, die sowohl mit dem Purin-Stoffwechsel als auch mit Alzheimer und den damit verbundenen kognitiven Störungen sowie einer Herzmuskelerkrankung in Verbindung stehen könnten.

4.2.5 Auswirkung des Rauchens auf die genomweite DNA-Methylierung

Zur Untersuchung des Effekts des Zigaretten-Rauchens auf die genomweite DNA-Methylierung konnten nur 51/53 Zwillinge der SAME-Studie berücksichtigt werden, da von zwei Zwillingen der SAME-Studie die Angaben zum Rauchverhalten fehlten und ferner die Informationen über den Rauchstatus der Zwillinge der GEMINAKAR-Studie zum Zeitpunkt der Analyse nicht vorlagen. Der Rauchstatus wurde gemäß Fragebogen in die vier Kategorien „Raucher“, „Gelegenheitsraucher“, „ehemalige Raucher“ und „Nichtraucher“ eingeteilt, wobei hier die DNA-Methylierung der beiden Extreme „Raucher“ (n=12) und „Nichtraucher“ (n=25) untersucht wurde (vgl. Tabelle 3). Im Anschluss an die Analyse wurde das Methylierungsmuster der zwischen Rauchern und Nichtrauchern ermittelten differentiell methylierten CpG-Loci in den Gelegenheitsrauchern (n=4) und ehemaligen Rauchern (n=10) visualisiert.

4.2.5.1 Identifizierung der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte über die FDR

Der Vergleich des globalen Methylierungsmusters von Rauchern (n=12) und Nichtrauchern (n=25) lieferte mit den Einstellungen der FDR $q < 0,05$ und des Varianzfilters $\sigma/\sigma_{\max} > 0,08$ 22 CpG-Positionen, die insgesamt 14 Genen zugeordnet werden konnten und allesamt bei den Rauchern im Vergleich zu den Nichtrauchern eine Hypomethylierung aufwiesen (Tabelle 10).

Tabelle 10: Liste der mittels Infinium HumanMethylation450 BeadChip identifizierten differentiell methylierten CpG-Regionen ($q < 0,05$; t-Test; $\sigma / \sigma_{\max} > 0,08$) im Vergleich von Rauchern ($n=12$) und Nichtrauchern ($n=25$). Die Loci sind nach Stärke der mittleren Methylierungsdifferenz (Raucher – Nichtraucher) geordnet.

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert
cg05575921	5	373378	AHRR	Body	N Shore	-0,258	0,549	0,807	$2,11 \times 10^{-11}$
cg21566642	2	233284661	N/A	N/A	Island	-0,186	0,294	0,480	$2,25 \times 10^{-05}$
cg19336497	11	14380999	RRAS2	TSS1500	S Shore	-0,177	0,556	0,733	0,0490
cg01940273	2	233284934	N/A	N/A	Island	-0,156	0,544	0,704	$5,63 \times 10^{-07}$
cg05951221	2	233284402	N/A	N/A	Island	-0,154	0,305	0,459	$3,24 \times 10^{-04}$
cg06126421	6	30720080	N/A	N/A	N/A	-0,146	0,635	0,780	0,002
cg03636183	19	17000585	F2RL3	Body	N Shore	-0,143	0,623	0,766	$9,06 \times 10^{-09}$
cg14671384	12	47219920	SLC38A4	TSS200	N/A	-0,128	0,429	0,557	0,037
cg14817490	5	392920	AHRR	Body	N/A	-0,084	0,201	0,285	0,005
cg26703534	5	377358	AHRR	Body	S Shelf	-0,078	0,629	0,707	$3,42 \times 10^{-04}$
cg12207248	3	59038598	N/A	N/A	S Shelf	-0,076	0,396	0,471	0,036
cg25648203	5	395444	AHRR	Body	N/A	-0,071	0,790	0,861	0,027
cg23161492	15	90357202	ANPEP	5'UTR	N Shore	-0,068	0,214	0,282	0,044
cg13972460	20	1165703	C20orf46	TSS1500	Island	-0,062	0,184	0,245	0,040
cg11146034	12	96590670	ELK3	5'UTR	S Shore	-0,061	0,335	0,396	0,050
cg01321189	12	32259338	BICD1	TSS1500	Island	-0,060	0,136	0,196	0,002
cg17268801	19	24269968	ZNF254	TSS200	N/A	-0,055	0,301	0,356	0,044
cg06852575	19	22235454	ZNF257	Body	N/A	-0,053	0,264	0,317	0,046
cg08143875	13	113784663	F10	Body	Island	-0,053	0,385	0,437	0,044
cg25940730	4	16228647	TAPT1	TSS1500	Island	-0,051	0,254	0,305	0,050
cg16255816	17	39889811	HAP1	Body	N Shore	-0,047	0,311	0,359	0,027
cg27025079	4	26865310	STM2	Body	S Shelf	-0,046	0,813	0,858	0,044

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position des analysierten CpG-Locus nach UCSC Genome Browser. q-Wert: p-Wert korrigiert nach Benjamini-Hochberg. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

Ergebnisse

Die korrespondierende Heatmap ist in Abbildung 28 dargestellt, in der eine klare Trennung zwischen Rauchern und Nichtrauchern beobachtet werden konnte. Die jeweiligen Methylierungsdifferenzen erstreckten sich dabei von 4,6% bis hin zu 25,8%. Der p-Wert lag zwischen $8,97 \times 10^{-6}$ und $1,72 \times 10^{-16}$, was einem korrespondierenden für multiples Testen korrigierten p-Wert von $q < 0,05$ bis $< 2,11 \times 10^{-11}$ entsprach. Die 22 differentiell methylierten CpGs waren auf insgesamt 12 Chromosomen lokalisiert. Hierbei deckten vier CpGs etwa 22kb des auf Chromosom 5 (5p15.3) lokalisierten etwa 130kb großen *AHRR*-Gens (*aryl hydrocarbon receptor repressor*) ab. Zwei der vier CpGs flankierten dabei eine 584bp große CpG-Insel, wobei cg05575921 in einem N-Shore *downstream* und cg26703534 in einem S-Shelf *upstream* der genannten Insel lagen. Weiterhin konnten drei CpGs in einer 533bp großen Region identifiziert werden, die innerhalb einer 2562bp großen CpG-Insel auf Chromosom 2 (2q37.1) lokalisiert war (<http://genome-euro.ucsc.edu/>), aber keinem Gen zugeordnet werden konnte.

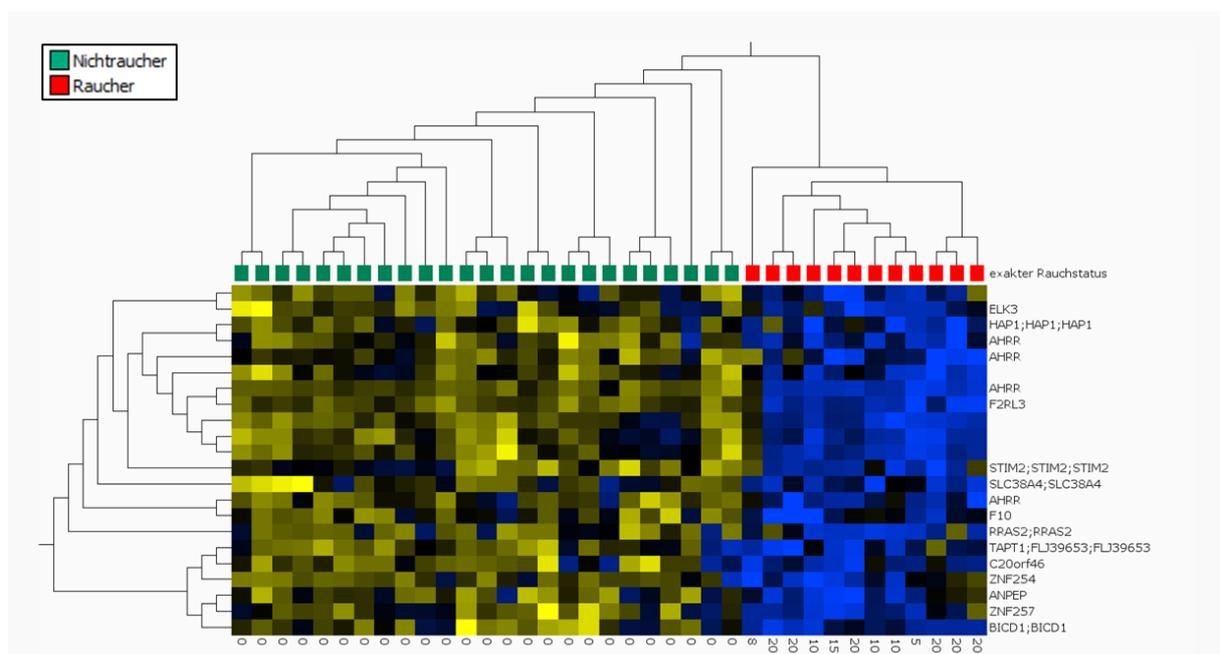


Abbildung 28: Differentielle Methylierungsanalyse von Rauchern (n=12) vs. Nichtrauchern (n=25). Eine hierarchische Clusteranalyse (*average linkage*) der 22 differentiell methylierten CpGs ($q < 0,05$; t-Test; $\sigma/\sigma_{\max} > 0,08$) zeigt für alle Loci eine Hypomethylierung in den Rauchern (rot) im Vergleich zu den Nichtrauchern (grün). Spalten repräsentieren Proben, Reihen hingegen CpG-Loci. Innerhalb der Heatmap gibt ein blauer Farbcode eine geringe Methylierung an, während gelb eine hohe Methylierung widerspiegelt. Auf der x-Achse der Heatmap ist die Anzahl der gerauchten Zigaretten pro Tag angegeben, auf der y-Achse ist, falls verfügbar, die zu jedem CpG-Locus korrespondierende Genzuordnung verzeichnet. Zur Visualisierung wurden die Mittelwerte sowie die Varianz für jeden CpG-Locus gegen 0 bzw. 1 normalisiert (mean=0, var=1).

Ergebnisse

Basierend auf ihrem Methylierungsmuster wurden die Gelegenheitsraucher und ehemaligen Raucher in der zur Heatmap (Abbildung 28) korrespondierenden PCA (Abbildung 29A) in Abbildung 29B dargestellt, wobei die Daten nicht in die statistische Analyse einbezogen wurden.

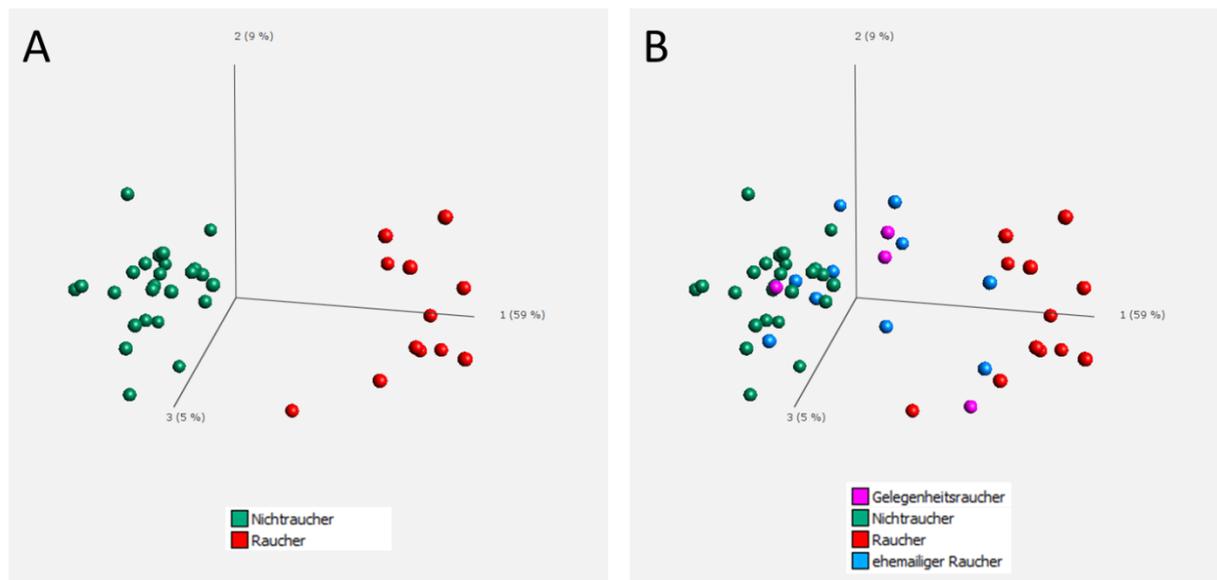


Abbildung 29: PCA der differentiellen Methylierungsanalyse von Rauchern (n=12) vs. Nichtrauchern (n=25). (A) zeigt die PCA der 22 differentiell methylierten CpGs ($q < 0,05$; t-Test; $\sigma/\sigma_{\max} > 0,08$) zwischen den Rauchern (rot) und den Nichtrauchern (grün). In (B) sind zusätzlich in der in (A) generierten PCA simultan die Gelegenheitsraucher (rosa) und die ehemaligen Raucher (blau) eingeblendet.

Hierbei ordneten sich die ehemaligen Raucher sowohl bei den Nichtrauchern als auch bei den Rauchern an, wobei die Tendenz in Richtung der Nichtraucher verlief. Dies deutet darauf hin, dass die durch das Rauchen veränderte DNA-Methylierung an den ausgewählten CpG-Loci nach Aufgabe des Rauchens wieder zum ursprünglichen Methylierungsniveau zurückkehren kann. Die Gelegenheitsraucher ordneten sich ebenfalls sowohl bei den Nichtrauchern als auch bei den Rauchern an, was den Vergleich des Methylierungsmusters mit dem der ehemaligen Raucher erlaubte.

4.2.5.1.1 Eigenschaften der 22 zwischen Rauchern und Nichtrauchern differentiell methylierten CpG-Regionen

Anreicherungen der mit dem Rauchen assoziierten CpG-Loci in DHS und DMR

Innerhalb der Gruppe der regulatorischen Elemente konnte eine signifikante Anreicherung von DHS (OR=4,05; $p < 0,001$; χ^2 Test) und DMR (OR=3,97; $p < 0,05$; χ^2 Test) ermittelt werden.

Somit konnte mit dem stringenten Ansatz gezeigt werden, dass Rauchen eine Abnahme der Methylierung von CpGs in regulatorischen Bereichen begünstigt. Interessanterweise wird für transkriptionell aktive DNA im Allgemeinen im Promotorbereich eine niedrige Methylierung erwartet.

Keine An-/Abreicherungen der mit dem Rauchen assoziierten CpGs im Gen-Zusammenhang

Hinsichtlich der Verteilung innerhalb eines Genes konnten keine statistisch signifikanten An- bzw. Abreicherungen von 5'-UTR, TSS1500/200, Exon1, Genkörper oder 3'-UTR in der Gruppe der mit Rauchen assoziiert differentiell methylierten CpG-Loci ermittelt werden ($p > 0,05$; χ^2 Test).

Anreicherung der mit dem Rauchen assoziierten CpGs in Gen-assoziierten Zelltyp-spezifischen Regionen

Eine signifikante Anreicherung der zwischen Rauchern und Nichtrauchern differentiell methylierten CpGs wurde in mit Gen-assoziierten Zelltyp-spezifischen Regionen (OR=24,33, $p < 0,0001$; χ^2 Test) ermittelt, wobei die zwei beteiligten CpG-Loci in dem *AHRR* Gen lokalisiert waren.

Anreicherung der mit dem Rauchen assoziierten CpGs in S-Shelf Regionen

Eine Anreicherungsanalyse bezüglich der CpG-Lokalisation (N- und S-Shelf, Insel oder N- und S-Shore) hat gezeigt, dass im Vergleich zu den auf dem Array abgebildeten CpG-Positionen die hypomethylierten CpGs im Blut von Rauchern verstärkt in S-Shelf Regionen (OR=3,35; $p < 0,05$; χ^2 Test) anzutreffen waren, in denen eine etwa 3-fache Anreicherung detektiert werden konnte.

4.2.5.1.2 GO-Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte (n=14)

Mittels einer GO-Analyse wurden in der funktionellen Hauptgruppe „*biological process*“ signifikant angereicherte GO-Kategorien ermittelt (Abbildung 30).

Ergebnisse

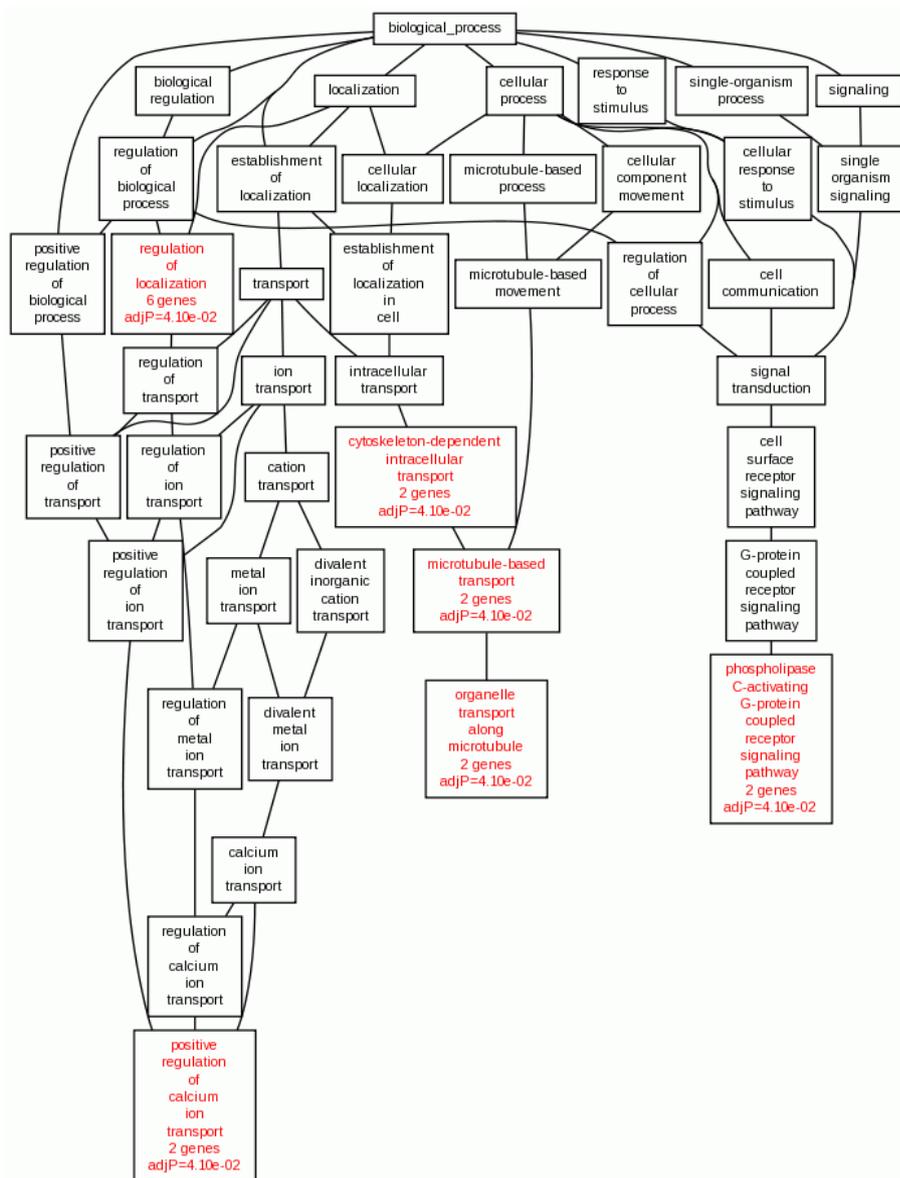


Abbildung 30: Anreicherungsanalyse für Genontologie-Kategorien. Im biologischen Prozess wurden für multiples Testen korrigierte (Methode nach Benjamini und Hochberg) signifikante GO-Kategorien identifiziert, die hier in roter Farbe hinterlegt sind.

Diese konnten den funktionellen Kategorien „G-Protein gekoppelter Rezeptor Signalweg“, „intrazellulärer Transport“ und „Regulation biologischer Prozesse“ zugeordnet werden, wobei die drei immer spezifisch werdenden Unterkategorien des intrazellulären Transports jeweils von den beiden Genen *BICD1* und *HAP1* gebildet wurden. Eine Auflistung der angereicherten GO-Prozesse ist in nachstehender Tabelle hinterlegt.

Ergebnisse

Tabelle 11: GO-Analyse der 14 zwischen Rauchern und Nichtrauchern differentiell methylierten Gene. Die identifizierten Prozesse sind entsprechend ihrer Signifikanz angeordnet.

GO-Kategorie	beobachtet	erwartet	OR	p-Wert	q-Wert
biological process					
Organellentransport entlang der Microtubuli	2	0,03	75,02	$3,00 \times 10^{-04}$	0,041
Regulation der Lokalisation	6	1,24	4,85	$7,00 \times 10^{-04}$	0,041
Mikrotubuli-basierter Transport	2	0,05	39,72	0,001	0,041
Phospholipase C-aktivierender G-Protein gekoppelter Rezeptor Signalweg	2	0,06	36,17	0,001	0,041
Zytoskelet-abhängiger intrazellulärer Transport	2	0,06	36,17	0,001	0,041
Positive Regulation des Kalziumionen Transport	2	0,06	36,17	0,001	0,041
molecular function					
-	-	-	-	-	-
cellular component					
-	-	-	-	-	-

GO: Genontologie. OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

4.2.5.1.3 Erweiterte Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte (n=14) mittels WebGestalt

Unter den mit dem Rauchen assoziierten differentiell methylierten Genen konnte eine Anreicherung von krankheitsassoziierten Genen in „Thrombose“ (OR=23,41; $q < 0,01$) und „Wechselwirkungen zwischen Arzneimitteln“ (OR=7,71; $q < 0,05$) identifiziert werden. Es ist bereits gezeigt, dass Rauchen neben dem Herz-Kreislauf-System [219] auch Blutgefäßveränderungen hinsichtlich einer Gefäßverengung beeinflusst [220]. Die unter den hier mit Thrombose assoziierten differentiell methylierten Gene waren der Koagulationsfaktor X (*F10*) sowie das Koagulationsfaktor II (Thrombin) Rezeptor-ähnliche Gen *F2RL3*. Die beiden besagten Gene waren ebenfalls mit den Wechselwirkungen zwischen Arzneimitteln assoziiert.

4.2.5.2 Identifizierung differentiell methylierter Genorte zwischen Rauchern und Nichtrauchern mittels p-Wert und $\Delta\beta$

Mittels für multiples Testen nicht angepasstem p-Wert, wurden bei $p < 0,0004$, $|\Delta\beta| > 0,05$ und einer Varianz $\sigma/\sigma_{\max} > 0,1$ 131 differentiell methylierte CpG-Positionen identifiziert, die

Ergebnisse

mit 68 Genorten assoziiert werden konnten (siehe Tabelle A5 im Anhang). Diese wurden graphisch mittels einer hierarchischen Clusteranalyse in einer entsprechenden Heatmap dargestellt (Abbildung 31). Die Raucher konnten von den Nichtrauchern durch zwei Haupt-Clusterarme getrennt werden, wobei 113/131 CpGs bei Rauchern mit einer Hypomethylierung und 18/131 CpGs mit einer Hypermethylierung einhergingen.

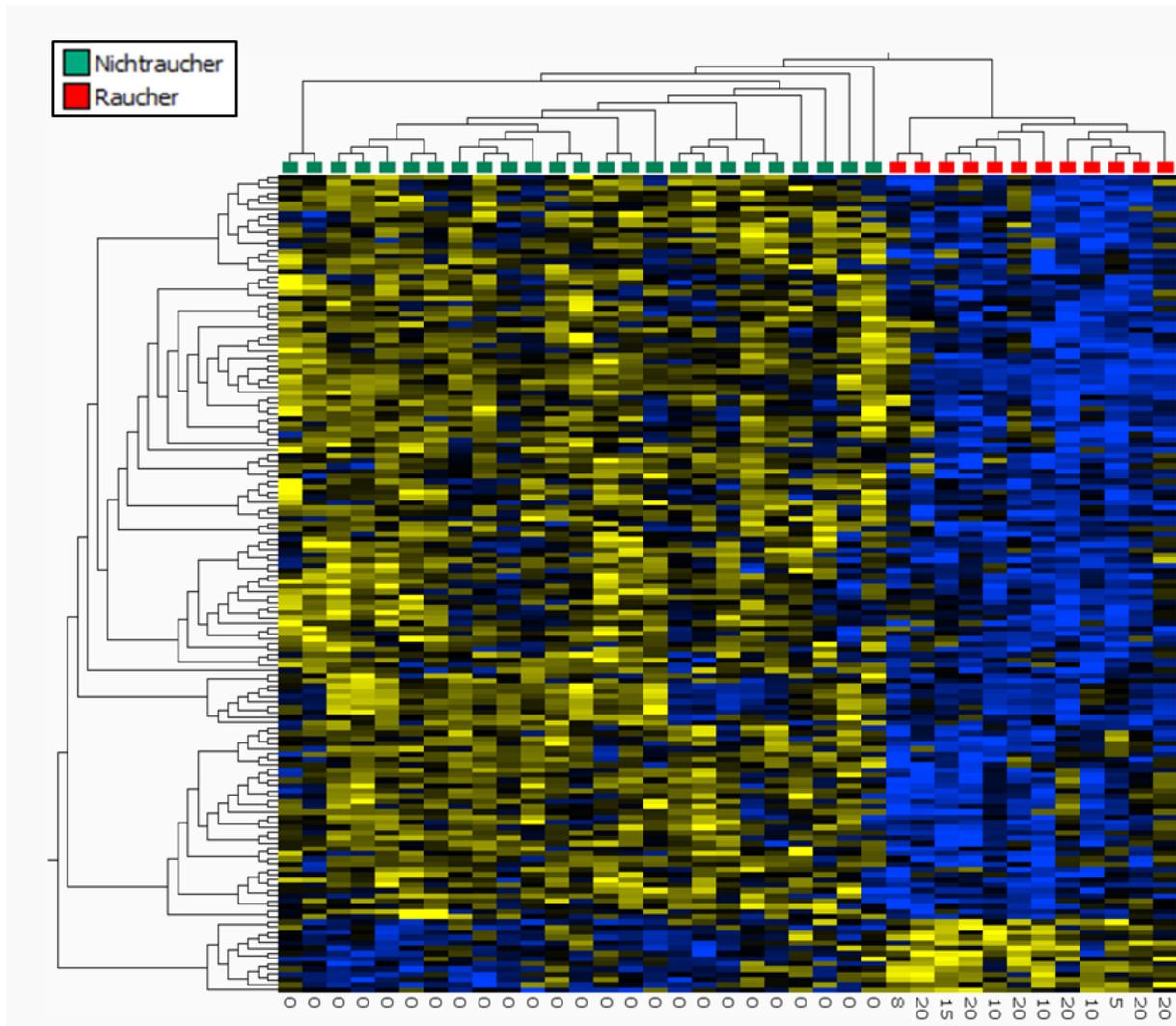


Abbildung 31: Differentielle Methylierungsanalyse von Rauchern (n=12) vs. Nichtrauchern (n=25). Eine hierarchische Clusteranalyse (*average linkage*) der 131 differentiell methylierten CpGs ($p < 0,0004$; t-Test; $|\Delta\beta| > 0,05$, $\sigma/\sigma_{\max} > 0,1$) zeigt, dass in den Rauchern (rot) 113/131 CpGs eine Hypermethylierung im Vergleich zu den Nichtrauchern (grün) zeigen. Eine Hypomethylierung wurde hingegen nur für 18/131 CpGs ermittelt. Spalten repräsentieren Proben, Reihen hingegen CpG-Loci. Innerhalb der Heatmap gibt ein blauer Farbcode eine geringe Methylierung an, während eine hohe Methylierung durch einen gelben Farbcode widergespiegelt wird. Auf der x-Achse der Heatmap ist die Anzahl der gerauchten Zigaretten pro Tag angegeben. Zur Visualisierung wurden die Mittelwerte sowie die Varianz für jeden CpG-Locus gegen 0 bzw. 1 normalisiert (mean=0, var=1).

Ergebnisse

Der p-Wert erstreckte sich in den 131 Positionen von $p < 0,0004$ bis hin zu $p < 1,72 \times 10^{-16}$. Der zuvor über die FDR ermittelte 533bp große differentiell methylierte Bereich auf Chromosom 2 (Chr2:233284402–233284934) konnte durch zwei weitere CpG-Loci um etwa 34kb *downstream* ausgeweitet werden. Hierbei waren die beiden zusätzlichen CpGs cg27241845 und cg03329539 jeweils in N-Shores von 2053bp bzw. 2562bp großen CpG-Inseln lokalisiert. Ferner konnte auch die bereits im *AHRR*-Genkörper identifizierte Region um ca. 49kb *downstream* und etwa 5kb *upstream* durch drei zusätzlich differentiell methylierte CpGs erweitert werden, die ebenfalls im Genkörper lokalisiert waren. Auch die durch den t-Test ermittelten CpGs in der 2q37.1- und *AHRR*-Region konnten bereits von Zeilinger *et al.* [133] und Dogan *et al.* [134] gezeigt werden, was den hier verwendeten Ansatz unterstützt.

Nachdem die Gelegenheits- und ehemaligen Raucher ohne statistischen Bezug in der zur Heatmap korrespondierenden PCA dargestellt wurden, zeigte sich eine zum stringenteren Ansatz vergleichbare Anordnung (Abbildung 32). Auch hier waren die Gelegenheitsraucher und die ehemaligen Raucher sowohl bei den Nichtrauchern als auch bei den Rauchern lokalisiert, wobei die Randbereiche von den Rauchern und Nichtrauchern gebildet wurden.

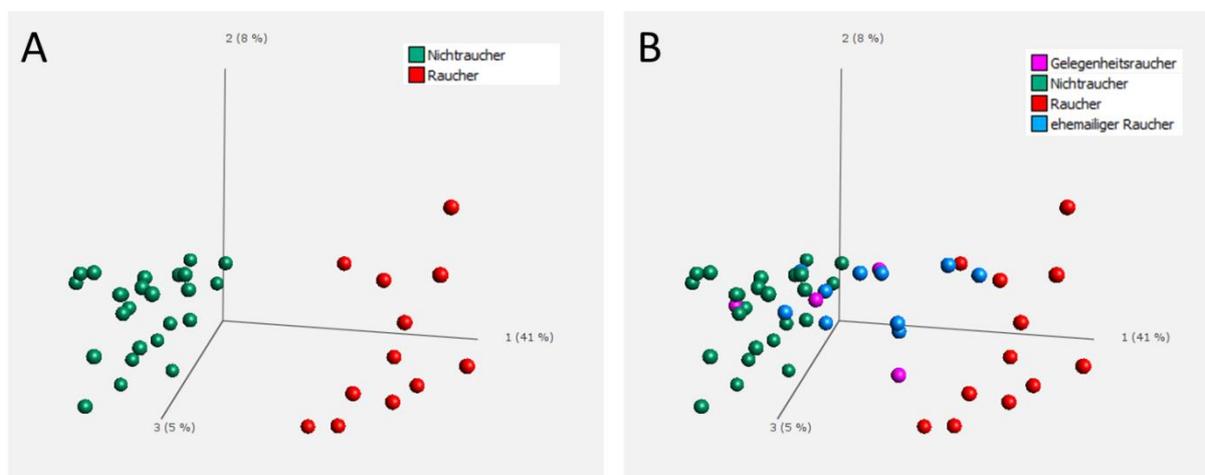


Abbildung 32: PCA der differentiellen Methylierungsanalyse von Rauchern (n=12) vs. Nichtrauchern (n=25). (A) zeigt die PCA der 131 differentiell methylierten CpGs ($p < 0,0004$; t-Test; $|\Delta\beta| > 0,05$, $\sigma/\sigma_{\max} > 0,1$) zwischen den Rauchern (rot) und den Nichtrauchern (grün). In (B) sind zusätzlich in der in (A) generierten PCA simultan die Gelegenheitsraucher (rosa) und die ehemaligen Raucher (blau) eingeblendet.

4.2.5.2.1 Eigenschaften der zwischen Raucher und Nichtraucher differentiell methylierten CpG-Regionen (n=131)

Keine An-/Abreicherungen der mit dem Rauchen assoziierten CpG-Loci in bestimmten regulatorischen Elementen

In der Gruppe der regulatorischen Elemente (Enhancer, DHS, DMR) konnten keine signifikanten An- oder Abreicherungen der mit dem Rauchen assoziierten Loci beobachtet werden ($p > 0,05$; χ^2 Test).

Keine An-/Abreicherungen der mit dem Rauchen assoziierten CpG-Loci im Gen-Zusammenhang

Auch hinsichtlich der Verteilung der identifizierten Loci im Gen-Zusammenhang (5'-UTR, TSS1500/200, Exon1, Genkörper oder 3'-UTR) lieferte die Anreicherungsanalyse keine signifikanten Ergebnisse ($p > 0,05$; χ^2 Test).

Abreicherung der mit dem Rauchen assoziierten CpG-Loci in Promotor-assoziierten Regionen

Eine signifikante Abreicherung der differentiell methylierten CpGs konnte in Promotor-assoziierten Regionen (OR=0,38; $p < 0,005$; χ^2 Test) ermittelt werden. Diese Regionen waren dabei ausschließlich auf die in Rauchern hypomethylierten Genen beschränkt (OR=0,45; $p < 0,001$; χ^2 Test).

Anreicherung der mit dem Rauchen assoziierten CpG-Loci in N-Shore Regionen bei simultaner Abreicherung in CpG-Inseln

Hinsichtlich der 131 differentiell methylierten CpGs konnte eine signifikante Anreicherung für N-Shore Regionen (OR=1,74; $p < 0,01$; χ^2 Test), die vor Allem in den bei Rauchern hypomethylierten Genen zu finden war (OR=2,00; $p < 0,005$; χ^2 Test), identifiziert werden. Weiterhin konnte eine signifikante Abreicherung für CpG-Inseln (OR=0,44; $p < 0,001$; χ^2 Test), für die auch hier überwiegend die in den Rauchern hypomethylierten Gene verantwortlich waren (OR=0,41; $p < 0,001$; χ^2 Test), ermittelt werden.

4.2.5.2.2 GO-Analyse der zwischen Raucher und Nichtraucher differentiell methylierten Genorte (n=68)

Die GO-Analyse lieferte sowohl für den „*biological process*“ als auch für „*molecular function*“ und „*cellular component*“ keine signifikanten An- oder Abreicherungen.

4.2.5.2.3 Erweiterte Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte (n=78) mittels WebGestalt

Eine signifikante Anreicherung der mit dem Rauchen assoziierten Gene konnte im *KEGG Pathway* „Renin-Angiotensin-System“ mit den Genen *ANPEP* und *ENPEP* ermittelt werden (OR=35,98; q=0,0070). Das Renin-Angiotensin-System reguliert neben dem Flüssigkeits- und Elektrolythaushalt auch den Blutdruck [221], der durch das Rauchen erhöht wird [222].

Eine Phänotyp-Analyse basierend auf der Phänotyp-Ontologie von Säugern und Menschen brachte ferner u.a. signifikante Anreicherungen (OR=2,61–47,65; q<0,05) für spezifische Defekte des kardiovaskulären Systems hervor (Tabelle 12), die größtenteils mit einer Fehlbildung des Herzventrikels assoziiert wurden.

Tabelle 12: Phänotyp-Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene. Die identifizierten Phänotypen sind entsprechend ihrer Signifikanz angeordnet.

Phänotyp	beobachtet	erwartet	OR	p-Wert	q-Wert
Fehlbildung des Herzventrikels	5	0,57	8,77	8,14x10 ⁻⁰⁵	0,019
Ventrikelseptumdefekt	4	0,40	10,12	3,00x10 ⁻⁰⁴	0,023
Wachstumsverzögerung	7	2,09	3,35	4,00x10 ⁻⁰⁴	0,023
Fehlbildung des Ventrikelseptums	4	0,40	9,94	4,00x10 ⁻⁰⁴	0,023
Fehlbildung des kardiovaskulären Systems	8	3,07	2,61	5,00x10 ⁻⁰⁴	0,023
Defekt des Atrioventrikularkanals	2	0,04	47,65	7,00x10 ⁻⁰⁴	0,027

OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

Eine weitere Analyse hinsichtlich assoziierter Krankheitsbilder lieferte analog zu den Ergebnissen der Phänotyp-Analyse signifikante Anreicherungen für Gene (OR=33,74–67,47; q<0,05), die in spezifischen Erkrankungen des Herzens involviert sind (Tabelle 13). Beteiligt waren die jeweiligen Genkombinationen *GATA4* und *CACNA1C*, *GATA4* und *EVC2* sowie *GATA4* und *NODAL*.

Ergebnisse

Tabelle 13: Krankheitsassoziations-Analyse der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene. Die assoziierten Erkrankungen sind entsprechend ihrer Signifikanz angeordnet.

Assoziierte Erkrankung	beobachtet	erwartet	OR	p-Wert	q-Wert
Atrioventrikulärer Septumdefekt	2	0,03	67,47	$4,00 \times 10^{-04}$	0,024
Persistierendes Foramen ovale	2	0,04	49,07	$7,00 \times 10^{-04}$	0,024
Pulmonalatresie	2	0,06	33,74	0,002	0,037

OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

Die ermittelten Anreicherungen für diverse das kardiovaskuläre System betreffende Fehlbildungen und Erkrankungen befinden sich in Einklang mit der Tatsache, dass Rauchen als großer Risikofaktor für Herz-Kreislauf-Erkrankungen angesehen wird [219].

4.2.5.2.4 Identifizierung von krebsrelevanten Genen innerhalb der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene

Da Rauchen allgemein mit einem erhöhten Risiko für diverse Tumorentitäten verbunden ist [223], wurde die Gruppe der zwischen Raucher und Nichtraucher differentiell methylierten Genorte auf bekannte Onkogene und Tumorsuppressorgene hin überprüft. Unter den insgesamt 68 mit Rauchen assoziierten aberrant methylierten Genorten konnten jeweils vier Gene (*GDNF*, *MUC4*, *PTTG1*, *RRAS2*) den Onkogenen und vier Gene (*AHRR*, *AKAP12*, *PTPRG*, *TNFRSF10A*) den Tumorsuppressorgenen zugeordnet werden.

4.2.5.3 Prüfung der mit dem Rauchen assoziierten über die FDR ermittelten 22 CpG-Signatur in Lungen-Proben

Laut Schätzungen der WHO sind etwa 70% der Lungenkrebsfälle allein auf das Rauchen zurückzuführen (<http://www.who.int/cancer/prevention/en/>). Da aufgrund einer anderen Studie am Institut für Humangenetik der Datensatz des Methylierungsmusters eines Lungengewebekollektivs bestehend aus 40 unauffälligen und 40 Krebsproben zur Verfügung stand, wurde in diesem die zwischen Rauchern und Nichtrauchern ermittelte Methylierungssignatur ($n=22$ CpGs; $q<0,05$) betrachtet. Es sollte somit überprüft werden, ob eine in Zellen des peripheren Blutes veränderte DNA-Methylierung der hier identifizierten mit dem Rauchen assoziierten CpGs auch in Lungenkrebsproben im Vergleich zu unauffälligen Lungenproben detektiert werden kann. Um gleiche Bedingungen zu schaffen,

Ergebnisse

wurde der Lungendatensatz nach erfolgter Qualitätskontrolle ebenfalls mit der SWAN-Methode normalisiert. Anschließend wurden im QIUCORE Omics Explorer in den verbliebenen 67 Lungenproben ausschließlich die Methylierungswerte der 22 CpG-Signatur betrachtet. Die unsupervidierte hierarchische Clusteranalyse und die korrespondierende PCA offenbarten die Tendenz, dass sich das Methylierungsmuster der 33 Lungenproben von den 34 Lungenkrebsproben an den genannten 22 CpGs unterscheiden ließ und sich die beiden Gruppen entsprechend trennen lassen könnten (Abbildung 33).

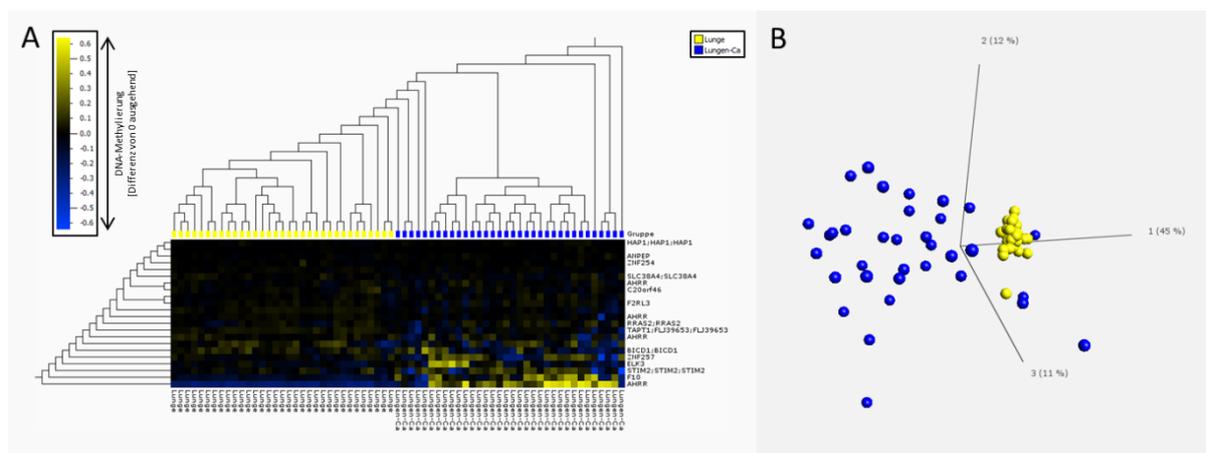


Abbildung 33: Unsupervidierte hierarchische Clusteranalyse (A) und PCA (B) des DNA-Methylierungsmusters der 33 Lungen- und 34 Lungenkrebsgewebeproben für die zwischen Rauchern und Nichtrauchern identifizierten differentiell methylierten CpGs (n=22). (A) zeigt eine Heatmap der 22 CpGs, die aus der Methylierungsanalyse zwischen Rauchern und Nichtrauchern ($q < 0,05$; t-Test) als differentiell methyliert identifiziert wurden (*average linkage*). Zur Visualisierung wurden die Mittelwerte für jeden CpG-Locus gegen 0 normalisiert (mean=0). (B) zeigt die ersten drei PCAs der korrespondierenden PCA.

Während alle 22 CpG-Loci aus der SAME Kohorte über eine FDR $q < 0,05$ identifiziert wurden, erfüllten nur etwa 27% (6/22) der CpGs in den Lungenproben diesen Schwellenwert (siehe Abbildung 34A). Diese sechs CpG-Loci konnten den Genorten *AHRR*, *BICD1*, *F10*, *SLC38A4* und *TAPT1* zugeordnet werden und wurden hinsichtlich der Richtung ihrer Methylierungsveränderung analysiert. In der vorangegangenen Analyse konnte im Blut der Raucher im Vergleich zu dem der Nichtraucher im Durchschnitt in allen Loci eine Hypomethylierung ermittelt werden, welche ebenfalls für 4/6 CpG-Loci in den Lungenkrebsproben festgestellt werden konnte. Die anderen zwei CpG-Loci (siehe Abbildung 34A und B) zeigten in den Lungenkrebsproben im Vergleich zum Normalgewebe eine durchschnittliche Hypermethylierung, wobei cg08143875 ($q < 1,7 \times 10^{-4}$) mit einer mittleren Methylierungsdifferenz von 13,4% im Genkörper des *F10*-Gens und der Locus cg05575921 ($q < 6,65 \times 10^{-14}$) mit einem mittleren Methylierungsunterschied von 37,4% im Genkörper des

Ergebnisse

AHRR-Gen lokalisiert waren. Während sich die DNA-Methylierung in beiden Loci im Blut von Rauchern und den Lungenkrebsproben im Median auf einem ähnlichen Niveau befand (cg08143875: 39,8% vs. 41,4% bzw. cg05575921: 54,0% vs. 54,2%), unterschied sich diese deutlich zwischen den im Blut von Nichtrauchern und Lungenproben (cg08143875: 44,1% vs. 29,0% bzw. cg05575921: 80,2% vs. 16,3%).

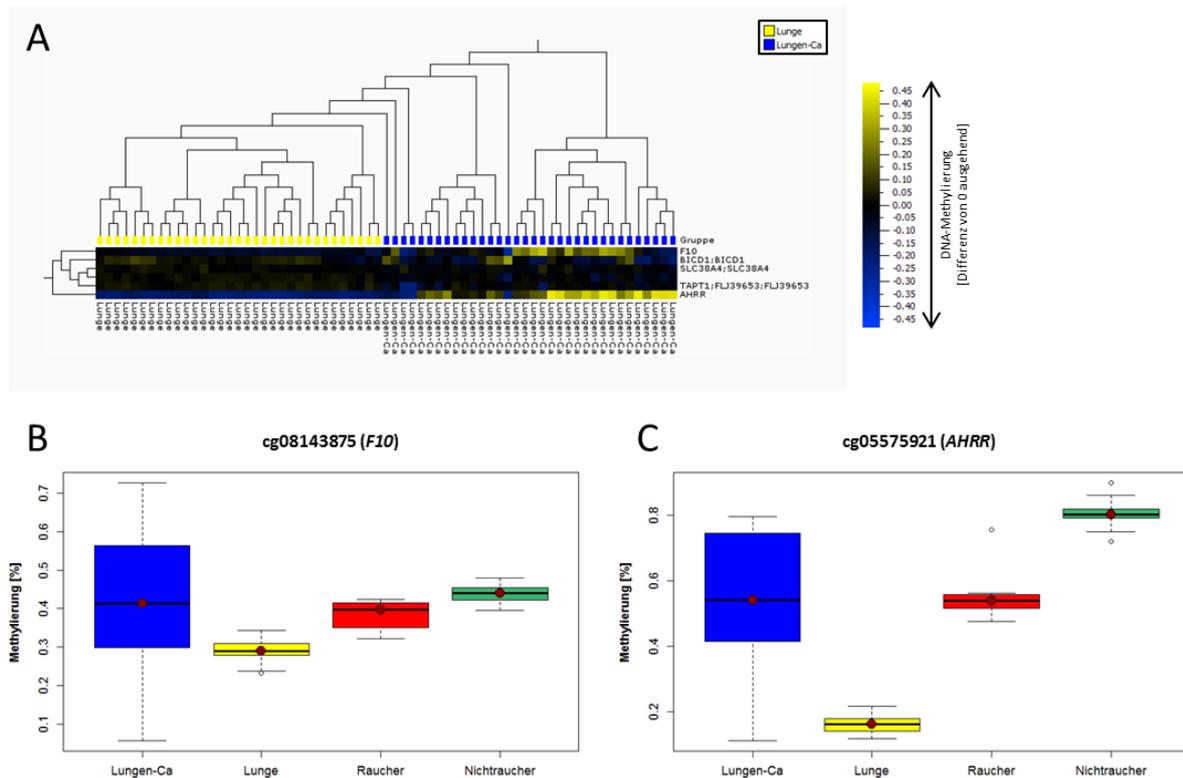


Abbildung 34: Unsupervidierte hierarchische Clusteranalyse (A) und Boxplot definierter CpG-Loci. Die Heatmap (*average linkage*) in (A) zeigt das Methylierungsmuster von Lungenkrebsproben (blau) und unauffälligen Lungenproben (gelb) in sechs CpG-Loci ($q < 0,05$; t-Test). Zur Visualisierung wurden die Mittelwerte für jeden CpG-Locus gegen 0 normalisiert (mean=0). Ein blauer Farbcodierung gibt ein niedriges, ein gelber Farbcodierung ein hohes Methylierungsniveau an. In (B) ist die DNA-Methylierung in zwei auffälligen CpG-Loci dargestellt. Während sich das Niveau der DNA-Methylierung in beiden Loci im peripheren Blut von Rauchern und im Lungenkrebsgewebe vergleichen lässt (39,8% vs. 41,4% bzw. 54,0% vs. 54,2%), verhielt sich dieses im peripheren Blut von Nichtrauchern und Lungengewebe unterschiedlich (80,2% vs. 16,3% bzw. 44,1% vs. 29,0%), was in einer Hypomethylierung im Blut von Rauchern, hingegen aber in einer Hypermethylierung im Lungenkrebsgewebe resultierte.

Die entgegengesetzte Richtung der differentiellen DNA-Methylierung in den genannten CpG-Loci im Blut von Nichtrauchern und Lungenproben könnte u.a. auf eine gewebespezifische DNA-Methylierung (Zellen des peripheren Blutes vs. Lungengewebe) oder auf unterschiedliche Funktionen der Gene in den beiden Kompartimenten zurückzuführen sein. Die Information über den Rauchstatus der Lungenkrebspatienten war nicht verfügbar, so

dass keine Rückschlüsse auf „reale“ mit dem Rauchen assoziierte Loci gezogen werden konnten.

4.2.6 Zwillingsstudie: Auswirkung von Schichtarbeit mit Nachtschichtdienst auf die genomweite DNA-Methylierung

Da in Schichtsystemen, die Nachtarbeit beinhalten, eine Veränderung der circadianen Rhythmik beobachtet werden kann [224], wurde das Methylierungsmuster von Zwillingen innerhalb eines Zwillingspaares, die sich hinsichtlich Schichtarbeit mit Nachtschicht diskordant verhielten, analysiert. In der Gesamtkohorte der Zwillinge qualifizierten sich hierfür acht monozygote sowie sechs dizygoten Zwillingspaare (vgl. Tabelle 3).

4.2.6.1 Identifizierung differentiell methylierter Loci zwischen Schichtarbeitern mit Nachtschicht und Tagarbeitern

Mittels Wilcoxon-Test wurden zuerst die Methylierungswerte aller Loci der für Schichtarbeit mit Nachtschichtdienst diskordanten Zwillinge der monozygoten Zwillingspaare (n=16 Zwillinge) auf Signifikanz hin überprüft und die ermittelten signifikanten Loci mit dem Methylierungsmuster dem der diskordanten dizygoten Zwillinge (n=12) verglichen. Es konnten zunächst 4685 signifikante Loci ermittelt werden, die sich in den monozygoten Schichtarbeitern von den am Tag arbeitenden Co-Zwillingen unterschieden. Die mittleren Methylierungsdifferenzen reichten hierbei von -7,8% bis hin zu 9,6%. Für die weitere Analyse sollten nur diejenigen Loci berücksichtigt werden, die im Methylom der schichtarbeitenden MZ durchwegs entweder eine Hyper- oder Hypomethylierung im Vergleich zu den tagarbeitenden Co-Zwillingen aufwiesen. 792/4685 Loci erfüllten diese Bedingung, wobei 406 Loci eine Hypermethylierung und 386 Loci eine Hypomethylierung in den Schichtarbeitern zeigten. Da bei einem stringenteren Filtern kein Locus mit einer Methylierungsdifferenz von mindestens 5% in allen acht MZ-Paaren beobachtet werden konnte, wurde das Methylierungsmuster aller sich bis dahin qualifizierten Loci in den analog DZ-Paaren überprüft und konnte hinsichtlich einer einheitlichen Methylierungsrichtung für 14/792 Loci bestätigt werden. Folglich konnten in 14 diskordanten Zwillingspaaren zwischen Schichtarbeitern und Tagarbeitern 14 differentiell methylierte CpG-Loci mit einer gleichgerichteten Methylierungsdifferenz identifiziert bzw. bestätigt werden (siehe Tabelle 14). Hierbei zeigten sechs CpGs in den Schichtarbeitern eine Hypermethylierung, während

Ergebnisse

acht CpGs von einer Hypomethylierung betroffen waren. Die 14 CpG-Loci unterschieden sich zwar signifikant zwischen Schichtarbeitern und Tagarbeitern ($p=1,22 \times 10^{-04}$), die einzelnen Methylierungsdifferenzen ließen aber aufgrund sehr geringer Unterschiede (mittlere Methylierungsdifferenz: 1,92%) keine verlässliche Schlussfolgerung zu (Abbildung 35).

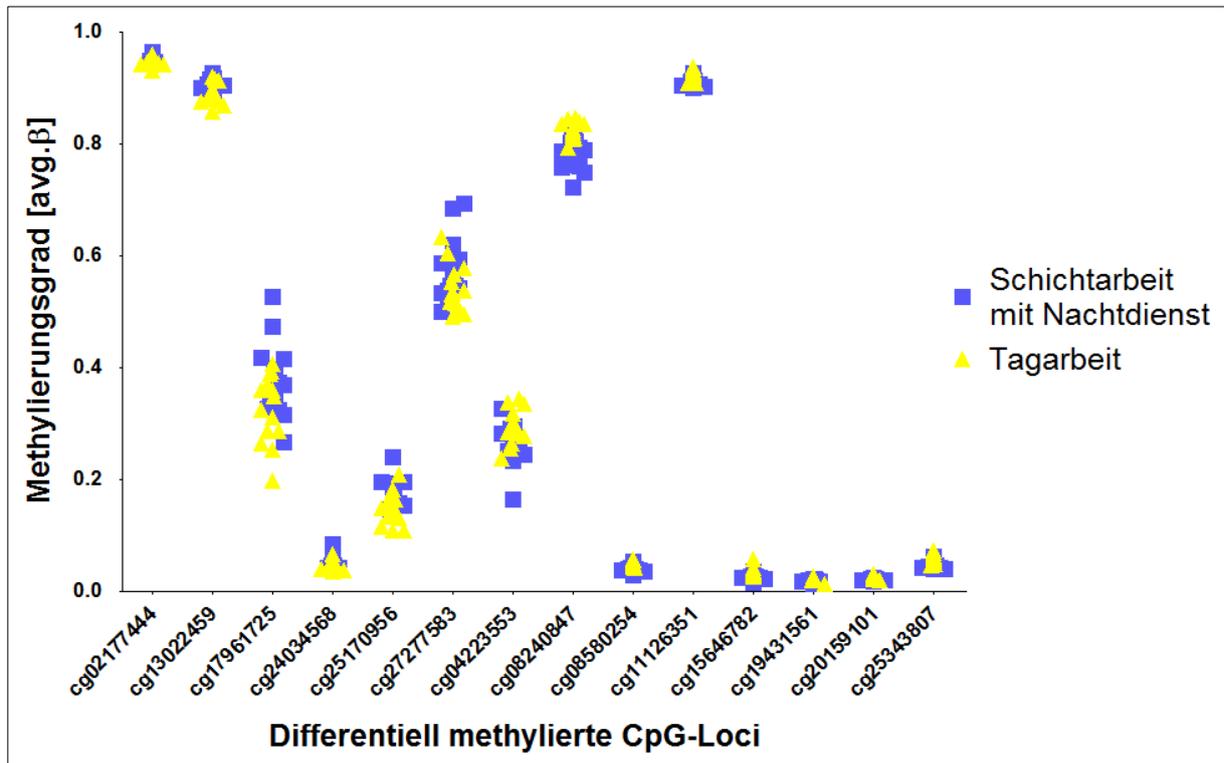


Abbildung 35: Verteilung der Methylierungswerte auf die 14 differentiell methylierten Loci. 6/14 CpGs zeigten in den Schichtarbeitern mit Nachtdienst im Vergleich zu den Tagarbeitern jeweils eine einheitliche Hypermethylierung, während 8/14 CpGs eine einheitliche Hypomethylierung aufwiesen.

Aufgrund dieser Resultate wurde von einer tiefergehenden Analyse bezüglich funktioneller bzw. biologischer Relevanz, abgesehen.

Tabelle 14: Liste der mittels Infinium HumanMethylation450 BeadChip identifizierten 14 CpG-Loci, die zwischen den Zwillingen mit Schichtarbeit inklusive Nachtschichtdienst und den Zwillingen mit einer Arbeit zur üblichen Tageszeit eine signifikante Methylierungsveränderung zeigen. Die Loci sind nach aufsteigender Position auf den Chromosomen geordnet.

CpG ID	Ch r	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierung [β] Zwilling Schichtarbeit	Mittlere Methylierung [β] Zwilling Tagarbeit	Mittlere Methylierungsdifferenz [Δβ] (Schicht – Tag)	p-Wert
cg25343807	2	74007420	DUSP11	TSS200	S-Shore	0,048±0,008	0,056±0,008	-0,008	1,22x10 ⁻⁰⁴
cg17961725	2	150454027	N/A	N/A	N/A	0,377±0,067	0,325±0,061	0,053	1,22x10 ⁻⁰⁴
cg02177444	2	216795947	N/A	N/A	N/A	0,954±0,005	0,948±0,007	0,006	1,22x10 ⁻⁰⁴
cg15646782	3	101498416	FAM55C	5'UTR	Island	0,026±0,005	0,033±0,008	-0,008	1,22x10 ⁻⁰⁴
cg24034568	3	184080707	CLCN2	TSS1500	Island	0,056±0,012	0,049±0,009	0,008	1,22x10 ⁻⁰⁴
cg04223553	5	153852934	N/A	N/A	N-Shore	0,262±0,037	0,294±0,032	-0,032	1,22x10 ⁻⁰⁴
cg13022459	6	112374876	WISP3	TSS1500	N/A	0,909±0,012	0,889±0,018	0,020	1,22x10 ⁻⁰⁴
cg20159101	8	58256057	N/A	N/A	N/A	0,021±0,002	0,024±0,003	-0,003	1,22x10 ⁻⁰⁴
cg08240847	8	134354024	N/A	N/A	N/A	0,781±0,028	0,822±0,016	-0,042	1,22x10 ⁻⁰⁴
cg25170956	11	68816067	TPCN2	TSS1500	N-Shore	0,182±0,023	0,148±0,029	0,034	1,22x10 ⁻⁰⁴
cg11126351	14	35594356	KIAA0391	Body	S-Shelf	0,912±0,007	0,923±0,009	-0,012	1,22x10 ⁻⁰⁴
cg27277583	15	76484625	C15orf27	Body	S-Shore	0,574±0,058	0,541±0,041	0,033	1,22x10 ⁻⁰⁴
cg08580254	17	38478807	RARA	5'UTR	S-Shelf	0,042±0,006	0,049±0,005	-0,007	1,22x10 ⁻⁰⁴
cg19431561	21	47648908	MCM3APAS	TSS1500	Island	0,020±0,002	0,022±0,003	-0,003	1,22x10 ⁻⁰⁴

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position des analysierten CpG-Locus nach UCSC Genome Browser. p-Wert ermittelt über Wilcoxon-Test. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

4.2.7 Zwillingsstudie: Auswirkung von Nachtarbeit auf die genomweite DNA-Methylierung

Als nächstes wurde der Effekt von reiner Nachtarbeit auf die DNA-Methylierung untersucht. Hierfür wurde das Methylierungsmuster von Zwillingen innerhalb eines Zwillingspaares, die sich hinsichtlich einer Arbeit bei Nacht diskordant verhielten, analysiert. Für diese Analyse qualifizierten sich drei monozygote und drei dizygoten Zwillingspaare (vgl. Tabelle 3).

4.2.7.1 Identifizierung differentiell methylierter Loci zwischen Nacht- und Tagarbeitern

Für die Identifikation von differentiell methylierten Loci, die in direktem Zusammenhang mit den Auswirkungen der Nachtarbeit auf das Methylom stehen könnten, wurde das Methylierungsmuster von monozygoten Zwillingspaaren ($n=3$), in denen je einer der Zwillinge einer Beschäftigung ausschließlich bei Nacht und der andere Zwilling einer Beschäftigung zur gewöhnlichen Tageszeit nachging, verglichen und dem Methylierungsmuster der analog dizygoten Zwillingspaare ($n=3$) gegenübergestellt. Da aufgrund der geringen Fallzahl ($n=3$ pro Gruppe) keine statistisch signifikanten Aussagen getroffen werden können, wurde manuell nach differentiell methylierten Loci gefiltert (siehe 3.3.3.3). Über eine definierte Intra-Paar-Differenz ($|\Delta\beta| \geq 0,05$) konnten 452 Loci ermittelt werden, die in den drei MZ-Paaren eine gleichgerichtete Methylierungsdifferenz aufwiesen. Für 132/452 Loci konnte eine gleichgerichtete Methylierungsdifferenz ebenfalls in den drei DZ-Paaren bestätigt werden (114 hyper- und 18 hypomethylierte Loci in den Nachtarbeitern). Da die einzelnen Intra-Paar-Differenzen der DZ größtenteils sehr gering waren, wurden stringenter Bedingungen gewählt, bei denen für die DZ ein gleiches $\Delta\beta \geq 0,05$ bzw. $\Delta\beta \leq -0,05$ angenommen wurde. Dieses Kriterium erfüllten 20 CpGs sowie ein CpN Locus, die in den Nachtarbeitern der MZ und DZ im Vergleich zu den korrespondierenden Tagarbeitern jeweils hypermethyliert ($\Delta\beta > 0,05$) vorlagen (siehe Abbildung 36). Die 21 differentiell methylierten Loci zwischen den sechs Nachtarbeitern (jeweils drei MZ und DZ) und ihren entsprechenden Zwillings-Geschwistern (sechs Tagarbeiter) wurden mit dem Wilcoxon-Test analysiert und erwiesen sich als statistisch signifikant ($p < 0,05$). Die ermittelten 21 Loci konnten laut Infinium Manifest zehn Genen zugeordnet werden: *CLSTN2*, *CREB5*, *FGA*, *MED1*, *SH2D1B*, *SLC6A1*, *SLC11A1*, *SPATA19*, *TMEM132D* und *TRIM10*. *CREB5* und *TRIM10* waren hierbei von jeweils zwei differentiell methylierten Loci betroffen. Darüber

Ergebnisse

hinaus konnte der Locus cg04638710 nachträglich mittels UCSC Genome Browser dem Gen *TENM3* zugeordnet werden (siehe Tabelle 16).

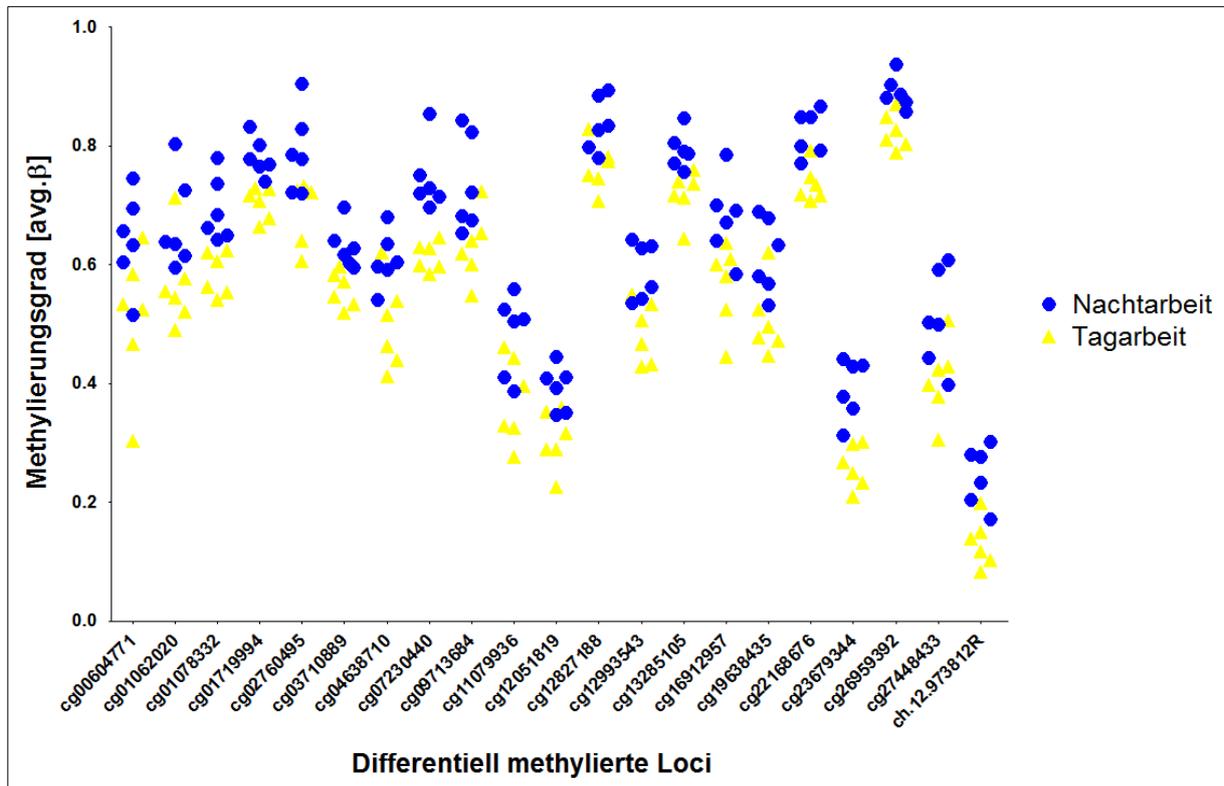


Abbildung 36: Verteilung der Methylierungswerte auf die 21 differenziell methylierten Loci, die in den Nachtarbeitern im Vergleich zu den Tagarbeitern jeweils eine einheitliche Hypermethylierung zeigten.

4.2.7.2 Eigenschaften der zwischen Nacht- und Tagarbeitern differenziell methylierten CpG-Regionen

Über die Anreicherungsanalyse für bestimmter Eigenschaften und Merkmale der zwischen Nacht- und Tagarbeitern differenziell methylierten CpG-Regionen konnte lediglich eine signifikante Abreicherung in CpG-Inseln ermittelt werden (OR=0,11; $p < 0,01$; χ^2 Test).

4.2.7.3 GO-Analyse der zwischen Nacht- und Tagarbeitern differenziell methylierten Genorte

Eine GO-Analyse (WebGestalt) lieferte in der Kategorie „*biological process*“ zwei signifikant angereicherte GO-Kategorien, die in nachfolgender Abbildung dargestellt sind.

Ergebnisse

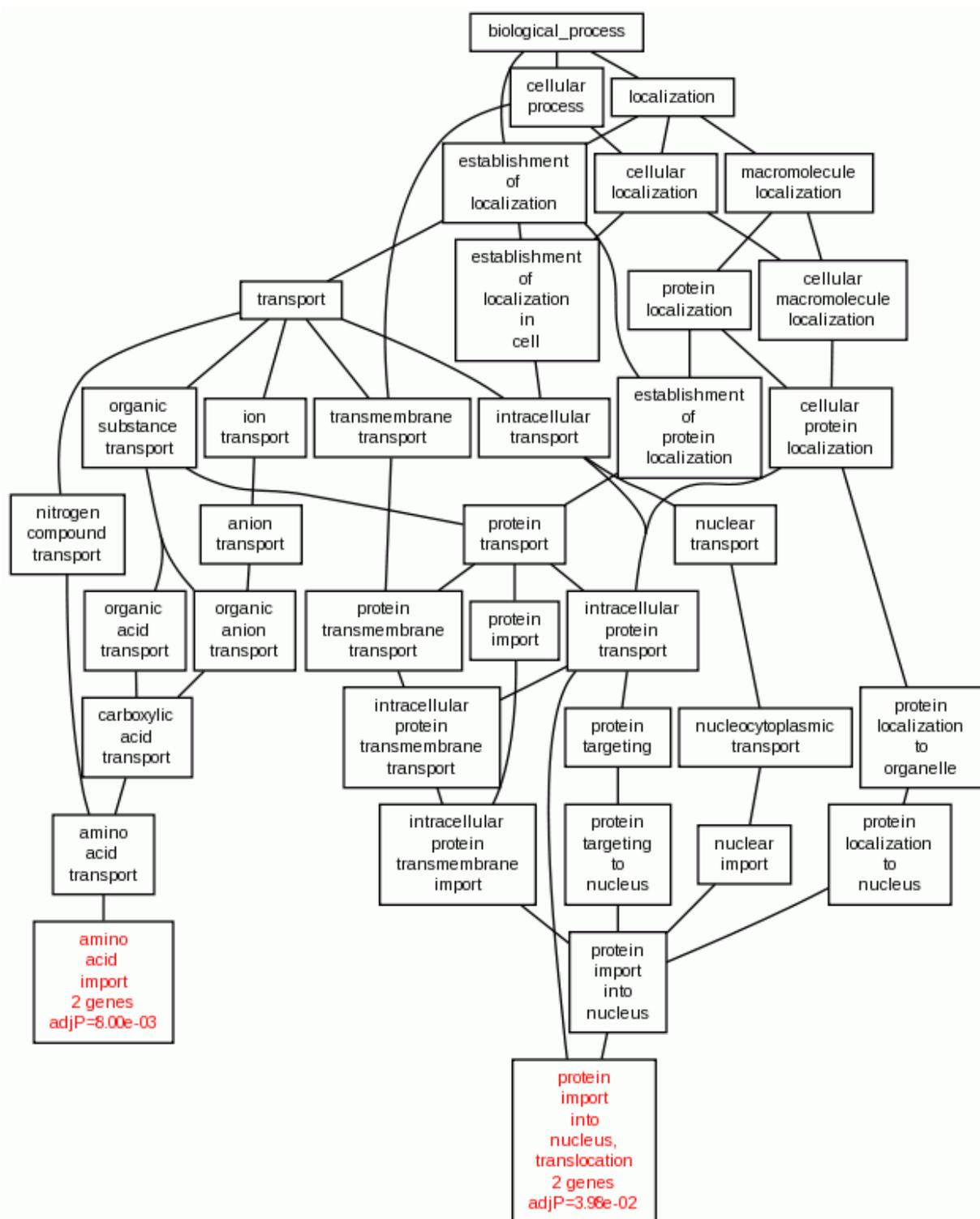


Abbildung 37: Anreicherungsanalyse für Genontologie-Kategorien. Für multiples Testen korrigierte (Methode nach Benjamini und Hochberg) signifikante GO-Kategorien sind in hellrot hinterlegt ($q < 0,05$).

An dem für „Aminosäure-Import“ angereichertem Prozess ($OR=202,55$; $q < 0,01$) waren die Gene *SLC6A1* und *SLC11A1*, zwei Transporter aus der *solute carrier* Familie, involviert. Ebenfalls unter der Beteiligung von zwei Genen, *SLC11A1* und *MED1*, konnte der Prozess

„Protein-Import in den Zellkern“ als signifikant angereichert (OR=69,29; $q < 0,05$) identifiziert werden.

4.2.7.4 Erweiterte Analyse der 11 zwischen Nacht- und Tagarbeitern differentiell methylierten Genorte mittels WebGestalt

Unter den weiterführenden Analysen fielen in einer Anreicherungsanalyse für krankheitsassoziierte Gene hauptsächlich Gene auf, die mit Darmerkrankungen und psychischen Erkrankungen assoziiert waren.

Tabelle 15: Krankheitsassoziations-Analyse der zwischen Nacht- und Tagarbeitern differentiell methylierten Gene. Die assoziierten Erkrankungen sind entsprechend ihrer Signifikanz angeordnet.

Krankheits-Assoziation	beobachtet	erwartet	OR	p-Wert	q-Wert
Panikstörung	2	0,06	33,99	0,002	0,008
Morbus Crohn	2	0,08	24,97	0,003	0,008
Chronisch-entzündliche Darmerkrankungen	2	0,10	20,97	0,004	0,008
Gastroenteritis	2	0,10	20,51	0,004	0,008
Wechselwirkungen zwischen Arzneimitteln	2	0,19	10,80	0,014	0,022
Translokation, genetisch	2	0,21	9,68	0,017	0,023
Psychische Störung	2	0,25	7,86	0,025	0,029

OR: *odds ratio*. p-Wert ermittelt über hypergeometrischen Test. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

Für die Anreicherung der Darmerkrankungen waren jeweils die Gene *CLSTN2* und *SLC11A1* verantwortlich, während für die psychischen Erkrankungen die Gene *SLC6A1* und *TMEM132D* (Panikstörung) bzw. *SLC6A1* und *CLSTN2* (psychische Störung) eine Rolle spielten. Neben einem erhöhten Risiko für Herz-Kreislauf-Erkrankungen [225] sind Erkrankungen des Gastrointestinaltrakts [226,227] sowie psychische Symptome wie Angststörungen [228,229] sowohl bei rotierenden Schichtarbeitern als auch bei Nachtarbeitern bekannt.

Tabelle 16: Liste der mittels Infinium HumanMethylation450 BeadChip identifizierten 21 Loci (20 CpGs und 1 CpN), die zwischen den Zwillingen mit einer Beschäftigung bei Nacht und den Zwillingen mit einer Arbeit zur üblichen Tageszeit eine differentielle Methylierung zeigen. Die Loci sind nach aufsteigender Position auf den Chromosomen geordnet.

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierung [β] Zwilling Nachtarbeit	Mittlere Methylierung [β] Zwilling Tagarbeit	Mittlere Methylierungsdifferenz [Δβ] (Nacht – Tag)	p-Wert
cg01062020	1	162382848	SH2D1B	TSS1500	N/A	0,669±0,080	0,567±0,078	0,102	0,031
cg01719994	1	32890253	N/A	N/A	N-Shelf	0,781±0,032	0,704±0,027	0,078	0,031
cg12051819	1	114717638	N/A	N/A	N/A	0,393±0,038	0,306±0,050	0,087	0,031
cg03710889	2	219252818	SLC11A1	Body	S-Shore	0,630±0,036	0,559±0,030	0,071	0,031
cg01078332	3	11033246	SLC6A1	TSS1500	N-Shore	0,693±0,054	0,585±0,037	0,108	0,031
cg02760495	3	139678075	CLSTN2	Body	N/A	0,790±0,070	0,691±0,054	0,098	0,031
cg04638710	4	183167912	TENM3	Body	N/A	0,609±0,046	0,498±0,076	0,110	0,031
cg12827188	4	155511987	FGA	TSS200	N/A	0,836±0,046	0,765±0,041	0,072	0,031
cg13285105	4	24722947	N/A	N/A	N/A	0,793±0,032	0,718±0,040	0,075	0,031
cg09713684	5	134372398	N/A	N/A	N-Shore	0,733±0,081	0,630±0,059	0,103	0,031
cg11079936	6	30124299	TRIM10	Body	N/A	0,482±0,068	0,372±0,073	0,110	0,031
cg16912957	6	30124214	TRIM10	Body	N/A	0,679±0,067	0,566±0,070	0,113	0,031
cg22168676	6	10060780	N/A	N/A	N/A	0,821±0,038	0,736±0,031	0,085	0,031
cg19638435	7	28452093	CREB5	5'UTR	S-Shelf	0,613±0,064	0,507±0,062	0,107	0,031
cg27448433	7	28452066	CREB5	5'UTR	S-Shelf	0,507±0,082	0,407±0,066	0,101	0,031
cg26959392	11	133715511	SPATA19	TSS200	N/A	0,890±0,028	0,825±0,030	0,065	0,031
cg07230440	12	130357212	TMEM132D	Body	N/A	0,744±0,057	0,614±0,024	0,131	0,031
cg12993543	12	108784905	N/A	N/A	N/A	0,590±0,049	0,487±0,052	0,104	0,031
ch.12.973812R	12	49046495	N/A	N/A	N/A	0,245±0,050	0,132±0,041	0,113	0,031
cg00604771	16	32213944	N/A	N/A	Island	0,642±0,079	0,510±0,118	0,133	0,031
cg23679344	17	37560925	MED1	3'UTR	S-Shore	0,392±0,051	0,261±0,037	0,131	0,031

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position nach UCSC Genome Browser. p-Wert ermittelt über Wilcoxon-Test. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

4.2.7.5 Validierung der Nacharbeit-Signatur für ausgewählte Genorte

Da in dieser Studie sowohl ein zwillingsbasierter Ansatz als auch der Effekt von Schichtarbeit auf die DNA-Methylierung im Vordergrund standen, wurde schließlich die 21 Loci-Signatur der Nacharbeiter für die Validierung ausgewählt. Die Methylierungswerte der mittels HM450K BC identifizierten 21 Loci-Signatur befanden sich zwar größtenteils im mittleren Methylierungsbereich, in dem bereits eine schwächere Korrelation zwischen dem Array und der Bisulfit-Pyrosequenzierung gezeigt werden konnte (vgl. Kapitel 4.1), da aber nicht die spezifischen Methylierungswerte selbst, sondern nur die Methylierungsdifferenzen der jeweiligen Loci zwischen Nacht- und Tagarbeitern validiert werden sollten, wurde die Validierung mittels Bisulfit-Pyrosequenzierung durchgeführt. Hierfür wurden 10/21 Genorte aufgrund der Stärke ihrer Methylierungsdifferenzen sowie ihrer biologischen Relevanz ausgewählt (Abbildung 38).

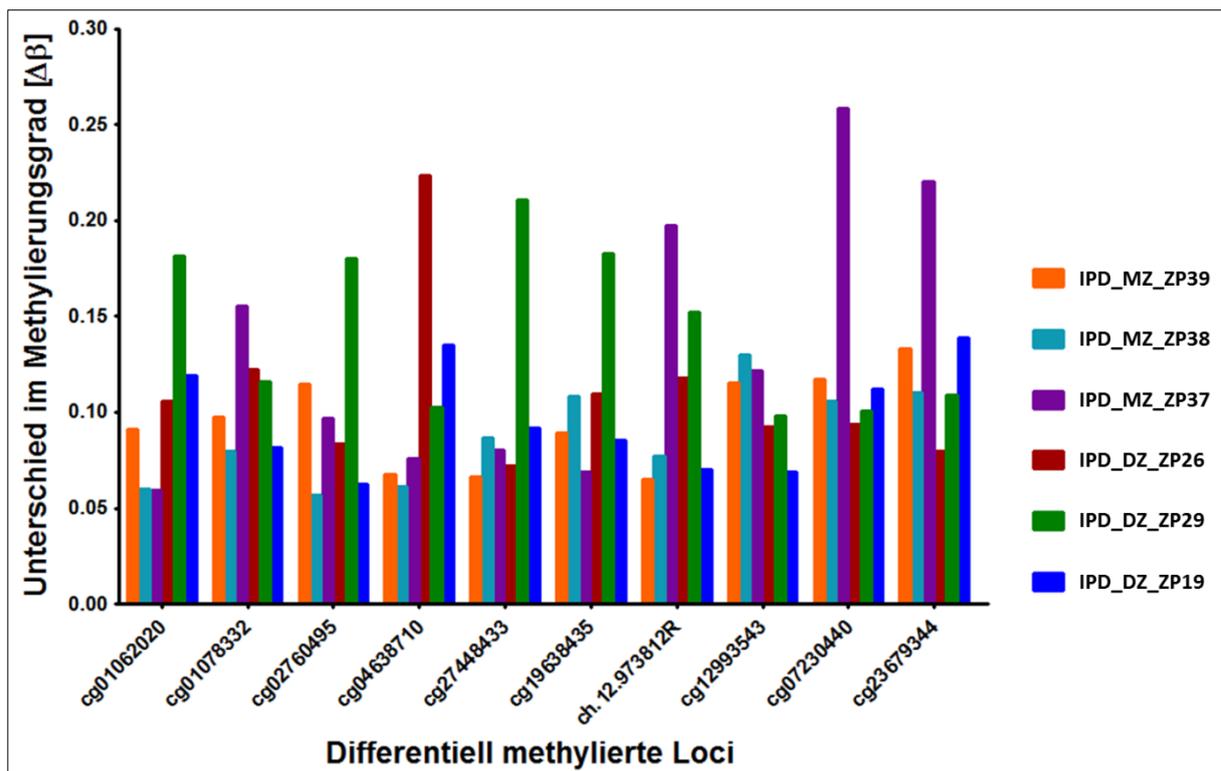


Abbildung 38: Auswahl der differentiell methylierten Loci für die Validierung mittels Bisulfit-Pyrosequenzierung. Dargestellt sind die ermittelten Intra-Paar-Differenzen der für Nacharbeit diskordanten Zwillinge. IPD: Intra-Paar-Differenz. MZ: monozygot. DZ: dizygot. ZP(n) gibt die Zwillingspaar-ID an.

Das Primerdesign konnte für alle ausgewählten Genorte erfolgreich durchgeführt werden. Die Validierung erfolgte in allen zuvor bereits auf den HM450K BC hybridisierten Zwillingspaarproben, die sich für Nacharbeit diskordant verhielten (n=12 Zwillinge) sowie in

Ergebnisse

zwei zusätzlichen für Nachtarbeit diskordanten dizygoten Zwillingspaaren (n=4 Zwillinge), von denen bei drei Zwillingen noch keine DNA-Methylierungsanalyse mittels HM450K BC durchgeführt wurde. Die DNA-Proben aller acht Zwillingspaare wurden für die Validierung erneut von unseren Kooperationspartnern angefordert. Nach erfolgter Bisulfitkonvertierung wurde an den 16 Proben die Pyrosequenzierung von jeweils zehn Genorten durchgeführt (siehe 3.4.1). Für 12 der 16 Proben konnte an den ausgewählten Genorten zuvor mittels HM450K BC eine Methylierungsdifferenz von mindestens 5% nachgewiesen werden ($p < 0,05$; Wilcoxon-Test; vgl. Tabelle 16 und siehe Abbildung 39), während in die Validierung vier weitere, zuvor nicht untersuchte Proben einbezogen wurden. Die Ergebnisse der Bisulfit-Pyrosequenzierung sind in nachstehender Abbildung dargestellt.

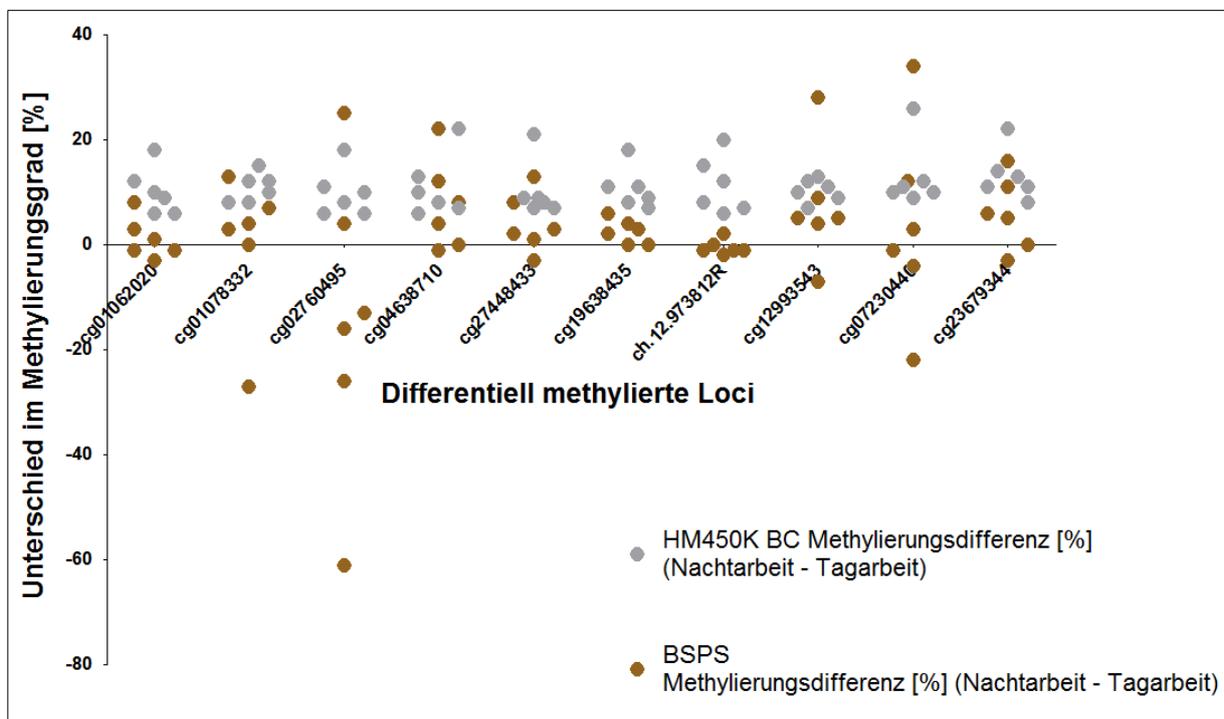


Abbildung 39: Ermittelte Unterschiede im Methylierungsgrad zehn ausgewählter Genorte mittels Infinium HM450K BC und Bisulfit-Pyrosequenzierung (BSPS). Ein Datenpunkt gibt die jeweilige Methylierungsdifferenz zwischen Nacht- und Tagarbeitern an. Graue Datenpunkte entsprechen den mittels HM450K BC ermittelten Methylierungsdifferenzen, braune Datenpunkte spiegeln die über die BSPS detektierten Methylierungsunterschiede wider.

In keinem der zehn Loci konnte eine gleichgerichtete Methylierungsdifferenz von mindestens 5% zwischen Nachtarbeitern und Tagarbeitern detektiert werden. Darüber hinaus konnte in nahezu allen Loci der Nachtarbeiter, mit Ausnahme von cg19638435, keine durchgehende Hypermethylierung (im Vergleich zu den Tagarbeitern) festgestellt werden.

Ergebnisse

Die einzelnen Methylierungswerte der zehn Genorte in jeweils 12 Zwillingen wurden daraufhin aus beiden Methoden verglichen (siehe Abbildung 40A). Hierbei konnte eine signifikante Gesamtkorrelation zwischen den beiden Methoden festgestellt werden ($R=0,78$; $p < 2,2 \times 10^{-16}$). Beim Vergleich der Intra-Paar-Differenzen zwischen Nacht- und Tagarbeitern aus beiden Methoden (siehe Abbildung 40B) betrug die Gesamtkorrelation hingegen nur $R=0,28$ ($p < 0,05$). Insgesamt wurde mittels HM450K BC überwiegend ein höherer Methylierungsgrad berechnet als mit der BSPS.

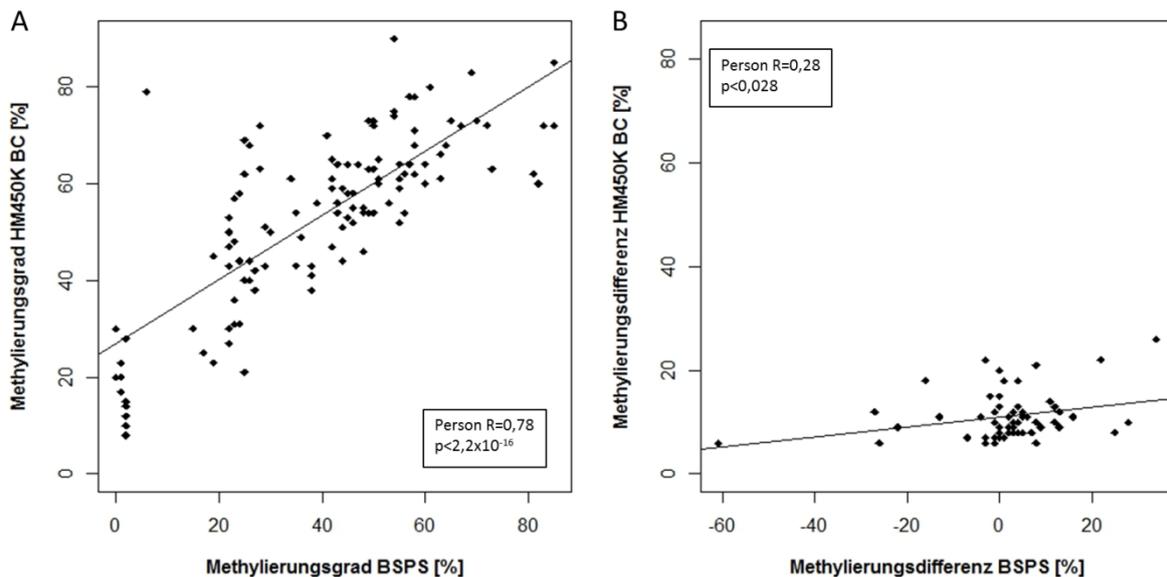


Abbildung 40: Validierung der HM450K BC Methylierungsdaten. (A) Korrelation der Methylierungswerte HM450K BC und Bisulfit-Pyrosequenzierung (BSPS). Die Datenpunkte aller 10 Loci in 12 Zwillingen werden jeweils durch eine Raute symbolisiert. Auf der x-Achse sind die einzelnen Methylierungswerte der Bisulfit-Pyrosequenzierung, auf der y-Achse die entsprechenden Methylierungswerte der HM450K BC Analyse aufgetragen. Der mittels Bisulfit-Pyrosequenzierung ermittelte Methylierungsgrad wurde zum besseren Vergleich mit 100 multipliziert. **(B) Korrelation der Intra-Paar-Methylierungsdifferenzen HM450K BC und BSPS.** Die Datenpunkte aller 10 Loci in 6 Zwillingspaaren werden jeweils durch eine Raute symbolisiert. Auf der x-Achse sind die Methylierungsdifferenzen (nachtarbeitender Zwilling–tagarbeitender Zwilling) der Bisulfit-Pyrosequenzierung, auf der y-Achse die entsprechenden Methylierungsdifferenzen der HM450K BC Analyse aufgetragen. Die mittels Bisulfit-Pyrosequenzierung ermittelte Methylierungsdifferenz wurde zum besseren Vergleich mit 100 multipliziert.

Da die zehn Genorte nicht über eine für multiples Testen angepasste Statistik ermittelt wurden, kann nicht ausgeschlossen werden, dass es sich bei diesen um falsch-positive Loci handelte, was für die hier nicht erfolgreich durchgeführte Validierung spricht.

4.2.8 Zusammenfassung der DNA-Methylierungsanalysen

Die einzelnen Fragestellungen betreffend, konnten mit den jeweiligen Analysen zahlreiche Gene identifiziert werden, die teilweise von mehreren benachbarten differentiell methylierten CpGs abgedeckt wurden. In Abbildung 41 sind die differentiell methylierten Gene aus den einzelnen DNA-Methylierungsanalysen in einem Venn-Diagramm dargestellt. Mit dem Alter korrelierten dabei insgesamt 63, mit einem hohen BMI 43, mit dem Rauchen von Zigaretten 68 und mit der Nachtarbeit elf differentiell methylierte Gene.

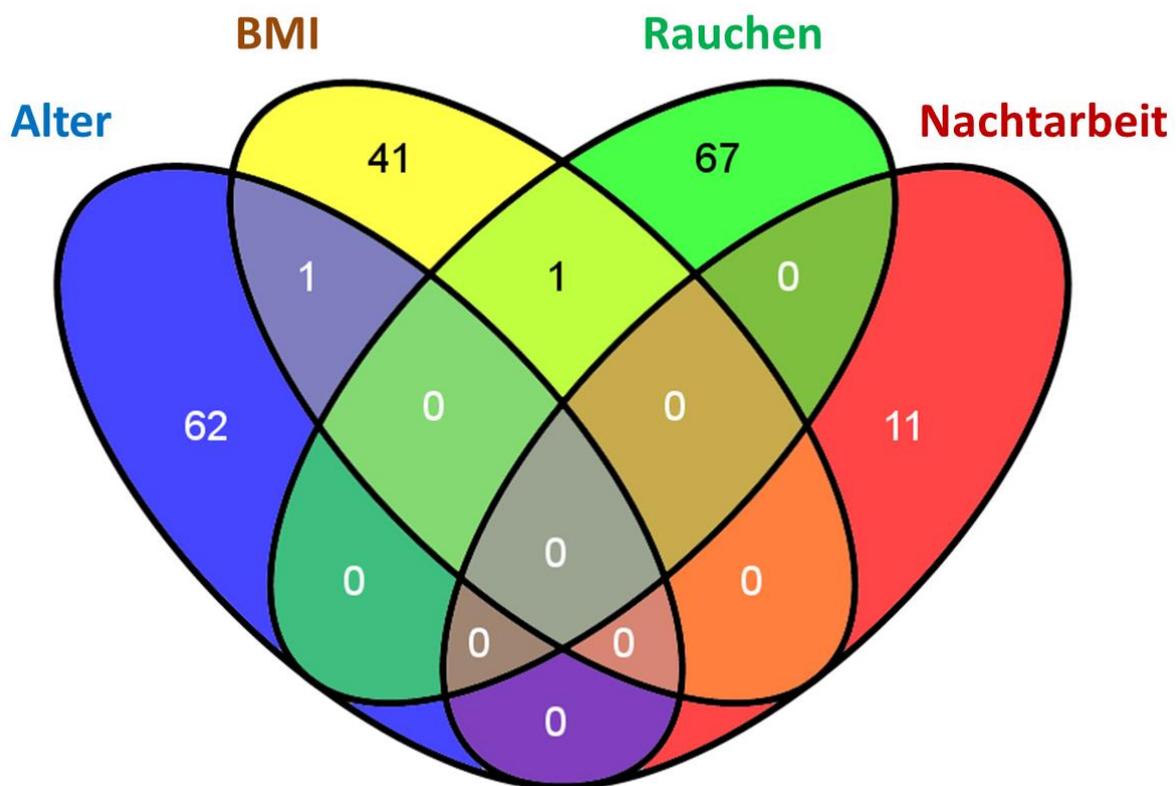


Abbildung 41: Darstellung des Vergleichs der differentiell methylierten Gene aus den vier DNA-Methylierungsanalysen. Abbildung erstellt mit Hilfe von <http://bioinfogp.cnb.csic.es/tools/venny/>.

Insgesamt konnte ein differentiell methyliertes Gen sowohl mit einem hohen BMI als auch mit einem hohen Alter sowie ein weiteres differentiell methyliertes Gen mit einem hohen BMI und gleichzeitig mit dem Rauchen assoziiert werden. Bei den zwei identifizierten Genen handelte es sich um *TNXB* (*tenascin XB*) und *LMO7* (*LIM domain 7*).

TNXB zeigte beim adipösen Phänotyp in einem CpG-Locus in der 5'-UTR eine Hypomethylierung von 5,8% im Vergleich zum normalgewichtigen Phänotyp, während ein

Ergebnisse

etwa 74kb entferntes CpG-Dinukleotid im Genkörper von *TNXB* im fortschreitendem Alter eine Hypermethylierung aufwies.

LMO7 zeigte beim adipösen Phänotyp in einem CpG-Locus im Genkörper von *TNXB* eine Hypomethylierung von 6,8% im Vergleich zum normalgewichtigen Phänotyp und bei Rauchern ebenfalls im Genkörper in einem etwa 5kb entfernten CpG-Locus eine Hypomethylierung von 5,4% im Vergleich zu den Nichtrauchern.

Zusammenfassend lässt sich daraus ableiten, dass die Auswirkungen der in dieser Arbeit untersuchten Parameter auf die genomweite DNA-Methylierung zwar mit Veränderungen des Lebensstils zusammenhängen, die analysierten Einflussfaktoren aber auf unterschiedlichen biologischen Prozessen basieren und somit wenig gemeinsam haben. Dies spiegelt sich ebenfalls auf der Ebene des DNA-Methyloms in Form von spezifischen Veränderungen wider, da in allen hier durchgeführten Teilstudien unterschiedliche Loci von der differentiellen Methylierung betroffen sind. Dennoch konnte eine Überschneidung der differentiell methylierten Gene *TNXB* aus der Alters- und BMI-Analyse sowie *LMO7* aus der BMI- und Rauchstudie identifiziert werden. Dies deutet darauf hin, dass ein hoher BMI bzw. Adipositas durchaus mit DNA-Methylierungsveränderungen von Genen einhergehen könnte, die ebenfalls mit dem Rauchen und dem Alterungsprozess korrelieren könnten.

5 Diskussion

In dieser Arbeit wurde der Einfluss unterschiedlicher Lebensstile auf die genomweite DNA-Methylierung monozygoter und dizygoter Zwillinge unter Verwendung des HM450K BC untersucht. Nach Etablierung des Arrays sollten so mittels verschiedener Auswertestrategien spezifische, den Lebensstil betreffende Fragestellungen analysiert werden. So konnten für den Konsum von Zigaretten signifikante Veränderungen im DNA-Methylierungsmuster bestimmter CpGs zwischen Rauchern und Nichtrauchern identifiziert werden. Ebenso konnten signifikante Veränderungen in der Methylierung einiger CpGs mit einem hohen BMI ($\geq 30 \text{ kg/m}^2$) assoziiert werden. Bezüglich des Einflusses von Schicht- bzw. Nachtarbeit konnten zwar differentiell methylierte Loci zwischen den korrespondierenden diskordanten Zwillingen ermittelt werden, die Unterschiede in der DNA-Methylierung fielen aber sehr gering aus bzw. waren aufgrund der geringen Gruppengröße und der damit verbundenen fehlenden Signifikanz nach Korrektur für multiples Testen mit einer unabhängigen Methode nicht validierbar. Weiterhin konnte sowohl eine mit dem chronologischen Alter assoziierte signifikante Zunahme als auch Abnahme der DNA-Methylierung für ausgewählte Loci im gesamten Studienkollektiv beobachtet werden.

Im Folgenden wird zunächst die Validierung des HM450K BC und der BSPS diskutiert und im Anschluss daran die Ergebnisse der einzelnen DNA-Methylierungsanalysen.

5.1 Validierung des Infinium HumanMethylation450 BeadChip und der Bisulfit-Pyrosequenzierung

Es konnte gezeigt werden, dass die Ergebnisse der mittels HM450K BC generierten Methylierungsdaten gut mit denen der für die Validierung eingesetzten BSPS korrelierten, obwohl für den mittleren Methylierungsbereich eine schwächere Korrelation beobachtet wurde [185]. Als mögliche Fehlerquelle für die im mittleren Methylierungsbereich beobachtete Abweichung zwischen HM450K BC und BSPS wurde die potentielle Anwesenheit von SNPs in Betracht gezogen. Daher wurden alle Assays, deren Methylierungswerte sich in den Bland-Altman-Diagrammen außerhalb des 95% Konfidenzintervalls befanden, auf SNPs hin überprüft. Als Kontrollen dienten 20 willkürlich gewählte Assays, die zwischen beiden Methoden eine sehr hohe Übereinstimmung zeigten.

Diskussion

Insgesamt konnte zwischen dem 95% Konfidenzintervall und der potentiellen Anwesenheit von SNPs jedoch keine deutliche Korrelation identifiziert werden.

Einige publizierte Arbeiten [51,52,193] bestätigten die eigenen Beobachtungen hinsichtlich des DNA-methylierungsbasierten Methodenvergleichs an Zelllinien. Auch in diesen Studien wurde deutlich, dass die sehr hohe Korrelation ($R^2=0,97$; $R^2=0,97$; $R^2=0,94$ bzw. $R^2=0,87$) zwischen zwei unabhängigen Methoden (HM450K BC vs. HM27K BC; HM450K BC vs. BSPS; HM450K BC vs. HM27K BC bzw. HM450K BC vs. Illumina GoldenGate DNA Methylation BeadArray) von der für Zelllinien typischen dichotomen Verteilung der Methylierungswerte abhängig war. Somit wurde auch in diesen Arbeiten der hohe Korrelationskoeffizient insbesondere von den hohen Korrelationen bei niedrigen und hohen Methylierungswerten, die von der überwiegenden Mehrzahl der Loci stammten, beeinflusst. In den Punktdiagrammen, die den Vergleich zweier Methoden darstellten, dominierten die Methylierungswerte wie in der eigenen Arbeit (vgl. Abbildung 13A) im unteren (0%) und oberen (100%) Methylierungsbereich. Darüber hinaus kann der Arbeit von Bibikova *et al.* [51] weiterhin entnommen werden, dass die Methylierungswerte einer humanen Krebsgewebe-Probe im durchgeführten Methodenvergleich (HM450K BC vs. Bisulfitsequenzierung) eine hohe Variabilität (im Mittel 30–40% Unterschied im Methylierungsgrad) zeigten, die ebenfalls im unteren und oberen Methylierungsbereich (0% und 100%) präsent war. So wurde mittels Bisulfitsequenzierung für viele Loci eine Methylierung nahe 0% gemessen, während das Methylierungsniveau laut Array-Analyse bei Bibikova *et al.* für diese Loci zwischen 0% und 50% betrug. In der eigenen Arbeit wurde für die Loci, die eine über die BSPS ermittelte Methylierung nahe 0% aufwiesen, mittels HM450K BC ein Methylierungsniveau zwischen 0% und 20% bestimmt. Weiterhin betrug die mittels Bisulfitsequenzierung gemessenen Methylierungswerte nahe 100% bei Bibikova *et al.* in den Messungen mit dem HM450K BC zwischen 40% und 100%. In der eigenen Arbeit wurden keine Loci ausgewählt, deren Methylierungsniveau bei 100% lag. Der geringere Unterschied in der Variabilität in der eigenen Arbeit (20% vs. 50%) könnte darin begründet liegen, dass Illumina ihren Methodenvergleich für humanes Krebsgewebe an 48809 CpG-Loci durchführten, während für den eigenen Vergleich 352 CpGs verwendet wurden.

5.2 Altersassoziierte Veränderungen des DNA-Methylierungsmusters

In dieser genomweiten Analyse des DNA-Methylierungsmusters von 116 Zwillingsproben im Alter von 20–63 Jahren konnten insgesamt 111 CpG-Positionen auf den Chromosomen 1–20 identifiziert werden, die mit 63 Genen assoziiert wurden und signifikant ($q < 1 \times 10^{-11}$, FDR) mit dem Alter korrelierten. Sie wiesen altersbedingte gerichtete Veränderungen in ihrer Methylierung auf, wobei in 82% der betroffenen Loci mit zunehmendem Alter eine Zunahme der DNA-Methylierung dominierte.

In vielen Studien wurde mit fortschreitendem Alter eine Abnahme der genomweiten DNA-Methylierung beschrieben, die in verschiedenen Wirbeltierarten, Organen und Zelltypen beobachtet werden konnte [151,230–235]. Der Methylierungsverlust scheint dabei hauptsächlich auf der Demethylierung transposabler und repetitiver Elemente sowie endogener Retroviren zu beruhen, die in ausdifferenzierten und gesunden Zellen normalerweise durch die DNA-Methylierung stillgelegt sind [234,235]. Da auf dem HM450K BC auch einige CpGs präsent sind, die in repetitiven Sequenzen liegen, wurden die 111 altersassoziierten CpG-Loci auf Lokalisation innerhalb dieser Regionen überprüft. Von den mit einer mit zunehmendem Alter verminderten Methylierung einhergehenden Loci ($n=20$) konnte ein CpG-Locus (cg12899747) mit einer repetitiven Region assoziiert werden. Von den 91 mit fortschreitendem Alter stärker methylierten CpG-Loci konnte hingegen kein Loci in der untersuchten Region ermittelt werden.

Die Auswirkungen der DNA-Methylierung auf die Genexpression werden von der Lokalisation der betroffenen CpG-Loci bestimmt. So wird eine Methylierung von CpG-Inseln im Promotorbereich mit einer langfristigen Genrepression in Verbindung gebracht, während eine Methylierung im Genkörper positiv mit einer Genexpression korreliert [236–238]. Von den 111 altersassoziierten CpG-Loci lagen acht CpGs in Promotor-assoziierten Bereichen, von denen 75% in CpG-Inseln lokalisiert waren und dort alle (100%) mit fortschreitendem Alter eine Zunahme in der Methylierung zeigten und somit mit einer verminderten Genexpression in Verbindung gebracht werden könnten. Ferner lagen 33% der 111 altersassoziierten CpG-Stellen in Genkörpern, von denen wiederum 73% in CpG-Inseln lokalisiert waren und dort alle (100%) eine altersbedingte Abnahme in der Methylierung aufwiesen, was mit einer erhöhten Expression einhergehen könnte. Weiterhin wird analog zur Promotormethylierung eine Methylierung im ersten Exon mit einer transkriptionellen Repression in Verbindung

Diskussion

gebracht [237,239]. 11/111 CpGs lagen jeweils im ersten Exon der entsprechenden Gene, wobei 9/11 Loci jeweils in CpG-Inseln lagen und insgesamt 10/11 CpGs eine altersabhängige Zunahme der Methylierung zeigten.

Insgesamt konnte eine signifikante Anreicherung von stärker methylierten CpG-Positionen in CpG-Inseln während des chronologischen Alterns ermittelt werden. Dies unterstützt die Ergebnisse vorangegangener Studien, in denen in verschiedenen Geweben gesunder Erwachsener ebenfalls eine altersabhängige Methylierungszunahme präferentiell in CpG-Inseln gefunden wurde [141,146,240].

Für insgesamt 58 der in peripheren Blutzellen identifizierten 111 CpG-Stellen konnte bereits eine Korrelation mit dem Alter sowohl in Vollblut als auch in isolierten Leukozyten sowie CD4⁺- und CD14⁺-Zellen, Wangenzellen, Nabelschnurblut und Speichel beobachtet werden [141,148,150,154,155,238,240]. Dies deutet darauf hin, dass somit mindestens 52% (58/111) der hier identifizierten altersassoziierten CpG-Positionen in verschiedenen Gewebetypen und Zellen von Individuen verschiedenen Alters, Geschlechtes und Ethnie konserviert zu sein scheinen und dass diese Veränderungen mit biologischen Prozessen einhergehen, die im Alterungsprozess eine wesentliche Rolle spielen könnten.

Insgesamt konnten hier demnach 53 neue altersassoziierte CpG-Loci identifiziert werden, von denen 15 CpGs keinem annotierten Gen zugeordnet werden konnten. Die verbliebenen 38 CpG-Positionen konnten mit 33 Genregionen assoziiert werden. Die sich mit dem Alter verändernde DNA-Methylierung bestätigte für 15 der 33 Genregionen, die von insgesamt 19 differentiell methylierten CpG-Loci abgedeckt wurden, die Resultate der o.g. Studien, obwohl in diesen nicht die hier identifizierten CpG-Positionen, sondern benachbarte CpG-Loci ermittelt wurden. Darüber hinaus konnten unter den verbliebenen 18 Genregionen zwei Genregionen (*GRM2*, *ZNF57*) ermittelt werden, in denen zu den schon publizierten CpG-Stellen zusätzlich in dieser Arbeit insgesamt drei bisher nicht publizierte benachbarte CpG-Positionen identifiziert wurden. Zusammenfassend wurden in der hier durchgeführten Analyse 16 neue CpG-Stellen identifiziert, die mit 16 ebenfalls neuen Genorten (*ANKRD43*, *B3GNT9*, *C6orf174*, *HPCAL4*, *IGSF11*, *LDLRAD3*, *LHFPL4*, *LHX3*, *LMTK3*, *LOC729991-MEF2B*, *LRRC24*, *MANEAL*, *MGC70857*, *NXP1*, *RGS22*, *TBR1*) assoziiert werden konnten, an denen sich eine altersabhängige Methylierung abzeichnete. Während in 14 der genannten 16 Genorte eine altersabhängige Zunahme der Methylierung beobachtet wurde, wurde für die

Genorte *IGSF11* und *LDLRAD3* eine Abnahme der Methylierung ermittelt. Die beiden weniger methylierten Loci waren dabei im Genkörper lokalisiert und könnten mit fortschreitendem Alter mit einer schwächeren Expression assoziiert sein. Während 12/14 mit zunehmendem Alter stärker methylierten Loci mit einer erhöhten Genexpression korrelieren könnten, da sie jeweils in CpG-Inseln (11/12) und im S-Shore-Bereich einer CpG-Insel (1/12) im Genkörper lagen, würden hingegen 2/14 Loci mit einer verminderten Expression korrelieren, da diese in CpG-Inseln in der 5'-UTR und im ersten Exon lokalisiert waren. Demnach könnte eine altersabhängige differentielle Methylierung für insgesamt vier der 16 neu identifizierten Genorte mit einer schwächeren und für 12 Genregionen mit einer erhöhten Genexpression korrelieren.

5.2.1 Eigenschaften der neu identifizierten mit zunehmendem Alter schwächer methylierten Genorte

Über das Protein *LDLRAD3* (*low density lipoprotein receptor class A domain containing 3*) ist in der Literatur bisher nicht viel beschrieben. Es wird in Neuronen exprimiert, wo ihm ferner modulierende Eigenschaften in der Entstehung und Entwicklung der Alzheimerkrankheit zugesprochen werden, die mit einer erhöhten Expression von *LDLRAD3* korrelieren [241].

Das Gen *IGSF11* (*immunoglobulin superfamily, member 11*) wird vorrangig in Gehirn und Hoden exprimiert und spielt eine Rolle in Entwicklung und Funktion des ZNS [242]. Eine erhöhte *IGSF11*-Expression konnte sowohl in kolorektalen und hepatozellulären Karzinomen als auch in Magenkrebs beobachtet werden [243,244].

Die identifizierten Assoziationen der beiden Gene würden zwar zum Bild altersabhängiger Erkrankungen passen, die Änderung des Methylierungsgrads in der hier durchgeführten Analyse (jeweils mit zunehmendem Alter eine Abnahme der Methylierung des jeweiligen CpG-Locus in Intron 1 von *IGSF11* bzw. in Intron 3 von *LDLRAD3*) spricht hingegen für eine verminderte Expression. Aufgrund fehlender funktioneller Analysen kann derzeit nicht zweifelsfrei bestätigt werden, ob die differentielle Methylierung in den betroffenen Loci zu einer verminderten bzw. überhaupt zu einer Änderung in der Expression führt.

5.2.2 Eigenschaften der neu identifizierten mit zunehmendem Alter stärker methylierten Genorte

Für sechs der 14 hypermethylierten Gene (*ANKRD43* (*SOWAHA*), *B3GNT9*, *C6orf174* (*SOGA3*), *MANEAL*, *MGC70857* (*C8orf82*) und *TBR1*) konnten keine Informationen zu Funktion, Schlüsselrolle oder spezifischen Assoziationen gefunden werden. Für die übrigen acht Gene konnte zumindest einer der genannten Aspekte erfüllt werden. Die Zunahme der Methylierung in den CpG-Loci dieser Gene könnte dabei mit einer erhöhten Expression einhergehen.

Die spezifische Funktion von *HPCAL4* (*hippocalcin like 4*) ist bisher nicht bekannt, eine erhöhte Expression konnte aber bereits im Lungenkarzinom beobachtet werden sowie in Zelllinien nach Exposition gegenüber Nikotin [245,246]. Des Weiteren wurde die Chromosomenregion, in der *HPCAL4* lokalisiert ist, mit vielen Krebsentitäten assoziiert, wie Brust- und Lungenkrebs, Neuroblastom, Phäochromozytom und Kolorektalkrebs [245].

Das Gen *LHFPL4* (*lipoma HMGIC fusion partner-like 4*) gehört zur Familie der *tetraspan* Transmembranproteine, in der Mutationen in *LHFPL5* bereits mit Gehörverlust im Menschen assoziiert wurden [247]. Eine Hypermethylierung in CpG-Inseln innerhalb des *LHFPL4*-Gens konnte in Gebärmutterhalskrebszellen beobachtet wurde [248].

LHX3 (*LIM homeobox 3*) kodiert für einen Transkriptionsfaktor, der für die Entwicklung der Hypophyse und motorischen Neuronen benötigt wird und auch im menschlichen Gehörsystem exprimiert wird. Mutationen in diesem Gen konnten u.a. mit einer Verschlechterung des Gehörsinns assoziiert werden [249,250].

LMTK3 (*Lemur tyrosine kinase-3*) gehört zur Familie der Serin-Threonin-Tyrosin Kinasen. Eine erhöhte Expression wurde bereits mit einer Vielzahl verschiedener Krebsentitäten assoziiert, so wie für Brust- [251,252], Lungen- [253], Magen- [254] und Kolorektalkrebs [255]

LOC729991-MEF2B gehört zur Familie der MEF2-Transkriptionsfaktoren (*myocyte enhancer factor-2*), die wichtige Rollen bei der Entwicklung und Funktion von T-, Nerven- und Muskelzellen spielen [256]. Eine Aktivierung von MEF2 konnte in der Reaktion des Herzens auf Stress beobachtet werden, die in einer Herzmuskelhypertrophie resultierte und mit einem erhöhten Risiko für Morbidität und Mortalität verbunden war [257].

Eine spezifische Funktion für das Gen *LRRC24* (*leucine rich repeat containing 24*) ist bisher nicht bekannt, es konnte jedoch mit einer Beteiligung an der Pathogenese altersrelevanter neurodegenerativer Erkrankungen in Zusammenhang gebracht werden [258].

Eine Hypermethylierung in einer CpG-Insel im Genkörper von *Neurexophilin 1* (*NXPH1*) konnte in Brustkrebsgewebe im Vergleich zu normalem Gewebe beobachtet werden [259]. Ferner konnte in intraduktalen papillär-muzinösen Neoplasien (IPMN) des Pankreas mit hochgradiger Dysplasie eine höhere CpG-Insel-Methylierung beobachtet werden als für IPMNs mit niedriger Dysplasie [260].

RGS22 (regulator of G-protein signaling 22) ist ein wenig charakterisiertes Gen, das in vielen Tumortypen, insbesondere in epithelialen Tumoren, exprimiert wird. Aufgrund der in Krebszelllinien aus einer Überexpression resultierenden Hemmung des Zellwachstums, wurde *RGS22* als Tumorsuppressorgen vorgeschlagen [261].

Insgesamt konnten die hier aufgeführten neu identifizierten Gene, deren locusspezifische Methylierungszunahme mit dem Alter signifikant korrelierte, mit möglicherweise altersbedingten Prozessen in Verbindung gebracht werden, die Erkrankungen, einschließlich Krebserkrankungen beinhalteten. Dabei stimmte die zu der veränderten Methylierung möglicherweise korrelierende Expression mit dem Expressionsmuster in den genannten Tumoren überein. Dies unterstützt die Ansicht, dass eine altersabhängige Methylierung einiger CpG-Inseln ein früher Schritt in der Karzinogenese sein könnte [262].

5.2.3 Signifikante Anreicherungen altersassoziierter Gene in biologischen Prozessen

Insgesamt konnten die einer altersabhängigen Methylierung unterliegenden Gene (n=63) verschiedenen biologischen Prozessen zugeordnet werden, wobei einige GO-Kategorien in Bereichen der Blutzirkulation, des Herzprozesses und des Nervensystems signifikant überrepräsentiert vorlagen.

5.2.3.1 Anreicherung altersassoziierter Gene in Prozessen des Blutdrucks und der Herzkontraktion

In den GO-Kategorien „Positive Regulation des systemischen arteriellen Blutdrucks“, „Regulation der Herzkontraktion“ und „Regulation der Herzfrequenz“ konnte eine signifikante Anreicherung der identifizierten altersassozierten Gene ermittelt werden. Mit zunehmendem Alter steigt das Risiko einer verminderten Schlagkraft des Herzens sowie einer Hypertonie, was auf altersbedingte Veränderungen an Herz und Blutgefäßen zurückzuführen ist [263]. Insgesamt konnten die in den hier identifizierten Prozessen des Blutdrucks und der Herzkontraktion involvierten altersassozierten Gene (*ADRB1*, *AVPR1A*, *CACNA1G*, *DES*) bereits mit einer altersabhängig veränderten Expression assoziiert werden [264–270]. An dieser altersbasierten Veränderung des Expressionsmusters könnte die Epigenetik, speziell die DNA-Methylierung, demnach eine wesentliche Rolle spielen.

5.2.3.2 Anreicherung altersassoziierter Gene in Prozessen des Nervensystems

Die Alterung bewirkt u.a. in Nervenzellen eine Senkung in der Fähigkeit, Nervenimpulse zu übertragen [271]. Am hier ermittelten überrepräsentierten Prozess der Übertragung eines Nervenimpulses waren insgesamt zehn Gene beteiligt, von denen die für die Gene *ADRB1*, *GLRA1* und *GRIK5* kodierenden Proteine die Regulation des postsynaptischen Membranpotenzials vermittelten. Des Weiteren waren neun Gene im Prozess der Entwicklung des Zentralnervensystems signifikant angereichert, wovon sieben Gene (*AVPR1A*, *BOK*, *CNTNAP2*, *IRS2*, *LHX3*, *TBR1*, *ZIC1*) in der Entwicklung des Gehirns involviert waren und somit eine Rolle in der Gehirnalterung spielen könnten, die durch einen altersbedingten Rückgang der kognitiven Fähigkeiten charakterisiert ist [272].

Eine altersabhängige Veränderung in der Regulation einer bestimmten Zellantwort könnte aufgrund epigenetischer Modifikationen der betreffenden Gene zur Entwicklung altersbedingter neurologischer Erkrankungen beitragen.

5.2.4 Stärken und Limitationen der Alters-Analyse

Eine Stärke der altersassozierten DNA-Methylierungsanalyse ist, im Gegensatz zu den nachfolgenden Methylierungsanalysen, die relativ große Fallzahl ($n=116$). Hierfür konnten alle rekrutierten Zwillinge der SAME- und GEMINAKAR-Studie, welche die Qualitätskontrolle

Diskussion

erfolgreich durchliefen, berücksichtigt werden. Die hohe Anzahl der Probanden wirkte sich hier letztendlich maßgeblich auf die Signifikanz der Ergebnisse aus.

Eine mögliche Limitation der durchgeführten Methylierungsanalyse liegt im potentiellen Einfluss der zellulären Heterogenität des Blutes begründet, die primär auf altersabhängigen Unterschieden in der Zellzusammensetzung des peripheren Blutes beruht. So verändert sich mit zunehmendem Alter beispielsweise die Anzahl der Lymphozyten und natürlichen Killerzellen [204,273]. Die interindividuelle Variation in der Anzahl spezifischer Blutzellen, wie den Leukozyten, kann somit zusammen mit den zellspezifischen Methylierungsprofilen die Ergebnisse der DNA-Methylierungsanalyse beeinträchtigen [274]. Talens *et al.* konnten in ihrer Studie für die Mehrheit der ausgewählten Kandidaten-Loci zeigen, dass die interindividuellen Unterschiede in der Blutzellzusammensetzung keinen Effekt auf die Variation in der DNA-Methylierung ausübten bzw. nur einen geringen Anteil dieser Variation erklärten [275]. Dies konnten Bell *et al.* für ihre Studie ebenfalls bestätigen, indem sie für ihre 490 identifizierten altersassoziierten Loci keine signifikanten Assoziationen zwischen DNA-Methylierung und Mengenverhältnis von Neutrophilen, Eosinophilen oder Monozyten beobachten konnten. Für 3,9% dieser Loci (n=19) konnten sie hingegen eine signifikante Assoziation zwischen DNA-Methylierung und Lymphozytenzahl nachweisen, was darauf hindeutete, dass die Effekte an diesen Loci zum Teil die Variabilität in der Anzahl der Lymphozyten während des Alterns reflektieren könnten [240]. Weidner *et al.* konnten wiederum in ihrer Studie zeigen, dass ihr aus Vollblut erhobenes Altersvorhersagemodell ebenso in isolierten Zellpopulationen (CD4, CD14, CD16 und CD34) anwendbar war und das Mengenverhältnis von Monozyten, Lymphozyten, Neutrophilen, Basophilen oder Eosinophilen nicht mit dem vorhergesagten Alter korrelierte [154]. Da in dieser Studie keine Informationen über die zelluläre Heterogenität der jeweiligen Blutproben vorlagen, konnte für diese nicht korrigiert werden. Dieser Aspekt betraf alle hier durchgeführten DNA-Methylierungsanalysen.

Letztendlich lässt sich in Bezug auf die hier ermittelten angereicherten Prozesse schwer ableiten, in welchem Ausmaß die in Blutzellen gemessene differentielle DNA-Methylierung tatsächlich Auswirkungen auf diejenigen Gewebe hat, die in den identifizierten Prozessen direkt involviert sind und daher als potenzieller Marker genutzt werden könnte. Auch dieser Punkt muss bei allen hier erhobenen Resultaten berücksichtigt werden.

5.2.5 Verifizierung der Alterssignatur basierend auf der Arbeit von Steve Horvath

Die von Steve Horvath zur Bestimmung des epigenetischen Alters identifizierte 353 CpG-Signatur [152] wurde im Rahmen dieser Arbeit in dem hier verwendeten Zwillingskollektiv (n=116 Proben) überprüft und konnte mit einer sehr hohen Korrelation ($R=0,92$) zwischen chronologischem und epigenetischem Alter verifiziert werden. Unter den in dieser Arbeit identifizierten 111 altersabhängigen CpG-Loci konnte ein CpG-Locus (cg22736354) ermittelt werden, der ebenfalls zu der von Horvath identifizierten 353 altersassoziierten CpG-Signatur gehörte und sowohl bei Horvath als auch in dieser Arbeit mit fortschreitendem Alter eine Zunahme der Methylierung zeigte.

Innerhalb der einzelnen Zwillingspaare konnte kein einheitliches Muster in Form einer Diskordanz bezüglich verschiedener Umweltfaktoren festgestellt werden, mit der erklärt werden könnte, weshalb ein Zwilling innerhalb eines Paares epigenetisch älter oder jünger wirkte als der zweite Zwilling. Ferner konnte keine Assoziation zwischen den chronologischen Altersgruppen und einem älteren bzw. jüngeren epigenetischen Alter beobachtet werden. So wurde für die Zwillinge aus allen vier Altersquartilen sowohl ein jüngeres als auch ein älteres epigenetisches Alter bestimmt.

Des Weiteren war das Ausmaß der Intra-Paar-Unterschiede des ermittelten epigenetischen Alters von Zwillingspaaren, die zum Zeitpunkt der Blutentnahme das gleiche chronologische Alter hatten, von der jeweiligen Altersklasse unabhängig. So unterschied sich beispielsweise das epigenetische Alter von 30- bzw. 47-jährigen Zwillingspaaren (n=2 bzw. n=4) zwischen ein und vier Jahren, während die epigenetische Altersdifferenz für 31-jährige bzw. 37-jährige Zwillingspaare (n=2 bzw. n=3) zwischen drei und fünf Jahre, für 39-jährige Paare (n=3) zwischen drei und acht Jahre und für 52-jährige Paare (n=2) hingegen jeweils nur ein Jahr betrug.

Hingegen unterschieden sich die Intra-Paar-Differenzen des ermittelten epigenetischen Alters von MZ-Paaren zu zwei verschiedenen Zeitpunkten (GEMINAKAR und SAME) signifikant voneinander ($p<0,0092$; t-Test für gepaarte Stichproben). Dies unterstreicht zum einen, dass ältere MZPs epigenetisch unterschiedlicher sind als jüngere MZPs [143] und zum anderen, dass diese Unterschiede auch in der von Horvath ermittelten epigenetischen

Alterssignatur zu finden sind und sich somit zumindest teilweise auf das epigenetische Alter auswirken.

Abschließend lässt sich folgern, dass die von Horvath identifizierte Alterssignatur zur Ermittlung des epigenetischen Alters zwar eine sehr hohe Korrelation zum chronologischen Alter aufweist, diese aber nicht das reine chronologische Alter des untersuchten Materials widerspiegelt, sondern von äußeren Faktoren beeinflusst werden kann.

5.2.6 Fazit der altersassoziierten DNA-Methylierungsanalyse

Mit der hier durchgeführten genomweiten altersbasierten DNA-Methylierungsanalyse konnten zum einen bereits publizierte mit dem Alter korrelierende differentiell methylierte CpG-Loci und Genorte bestätigt werden. Zum anderen wurden neue Genorte identifiziert, die in altersbedingten Erkrankungen, einschließlich Krebs, eine Rolle spielen könnten.

5.3 Mit hohem BMI korrelierende Veränderungen des genomweiten DNA-Methylierungsmusters

Jüngste Publikationen zeigen, dass die DNA-Methylierung zur Entwicklung von Übergewicht beitragen und eine Verbindung zwischen Übergewicht und den damit assoziierten Folgeerkrankungen liefern könnte [276,277]. Übergewicht und Adipositas sind dabei wichtige Risikofaktoren für einige Erkrankungen einschließlich verschiedener Krebsentitäten [278,279] und Herz-Kreislauf-Erkrankungen [280].

Veränderungen im Methylierungsgrad konnten bereits für Gene gezeigt werden, die Schlüsselrollen in der Glukosehomöostase innehaben [281–283]. Darüber hinaus konnte bereits eine mit dem BMI signifikant korrelierende differentielle Methylierung an CpG-Positionen in der Nähe von Genen beobachtet werden, die an der Regulation des Körpergewichts oder T2D bereits bestätigte Rollen spielen [284,285].

Um der Frage nachzugehen, ob der BMI mit dem DNA-Methylierungsmuster korreliert, wurden im Rahmen dieser Arbeit die ermittelten Extreme des BMIs analysiert, welche hier vom normalgewichtigen und adipösen Phänotyp gebildet wurden. Da für die hier definierten BMI-Extreme unter den 23 Zwillingspaaren der SAME-Studie keine diskordanten Paare zur Verfügung standen, konnte eine auf dem diskordanten Zwillingsmodell basierende Analyse

Diskussion

nicht durchgeführt werden. Daher wurde das genomweite DNA-Methylierungsprofil von insgesamt fünf Zwillingen mit einem BMI ≥ 30 kg/m² dem DNA-Methylierungsprofil von 29 normalgewichtigen Zwillingen mit einem BMI zwischen 18,5 und 24,9 kg/m² gegenüber gestellt. Da die im stringenten Ansatz identifizierten differentiell methylierten CpG-Loci (n=3) auch im weniger stringenten Ansatz ermittelt wurden, wird im Folgenden nur auf die Ergebnisse der zweiten Analyse eingegangen. Die zwischen adipösen und normalgewichtigen Zwillingen ermittelten differentiell methylierten CpG-Loci (n=80) konnten insgesamt 43 annotierten Genorten zugeordnet werden. 4/43 Genorte (*CCDC144NL*, *DPCR1*, *EGFLAM*, *TMEM232*) beinhalteten dabei jeweils zwei und 2/43 Genorte (*CNTNAP1*, *PTPRN2*) jeweils drei benachbarte differentiell methylierte CpG-Positionen, was eine mögliche Assoziation zwischen den genannten Genen und der Adipositas unterstützen könnte.

In Zellen des peripheren Bluts konnten in genomweiten Analysen mittels HM27K BC und HM450K BC bereits Adipositas-assoziierte differentiell methylierte Loci identifiziert werden [286–288], wobei die ermittelten Loci aus dem Vergleich des schlanken mit dem adipösen Phänotyp resultierten. Die differentiell methylierten Loci lagen dabei sowohl in Kandidatengenen für Adipositas als auch in biologischen Funktionen wie Immunantwort, Regulation der Transkription und Zelldifferenzierung angereichert vor. In Anlehnung an den in dieser Arbeit identifizierten 43 Adipositas-assoziierten Genen konnte lediglich das Gen *DPCR1* in der Studie von Wang *et al.* in Leukozyten adipöser vs. schlanker Jugendlicher und junger Erwachsener unter den zehn signifikantesten Genen bestätigt werden [287]. Die zwei von Wang *et al.* in diesem Gen identifizierten differentiell methylierten CpG-Positionen verloren nach Korrektur für multiples Testen jedoch ihre Signifikanz. Beide CpG-Loci zeigten in *DPCR1* im adipösen Phänotyp eine Hypomethylierung (7,7%; $p < 6,90 \times 10^{-5}$ bzw. 0,6%; $p < 0,71$) gegenüber den schlanken Kontrollen, welche auch in der vorliegenden Arbeit beobachtet werden konnte (37,9%; $p < 9,65 \times 10^{-7}$ bzw. 10,4%; $p < 2,85 \times 10^{-4}$). Die von Wang *et al.* ermittelten CpG-Loci waren dabei an anderen Positionen im *DPCR1*-Gen als die hier ermittelten Loci lokalisiert. Dies könnte aufgrund mehrerer differentiell methylierter benachbarter CpG-Loci aus zwei unabhängigen Studien auf eine Beteiligung des *DPCR1*-Gens an der Pathologie von Adipositas schließen lassen. Die von Xu *et al.* [288] in Zellen des peripheren Bluts ermittelten zwischen adipösen und schlanken afroamerikanischen Jugendlichen und jungen Erwachsenen 23305 differentiell methylierten CpG-Stellen standen

(auch auf Nachfrage) nicht zur Verfügung, weshalb eine mögliche Überschneidung mit den hier identifizierten BMI-assoziierten Genen nicht ausgeschlossen werden kann.

5.3.1 Hoher BMI ist mit differentieller Methylierung von Loci in Genen, die in Prozessen des Nervensystems involviert sind, assoziiert

Unter den ermittelten 43 differentiell methylierten Genorten konnten über eine GO-Analyse insgesamt acht Gene (*ADORA1*, *CNTNAP1*, *COBL*, *EGFLAM*, *FABP7*, *PSEN1*, *PTPRN2* und *SYN2*) mit verschiedenen Kompartimenten der Nervenzelle (neuronaler Zellkörper, Axon, Endknöpfchen, Synapse) in Verbindung gebracht werden. Dies könnte auf eine Beteiligung an der Signalweiterleitung und Signalübertragung schließen lassen, zudem andere Arbeiten bereits Fälle von Adipositas zeigen konnten, die auf eine Dysfunktion im Hypothalamus zurückzuführen waren [289]. Die Nervenzellen stellten somit keine Signale mehr bereit, die unter physiologischen Bedingungen zur Beständigkeit der Körpermasse beitragen. In weiteren Studien konnte außerdem gezeigt werden, dass es unter bestimmten Umweltbedingungen und genetischen Konstellationen zum Verlust des Gleichgewichts zwischen Kalorienaufnahme und Energieverbrauch kommen kann, woraufhin sich die Adipositas ausbilden kann (zusammengefasst in [289]). Eine Beteiligung an solchen bzw. ähnlichen Signalkaskaden der hier identifizierten acht Gene wäre daher vorstellbar.

Ferner wurden bereits Beweise für einen direkten neuronalen Einfluss auf die Lipolyse des weißen Fettgewebes des Menschen gefunden [290]. Weiterhin konnte bereits gezeigt werden, dass adipöse Kinder und Jugendliche eine verminderte Aktivität des autonomen Nervensystems zeigten und daher Funktionsstörungen des autonomen Nervensystems durch Adipositas verursacht werden könnten [291].

Insgesamt wäre daher eine epigenetische Regulation spezifischer neurologischer Prozesse an der Ausbildung bzw. dem Voranschreiten der Adipositas vorstellbar, an denen die hier identifizierten Gene durchaus eine Rolle spielen könnten.

5.3.2 Identifizierte mit hohem BMI korrelierende differentiell methylierte Gene mit bekannter Assoziation zu Adipositas

Es konnte bereits für einige Gene mit Funktionen in der Regulierung des Körpergewichts sowie des Lipidmetabolismus, wie *LEP* und *TNF- α* , eine Regulation durch epigenetische

Diskussion

Mechanismen an der Entwicklung von Adipositas nachgewiesen werden [292,293]. Darüber hinaus konnte auch ein Einfluss der Risiko-Genvariante (rs9939609) des Fettmasse- und Adipositas-assoziierten *FTF*-Gens auf den Methylierungsgrad anderer Gene in adipösen Menschen gezeigt werden [286]. Des Weiteren konnte bereits gezeigt werden, dass eine fettreiche Ernährung von fünf Tagen die DNA-Methylierung eines wichtigen transkriptionellen Coaktivators (*PPARGC1A*) von Genen beeinflusste, die im oxidativen Energiemetabolismus involviert sind [294].

Aus diesem Grund wurden alle differentiell methylierten Genorte manuell nach einer bekannten auf Publikationen basierenden Assoziation zu Adipositas selektiert, um eine mögliche Verbindung zur veränderten Methylierung herstellen zu können. Dabei konnten zwei unabhängige Gene auf Chromosom 1 (*ADORA1* und *LPGAT1*) sowie das Gen *POM121C* auf Chromosom 7, welches bereits über die FDR ($q < 0,05$) identifiziert werden konnte, mit der Adipositas in Verbindung gebracht werden.

Der A1-Adenosinrezeptor *ADORA1* ist einer von vier Adenosinrezeptor-Subtypen, die zur Familie der G-Protein-gekoppelten Rezeptoren gehören. Sie werden im menschlichen Körper ubiquitär exprimiert, wobei *ADORA1* besonders stark in Gehirn, Herz und Fettgewebe exprimiert wird [295–297]. Adenosin wird vom subkutanen Fettgewebe kontinuierlich freigesetzt. Die Aktivierung von *ADORA1* in den Adipozyten führt infolge einer Hemmung der Adenylatzyklase und der damit verbundenen Senkung des cAMP-Spiegels zur Hemmung der Lipolyse [217]. Durch den antilipolytischen Effekt kommt es zur Reduktion der freien Fettsäuren im Blutkreislauf, was die Behandlung von Dyslipidämie und T2D begünstigen könnte [217,297,298]. Im adipösen Phänotyp wurde in der hier vorliegenden Studie in cg05160197 des *ADORA1*-Gens im Mittel eine Methylierungsdifferenz von -5,6% gegenüber dem normalgewichtigen Phänotyp in der 3'-UTR identifiziert ($p < 0,001$), die zusätzlich eine DNase I-Hypersensitivität aufwies. In einigen Studien konnte bereits gezeigt werden, dass die Methylierung in 3'-UTRs regulatorische Effekte auf die Genexpression inne hat. So konnte sowohl eine Hyper- als auch Hypomethylierung in diesen genomischen Regionen bestimmter Gene mit einer erhöhten Genexpression assoziiert werden [299–302]. Demnach könnte die hier im Vergleich zum normalgewichtigen Phänotyp identifizierte Hypomethylierung im letzten Intron von *ADORA1* aufgrund einer veränderten Genexpression für die beobachtete gesteigerte Lipolyse-Rate in Fettleibigen verantwortlich sein, die zu der

für diesen Phänotyp charakteristischen Erhöhung von freien Fettsäuren führt, welche wiederum auf Hyperplasie (Erhöhung der Adipozytenzahl) als auch auf Hypertrophie (Vergrößerung der Adipozyten) zurückzuführen sein könnte [303,304]. ADORA1 konnten ferner weitere Funktionen in den Adipozyten nachgewiesen werden, einschließlich einer erhöhten Lipogenese und Lipidakkumulation, Stimulation der Leptin-Ausschüttung und Schutz gegen die mit Adipositas verbundene Insulinresistenz [305,306].

Im CpG-Locus cg13064658, welcher in einer 849bp großen CpG-Insel im Promotorbereich innerhalb der 5'UTR des Lysophosphatidylglycerol Acyltransferase 1-Gens *LPGAT1* liegt, konnte in den adipösen Zwillingen (mittlere Methylierung: 14,0%) im Vergleich zu den normalgewichtigen Zwillingen (mittlere Methylierung: 7,0%) eine Hypermethylierung festgestellt werden ($p < 1 \times 10^{-4}$; t-Test). *LPGAT1* gehört zur Familie der Lysophosphatidsäure-Azyltransferasen. Azyltransferasen sind an einem breiten Spektrum von Stoffwechselwegen einschließlich Fettspeicherung, Nahrungsaufnahme, Insulinsensitivität, Energiehomöostase und Lipogenese involviert [218,307,308]. Ferner konnte bereits beobachtet werden, dass ein Knockdown von *LPGAT1* in der Leber von diabetischen Mäusen u.a. in einer Abnahme von Triglyzeriden und freien Fettsäuren in den Leberzellen sowie in einer Abnahme des Körpergewichts resultierte [309]. Folglich wurde dem Gen *LPGAT1* eine mögliche Beteiligung an der Synthese und Sekretion von Triglyzeriden in der Leber zugesprochen. Des Weiteren konnten beim Menschen bereits einige Varianten im *LPGAT1*-Gen identifiziert werden, die bei Pima-Indianern und anderen Ureinwohnern Nordamerikas mit Adipositas assoziiert werden konnten [218]. Dies unterstützt das Ergebnis, dass epigenetische Modifikationen in der Promotorregion von *LPGAT1* möglicherweise mit Adipositas in Verbindung stehen könnten.

Eine Abnahme um 5,2% im Methylierungsgrad konnte in den adipösen Zwillingen im S-Shore Bereich einer 553bp großen CpG-Insel im Genkörper des *POM121C*-Gens (cg10351692) nachgewiesen werden. *POM121C* ist ein Transmembran-Nukleoporin, das in der Literatur bisher wenig beschrieben wurde. Bei adipösen Kindern aus lateinamerikanischen Familien konnte eine genetische Variante des *POM121C*-Gens beobachtet werden, die sich mit der Pathophysiologie von Fettleibigkeit assoziieren ließ [310]. Somit könnte eine veränderte Methylierung in *POM121C* durchaus für eine mögliche Assoziation zu Adipositas sprechen.

5.3.3 Identifizierte mit hohem BMI korrelierende differentiell methylierte Gene mit Assoziation zu Typ-2-Diabetes (T2D)

Da Übergewicht mit einem erhöhten Risiko für Typ-2-Diabetes verbunden ist, wurden die differentiell methylierten Gene hierfür nach möglichen Assoziationen überprüft, wobei zwei Gene in den Vordergrund traten, die beide auf Chromosom 3 lokalisiert sind: *ADCY5* und *SYN2*.

Das *SYN2*-Gen kodiert für Synapsin 2 und gehört zur Familie der synaptischen Vesikel-assoziierten Phosphoproteine. *SYN2* bindet selektiv an die zytoplasmatische Oberfläche der synaptischen Vesikelmembran im Endknöpfchen präsynaptischer Neuronen, wo es mit der Modulation der Neurotransmitter-Freisetzung und der Synaptogenese in Verbindung gebracht wird [311,312]. Kaukasischen Trägern des SNPs rs17036101 konnte eine erhöhte Anfälligkeit für T2D durch Modulation der Expression des *SYN2*-Gens im Fettgewebe zugeschrieben werden [313]. Kuppuswamy *et al.* schlugen aufgrund der beobachteten Assoziation zu T2D und GO-Annotationen für *SYN2* eine Rolle in der Pathogenese von T2D vor [314]. Das CpG cg13523468 zeigte im Genkörper von *SYN2* in den adipösen Zwillingen (mittlere Methylierung: 80,9%) einen Methylierungsunterschied von -6,4% ($p < 0,001$; t-Test) gegenüber den normalgewichtigen Zwillingen (mittlere Methylierung: 87,3%).

ADCY5 kodiert für die Adenylatzyklase 5, die die Bildung von cAMP katalysiert. Da cAMP wiederum die Aktivierung der Proteinkinase A vermittelt, die den Kalzium-Einstrom, die anschließende Insulin-Freisetzung und die Transkription des Proinsulins induziert, wäre eine Beteiligung von *ADCY5* an der Ausschüttung von Insulin denkbar [315]. Mittels einer GWAS konnten Varianten in *ADCY5* bereits mit einem geringen Geburtsgewicht [315,316] und T2D einschließlich einer herabgesetzten β -Zell-Funktion [317] assoziiert werden. Der identifizierte differentiell methylierte CpG-Locus cg04908625 liegt in einer 2349bp großen CpG-Insel und ist im 1. Exon von *ADCY5* lokalisiert. In den fettleibigen Zwillingen zeigte dieser eine mittlere Methylierung von 32,8%, in den normalgewichtigen Zwillingen betrug diese 40,8% ($p < 0,001$; t-Test). Eine veränderte Methylierung des *ADCY5*-Gens könnte möglicherweise mit dem infolge einer vermehrten Freisetzung von Insulin hervorgerufenen Erschöpfungszustand der β -Zellen korrelieren.

5.3.4 Limitationen der BMI-Analyse

Die bedeutendste Schwäche dieser Studie ist die geringe Anzahl der adipösen Probanden, die sich nur auf fünf Fälle im Vergleich zu 29 normalgewichtigen Fällen belief. Um die erhobenen Resultate validieren zu können, sollte in einer zukünftigen Studie auf ein größeres Kollektiv zurückgegriffen werden. Somit könnte auch die damit verbundene Schwachstelle dieser Methylierungsanalyse, die mangelnde Signifikanz über die FDR, gelöst werden. Aufgrund der nicht für multiples Testen korrigierten Signifikanz des zweiten Ansatzes ($p < 0,0003$; t-Test; $|\Delta\beta| > 0,05$), sollten die Ergebnisse dieser Analyse vorsichtig interpretiert werden.

Als nächster Punkt ist anzumerken, dass in der hier durchgeführten Analyse davon ausgegangen wurde, dass die infolge der Adipositas differentiell methylierten Genorte zu Folgeerkrankungen der Adipositas beitragen könnten. Da es sich hier um eine Querschnittstudie (*cross-sectional study*) handelte und somit sowohl die Methylierungsdaten als auch der Phänotyp zum selben Zeitpunkt erhoben und gemessen wurden, kann nicht einwandfrei aufgeklärt werden, ob die Assoziation zwischen den ermittelten Veränderungen in der DNA-Methylierung an den ausgewählten Loci und der Adipositas Ursache oder Folge des adipösen Phänotyps ist.

Ferner handelte es sich bei den hier untersuchten adipösen Zwillingen um Zwillinge mittleren Alters, welches zwischen 36 und 48 Jahren lag. Somit bestand die Möglichkeit, dass die Ergebnisse der hier erhobenen Methylierungsstudie mit Folgeerkrankungen der Adipositas, z.B. dem T2D, korrelieren könnten. Da von den dänischen Kooperationspartnern aus dem Blut der Probanden der nüchterne Plasma-Glukose- und -Insulin-Spiegel gemessen wurde, konnte somit mittels HOMA (*homeostasis model assessment*)-Verfahren [318] die Insulinresistenz (IR), das Kennzeichen des T2D [319], berechnet werden. Tabelle 17 zeigt, dass sowohl die Insulin-Level als auch der Grad der IR basierend auf der HOMA-IR in der Gruppe der adipösen Zwillinge signifikant höher war als in der Gruppe der normalgewichtigen Zwillinge.

Diskussion

Tabelle 17: Ermittelte Glukose- und Insulin-Spiegel sowie die daraus berechnete Insulinresistenz (IR) der in der BMI-Analyse eingeschlossenen Probanden.

Merkmal	adipöse Zwillinge (n=5)	normalgewichtige Zwillinge (n=29)	p-Wert
Glukosespiegel [mmol/l]	6,0±0,7 (4,9–8,9)	5,4±0,1 (4,6–6,7)	0,1
Insulinspiegel [μ U/ml]	8,0±1,1 (4,8–11,2)	4,4±0,4 (1,5–9,8)	<0,001
IR (HOMA-IR)	2,2±0,4 (1,2–3,5)	1,1±0,1 (0,3–2,1)	<0,0001

Alle Variablen sind als Mittelwert \pm SD (Min-Max) dargestellt. p-Wert: t-Test für unabhängige Stichproben. Die IR wurde mittels HOMA-Verfahren nach folgender Formel berechnet: (Insulinspiegel nüchtern [μ U/ml] x Glukosespiegel nüchtern [mmol/l])/22,5 [318], wobei IR>2,0 Hinweis auf IR gibt, während IR>2,5 als bestätigte IR gilt. Konversionsfaktor für Insulin von pmol/l in μ U/ml ist 0,138.

Obwohl es sich bei allen Zwillingen der SAME-Studie um gesunde Probanden handelte, zeigten einige der Probanden Hinweise auf eine IR. Es ist bereits bekannt, dass IR mit einer aberranten DNA-Methylierung assoziiert ist [319–321]. Daher wäre es möglich, dass die Methylierungsveränderungen an den identifizierten CpG-Positionen zumindest teilweise durch die IR und nicht durch die Adipositas vermittelt wurden.

Zudem sollte darauf hingewiesen werden, dass es sich bei der DNA-Methylierung um einen gewebespezifischen Prozess handelt [322]. Da es sich bei dem hier verwendeten Untersuchungsmaterial um Zellen des peripheren Blutes handelte, sollte der identifizierte Bezug zur Adipositas und ihren Folgeerkrankungen vorsichtig interpretiert werden. Um die hier identifizierten Ergebnisse zu bekräftigen, wäre eine zusätzliche Methylierungsanalyse beispielsweise am Hypothalamus oder der Leber (wie bei [323]) der Probanden sinnvoll. Allerdings handelt es sich bei Gehirn- und Lebergewebe im Gegensatz zu peripherem Blut um keine leicht zu gewinnenden Untersuchungsmaterialien.

5.3.5 Fazit der DNA-Methylierungsanalyse für hohen BMI

Zusammenfassend konnte mit der hier durchgeführten Analyse ein erster Anhalt erhoben werden, dass sich die adipösen und normalgewichtigen Zwillinge an ausgewählten CpG-Positionen unterscheiden und sich beide Phänotypen somit anhand ihres DNA-Methylierungsmusters differenzieren lassen würden. Neben der bereits beschriebenen Identifizierung des *DPCR1*-Gens aus der Wang-Studie [287], konnten unter den differentiell methylierten Genorten Assoziationen zwischen Adipositas und einem damit einhergehenden Risikofaktor, dem T2D, identifiziert werden. Somit lieferte diese Methylierungsanalyse

Hinweise darauf, dass der epigenetische Mechanismus der DNA-Methylierung zumindest an einigen Loci in der Pathogenese von Adipositas involviert sein könnte.

5.4 Veränderungen des genomweiten DNA-Methylierungsmusters aufgrund des Zigarettenrauchens

Auch in dieser Analyse konnte der klassische Zwillingsansatz nicht angewandt werden. Unter den 23 Zwillingspaaren, von denen die Informationen zum Rauchverhalten zum Zeitpunkt der Analyse vorlagen, fanden sich insgesamt nur acht Zwillingspaare (vier MZ- und vier DZ-Paare), die sich im Rauchstatus diskordant verhielten. Dabei konnten in diesen Zwillingspaaren vier unterschiedliche Kombinationen beobachtet werden (Raucher vs. ehemaliger Raucher (n=3), Gelegenheitsraucher vs. Nichtraucher (n=2), Gelegenheitsraucher vs. ehemaliger Raucher (n=2), Raucher vs. Nichtraucher (n=1)), die aufgrund der jeweiligen geringen Fallzahl eine zwillingsbasierte Analyse nicht zuließen. Um den Einfluss des Zigarettenrauchens auf die genomweite DNA-Methylierung zu analysieren, wurde daher das Methylierungsmuster von insgesamt 12 Rauchern und 25 Nichtrauchern untersucht. In dem stringenten Ansatz wurden zwischen Rauchern und Nichtrauchern 22 differentiell methylierte CpG-Loci, die mit 14 Genen assoziiert werden konnten, identifiziert ($q < 0,05$; t-Test). Mit dem zweiten Ansatz wurden insgesamt 131 differentiell methylierte CpG-Loci ermittelt ($p < 3,13 \times 10^{-4}$; t-Test), die auf allen 22 Autosomen verteilt waren und insgesamt 68 Genen zugeordnet werden konnten.

5.4.1 Eigenschaften der über die FDR neu identifizierten zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte

Insgesamt wurden 50% (11/22) der hier ermittelten differentiell methylierten CpGs bereits in zwei Studien beschrieben, wo sie in peripherem Blut von Rauchern aus Süddeutschland bzw. in mononukleären Zellen des peripheren Blutes von rauchenden Afroamerikanerinnen jeweils unter Verwendung des HM450K BC als differentiell methyliert detektiert wurden [133,134]. Diese beinhalteten drei Loci der 533bp großen Region auf Chromosom 2, alle vier CpG-Loci innerhalb des *AHRR*-Gens, jeweils einen CpG-Locus in den Genen *ANPEP*, *F2RL3* und *HAP1*, sowie einen Locus auf Chromosom 6 (6p21.33), der entsprechend der Illumina Annotation mit keinem Gen assoziiert werden konnte.

Diskussion

Von den erstmals neu identifizierten elf CpG-Stellen konnte ein Locus (cg12207248) keinem Gen zugeordnet werden. Die übrigen zehn CpGs konnten mit zehn Genen assoziiert werden: *BICD1*, *C20orf46 (TMEM74B)*, *ELK3*, *F10*, *RRAS2*, *SLC38A4*, *STIM2*, *TAPT1*, *ZNF254* und *ZNF257*.

Der differentiell methylierte Locus cg01321189 liegt in der TSS1500 des *BICD1*-Gens in einer 916bp großen CpG-Insel einer Promotor-assoziierten Region. *BICD1 (bicaudal D homolog 1)* hat eine Funktion im retrograden Transport von Proteinen vom Golgi-Apparat zum endoplasmatischen Retikulum (ER) inne [324]. Weiterhin konnte *BICD1* beim Menschen als Regulator der Telomerlänge identifiziert werden [325], die generell als biologischer Marker für zelluläres Altern angesehen wird [326]. Verantwortlich hierfür war ein SNP, der mit einer verminderten *BICD1*-Expression korrelierte und mit einer verkürzten Telomerlänge in Leukozyten assoziiert werden konnte. Eine Verkürzung der Telomere ist weiterhin mit einer Vielzahl von Krankheiten assoziiert, einschließlich Krebs [327], kardiovaskulärer Erkrankungen [328,329], COPD [330], T2D [331] und Adipositas [332,333]. Ebenfalls konnte eine verkürzte Telomerlänge bereits in Leukozyten von Raucherinnen [334] sowie in alveolaren Makrophagen von Rauchern beider Geschlechtes [335] beobachtet werden. Ob die verminderte Methylierung in der TSS1500 von *BICD1* in Rauchern mit einer Telomerverkürzung in Zusammenhang stehen könnte, kann hier jedoch nicht zweifelsfrei belegt werden, zumal eine Demethylierung in dieser Region gewöhnlich mit einer erhöhten *BICD1*-Expression korrelieren müsste. Des Weiteren konnte ebenfalls ein SNP im *BICD1*-Gen mit der Pathogenese eines Lungenemphysems assoziiert werden [336], dessen Ausbildung stark mit dem Rauchen assoziiert ist [337,338]. Insgesamt könnte somit eine Assoziation zwischen der hier beobachteten verminderten Methylierung in der genannten *BICD1*-Region in Rauchern und den negativen gesundheitlichen Auswirkungen des Rauchens in Betracht gezogen werden.

RRAS2 gehört zu den unter 4.2.5.2.4 identifizierten Krebsgenen und zählt zu den Onkogenen. In Zellen des peripheren Blutes von Rauchern konnte hier in der TSS1500 eine verminderte Methylierung (-17,7%) beobachtet werden, die in einer verstärkten Expression resultieren sollte. Eine erhöhte Expression von *RRAS2* konnte bereits in vielen Tumorentitäten beobachtet werden, einschließlich Mundhöhlenkrebs [339], für dessen Ausbildung das

Rauchen einen großen Risikofaktor darstellt [340]. Somit könnte die veränderte Methylierung in Rauchern zur Karzinogenese beitragen.

STIM2 (*stromal interaction molecule*) ist für die Regulation der Kalziumionen-Konzentration im Cytosol und ER mitverantwortlich und spielt u.a. im hypoxischen Zelltod von Neuronen eine Rolle [341]. Eine Überexpression von *STIM2* konnte in Darmkrebs-Gewebe bzw. -Zelllinien beobachtet werden [342]. Diese korrelierte mit einem weniger invasiven Phänotyp bzw. einer verminderten Zellproliferation und Tumorwachstum, weshalb dem *STIM2*-Gen die Eigenschaft eines Tumorzellen-Wachstumssuppressors verliehen wurde. Die verminderte Methylierung in Genkörper von *STIM2* in Rauchern sollte mit einer verminderten Expression korrelieren. Aufgrund der mit einer erhöhten Expression assoziierten Zellwachstums-Suppression, müsste demnach eine verminderte Expression mit einer erhöhten Zellproliferation bzw. Tumorwachstum einhergehen. Somit könnte sich die Methylierungsänderung in *STIM2* bei Rauchern ungünstig auf das (Tumor-)Zellwachstum auswirken.

Eine mögliche Assoziation in der bei Rauchern verminderten Methylierung der Genorte *F10* und *SLC38A4* wird in 5.4.3.3 bzw. 5.4.2 diskutiert.

Die übrigen Gene *C20orf46*, *ELK3*, *TAPT1*, *ZNF254* und *ZNF257* wurden in der Literatur bisher teilweise wenig beschrieben. Des Weiteren konnten für diese Gene keine plausiblen Assoziationen zum Rauchen hergestellt werden.

5.4.2 Eigenschaften der aus dem zweiten Ansatz neu identifizierten zwischen Rauchern und Nichtrauchern differentiell methylierten Genorte

22/131 (16,8%) der aus dem zweiten Ansatz ermittelten CpG-Loci, die zwischen Rauchern und Nichtrauchern eine differentielle Methylierung aufwiesen, konnten in den Studien von Zeilinger und Dogan bestätigt werden [133,134]. Darunter waren sieben CpGs im Gen *AHRR*, fünf CpGs auf Chromosom 2 sowie zwei CpGs auf Chromosom 6, die keinem Gen zugeordnet werden konnten, jeweils ein Locus in den Genen *ANPEP*, *F2RL3*, *PIP5K1A* und *PRSS23* sowie insgesamt vier CpGs auf den Chromosomen 9, 14 und 19, die mit keinem annotierten Gen assoziiert werden konnten. Des Weiteren befanden sich unter den verbliebenen 109 differentiell methylierten CpGs, die mit 63 Genen assoziiert werden konnten, drei Gene

Diskussion

(*ANKRD11*, *CACNA1C* und *TBCD*), die in den oben genannten Studien ebenfalls zwischen Rauchern und Nichtrauchern als differentiell methyliert ermittelt wurden, aber mit anderen CpG-Positionen in Verbindung gebracht wurden. Weiterhin konnte Christensen *et al.* unter Verwendung des Illumina GoldenGate Arrays bereits einen CpG-Locus (P91) im *TNFRSF10A*-Gen mit einer Hypomethylierung im Lungengewebe von Rauchern im Vergleich zu Nichtrauchern identifizieren [146]. Auch hier konnte bei Rauchern eine Hypomethylierung (7,1%, $p < 1,55 \times 10^{-4}$; t-Test für unabhängige Stichproben) beobachtet werden. Da die übrigen 102 mit dem Rauchen assoziierten Loci mit 59 Genen in Verbindung gebracht wurden und nicht jedes Gen separat diskutiert werden kann, werden im Folgenden nur die Gene diskutiert, die von mehreren benachbarten differentiell methylierten CpG-Positionen betroffen waren: *SLC38A4* mit jeweils acht, *WFIKK2* mit jeweils drei und die Gene *BEGAIN*, *ENPEP*, *HDAC4* und *PPT2* mit jeweils zwei differentiell methylierten CpG-Loci.

Am auffälligsten waren acht differentiell methylierte CpGs, die mit dem Gen *SLC38A4* auf Chromosom 12 assoziiert waren (Abbildung 42) und bei Rauchern im Vergleich zu Nichtrauchern eine mittlere Hypomethylierung von 12,5% aufwiesen. Die Methylierungsdifferenzen der acht Loci reichten dabei von 7,8% (cg18064171; 63,1% vs. 70,9%; $p < 1,76 \times 10^{-4}$; t-Test für unabhängige Stichproben) bis 18,9% (cg11146691; 30,5% vs. 49,4%; $p < 1,94 \times 10^{-4}$; t-Test für unabhängige Stichproben). Der abgedeckte Bereich erstreckte sich über 331bp (12q13.11), wovon fünf CpGs in der TSS200, zwei CpGs im ersten Exon und ein CpG in der 5'-UTR innerhalb des ersten Introns lokalisiert waren. Interessanterweise konnte bereits einer der acht CpG-Loci (cg14671384) im *SLC38A4*-Gen in der vorangegangenen Analyse über die FDR ermittelt werden (vgl. Tabelle 10).

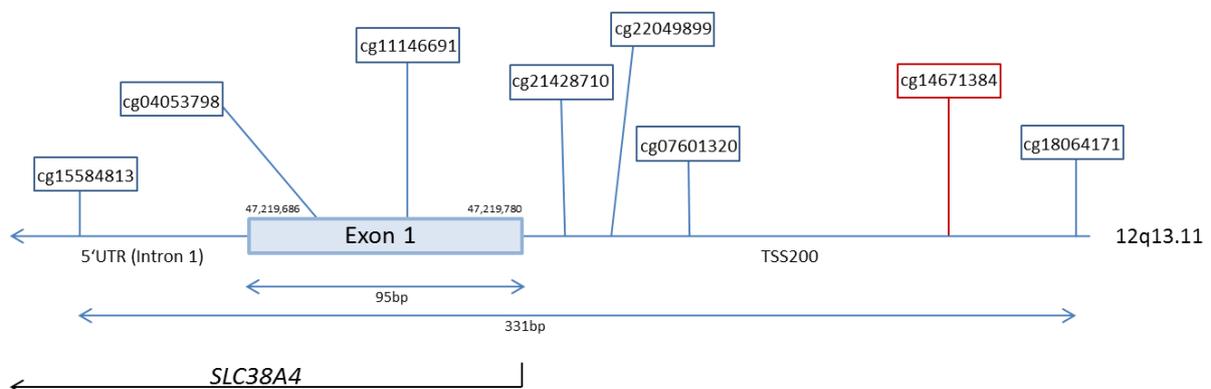


Abbildung 42: Ausschnitt der mit dem Gen *SLC38A4* assoziierten 12q13.11-Region, in der acht differentiell methylierte CpG-Loci jeweils eine Hypomethylierung bei den Rauchern im Vergleich zu den Nichtrauchern aufwiesen. Der rot markierte CpG-Locus konnte bereits über die FDR als signifikant differentiell methyliert ermittelt werden.

Diskussion

Das Gen *SLC38A4* (*solute carrier family 38 member 4*) kodiert einen neutralen Aminosäure-Transporter [343]. Die Familie der SLC-Transporter spielt in der Aufnahme von hydrophilen Arzneimitteln eine Rolle und konnte bereits mit Arzneimittelresistenz in Verbindung gebracht werden. Eine Überexpression von *SLC38A4* konnte dabei in einer gegen ein bestimmtes Medikament resistenten Eierstockkrebs-Zelllinie beobachtet werden [344]. Es ist bekannt, dass die Aufnahme von Nikotin die Chemotherapie-induzierte Apoptose von Krebszellen hemmt und somit die Resistenz von Tumoren fördert, so wie beim Lungenkrebs [345]. Demnach könnte die mit der verminderten Methylierung korrelierende verstärkte Expression von *SLC38A4* in Zellen des peripheren Blutes bei Rauchern ein Mechanismus sein, der zur schlechteren Prognose z.B. von Lungenkrebs beitragen könnte, die man bei Patienten beobachten kann, die während ihrer Chemotherapie weiterhin rauchten [346,347].

Im Gen *WFIKKN2* konnten zwischen Rauchern und Nichtrauchern drei differentiell methylierte CpG-Loci identifiziert werden, die in den TSS200 und TSS1500 bei den Rauchern im Durchschnitt um 6,6% hypomethyliert vorlagen. Das kodierte Protein enthält mehrere Protease-inhibierende Domänen, mit denen es verschiedene Typen von Proteasen steuern könnte [348]. Es konnte bereits gezeigt werden, dass im Blut von Rauchern eine erhöhte Anzahl inflammatorischer Zellen existiert, wie den Neutrophilen und Monozyten. Diese produzieren wiederum große Mengen an Proteasen [349,350], denen normalerweise Protease-Inhibitoren entgegenwirken [351]. Durch Tabakrauch können diese Antiproteasen jedoch inaktiviert [352]. Dies resultiert in einem Ungleichgewicht zwischen Proteasen und Antiproteasen, das beispielsweise in COPD und Lungenemphysemen beobachtet werden kann [353,354]. Die in den Rauchern verminderte Methylierung in den TSS von *WFIKKN2* könnte mit einer erhöhten Expression des protease-inhibierenden Proteins korrelieren. Dies könnte als Reaktionsmechanismus auf die vermehrten proteolytischen Enzyme angesehen werden.

Das Gen *BEGAIN* (*brain-enriched guanylate kinase-associated*) ist auf Chromosom 14 lokalisiert und wurde in der durchgeführten Analyse mit zwei in Rauchern differentiell methylierten benachbarten CpGs in Verbindung gebracht. Dabei wiesen beide CpG-Loci in Rauchern eine verminderte Methylierung in der TSS200 auf, wobei eine Differenz von 12,1% auf den Locus cg13521842 und ein Unterschied von 8,9% auf den Locus cg21193577 fiel. *BEGAIN* ist in der Literatur bisher wenig beschrieben. Im Gehirn von Ratten ist es in

Diskussion

Neuronen sowohl im Kern als auch an den Synapsen lokalisiert, was dem Genprodukt *BEGAIN* eine Rolle in der direkten Übermittlung von Signalen zuschreiben könnte [355]. Kürzlich konnte eine verminderte Methylierung von *BEGAIN* mit der Psychopathologie der Depression in Verbindung gebracht werden [356]. Neben einem erhöhten Krebsrisiko kann das Rauchen ebenfalls mit Depressionen in Zusammenhang gebracht werden. Dabei ist aber nicht eindeutig belegt, ob das Rauchen *per se* zur Entwicklung einer Depression führt oder ob die Depression ein Risikofaktor für die Tabakabhängigkeit darstellt [357]. So konnte zum einen in Rauchern im Vergleich zu Nichtrauchern über einen Zeitraum von mehreren Jahren eine signifikante Zunahme des Risikos zur Ausbildung einer Depression beobachtet werden [358], zum anderen bewirkte ein Rauchstopp einen deutlich weniger depressiven Phänotyp [359]. Somit könnte die in der vorliegenden Analyse in Rauchern beobachtete verminderte Methylierung in der TSS200 von *BEGAIN* für die mit Rauchen assoziierte Depression mitverantwortlich sein.

Im Genkörper von *HDAC4*, das für die Histon-Deazetylase 4 kodiert, konnte in zwei benachbarten CpG-Loci eine mittlere Zunahme der Methylierung in Rauchern um 12,2% beobachtet werden. Dabei war der Locus cg03568507 in einer 316bp großen CpG-Insel und der Locus cg12547218 im S-Shore der genannten Insel lokalisiert. *HDAC4* ist in der Histonazetylierung und Chromatinstrukturierung involviert. Eine erhöhte *HDAC4*-Expression könnte für die Deazetylierung von Histonen verantwortlich sein, was wiederum die Expression spezifischer Ziel-Proteine unterdrücken könnte. Im Zwerchfell von COPD-Patienten konnten erhöhte *HDAC4*-Mengen festgestellt werden, was als biologischer Anpassungsmechanismus zur besseren Bewältigung der kontinuierlichen Belastungen durch Inspiration auf die Atemwege gedeutet wurde [360]. Des Weiteren konnte in verschiedenen Krebszellen bereits eine Korrelation zwischen *HDAC4*-Aktivierung und inflammatorischer Prozesse identifiziert werden [361]. Da das Rauchen mit einem erhöhten Vorkommen inflammatorischer Marker assoziiert ist [362], könnte eine erhöhte Methylierung im Genkörper von *HDAC4*, die mit einer erhöhten Expression korrelieren müsste, entweder zu diesem beitragen oder als Konsequenz daraus hervorgehen.

In Zellen des peripheren Blutes von Rauchern im Vergleich zu Nichtrauchern konnten zwei hypomethylierte CpG-Positionen in TSS1500 und TSS200 des Gens *PPT2* (*palmitoyl-protein thioesterase 2*) identifiziert werden, die beide mit der Promotorregion assoziiert wurden und

im Mittel eine verminderte Methylierung von 8,4% aufwiesen. *PPT2* kodiert für ein lysosomales Enzym, das am Fettsäure-Metabolismus beteiligt ist [363]. In verschiedenen genomweiten Assoziationsstudien konnten genetische Varianten von *PPT2* bereits mit der Pathogenese chronischer Lungenerkrankungen wie COPD bzw. einem Verlust der Lungenfunktion in Verbindung gebracht werden [364–366]. Somit könnte auch eine veränderte Methylierung im Promotor von *PPT2*, was hier mit einer erhöhten *PPT2*-Expression in Rauchern korrelieren könnte, zur Pathogenese der mit Rauchen assoziierten chronischen Lungenerkrankungen beitragen.

Eine mögliche Assoziation zwischen einer veränderten Methylierung von *ENPEP* und dem Rauchen wird in 5.4.3.2 diskutiert.

Somit konnte den neu identifizierten mit dem Rauchen in Zusammenhang stehenden Genorten, in denen mehrere differentiell methylierte CpG-Positionen lokalisiert waren, eine mögliche Verbindung zu den durch das Rauchen verursachten Begleiterscheinungen hergestellt werden.

5.4.3 Differentiell methylierte mit dem Rauchen assoziierte Gene sind im kardiovaskulären System involviert

Über ermittelte signifikante Anreicherungen verschiedener Assoziations-Analysen konnten die durch das Rauchen aberrant methylierten Gene primär mit dem kardiovaskulären System in Zusammenhang gebracht werden. Die Resultate der dabei auf DNA-Methylierungsebene basierten Anreicherungen decken sich mit der Tatsache, dass Rauchen ein bedeutender Risikofaktor für kardiovaskuläre Erkrankungen mit Auswirkungen auf die Blutgerinnung ist [367].

5.4.3.1 Anreicherungen der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene in Herzerkrankungen

Aktives Zigarettenrauchen ist nicht nur mit einem erhöhten Risiko für verschiedene Krebsentitäten verbunden, sondern ist auch ein bedeutender Risikofaktor für Erkrankungen der Lunge sowie des kardiovaskulären Systems [219,368]. Über die hier durchgeführten Assoziations-Analysen konnten Erkrankungen des Herz-Kreislauf-Systems identifiziert werden, die somit eine zumindest partielle Beteiligung der DNA-Methylierung an den

Auswirkungen des Rauchens auf den Organismus nahelegen könnten. Die signifikant angereicherten differentiell methylierten Gene fanden sich dabei hauptsächlich in Herzerkrankungen sowie teils angeborener Herzfehlbildungen, insbesondere des Ventrikels. Veränderungen der diastolischen Ventrikelfunktion konnten sowohl in Rauchern direkt nach dem Konsum einer Zigarette [369–373] als auch in Passivrauchern, die unmittelbar dem Tabakrauch ausgesetzt waren [374,375], beobachtet werden.

5.4.3.2 Anreicherung der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene im Renin-Angiotensin-Signalweg

Die Gene *ANPEP* und *ENPEP* lagen im Signalweg des Renin-Angiotensin-Systems (RAS) signifikant angereichert vor. Das RAS wirkt über eine Kaskade regulatorischer Enzyme als wichtiger Regulator des Blutdrucks sowie der Aufrechterhaltung des Blutvolumens und Elektrolythaushalts [376]. Aufgrund seines bedeutenden Beitrags an der Regulierung des Blutdrucks könnten Störungen an einer der beteiligten Komponenten des Systems zu Bluthochdruck und Herz-Kreislauf-Erkrankungen führen [377]. Bei den hier identifizierten differentiell methylierten Genen handelte es sich um zwei Aminopeptidasen, die entscheidend am Metabolismus der Angiotensine mitwirken. *ENPEP* kodiert für die Aminopeptidase A (APA) und ist für die Umwandlung von Angiotensin II zu Angiotensin III verantwortlich, während *ANPEP* für die Aminopeptidase N (APN) kodiert und die Umwandlung von Angiotensin III zu Angiotensin IV katalysiert [378–380]. Ein detaillierter Überblick über die Komponenten des RAS ist in Abbildung 43 dargestellt. Durch spezifische Bindung an den entsprechenden Angiotensin-Rezeptoren (AT_1 , AT_2 und AT_4) können die genannten Proteine wichtige biologische Funktionen steuern.

Diskussion

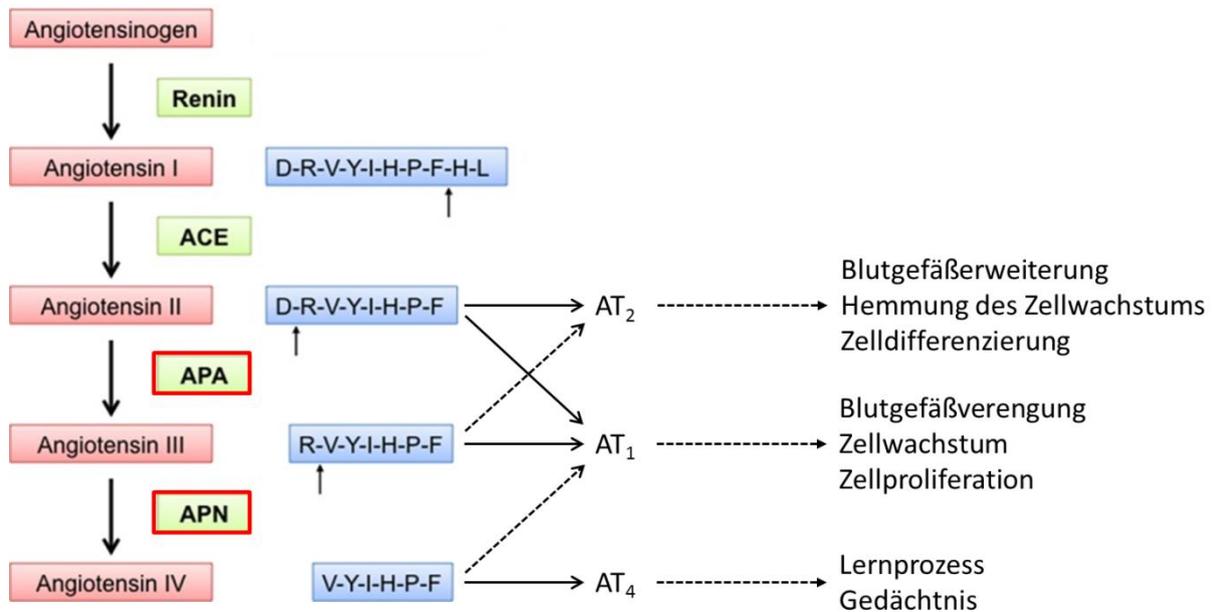


Abbildung 43: Das Renin-Angiotensin-System (RAS). Angiotensinogen ist die erste Komponente der enzymatischen Kaskade und das primäre Peptid-Substrat des RAS. Es wird von der Protease Renin zu Angiotensin I konvertiert, welches wiederum durch das Angiotensin-konvertierende Enzym ACE zu Angiotensin II hydrolysiert wird. Angiotensin II wird anschließend durch APA zu Angiotensin III konvertiert, welches im letzten Schritt der Kaskade schließlich durch APN zu Angiotensin IV hydrolysiert wird [380]. Die Wirkung der Angiotensine II-IV auf spezifische Rezeptoren und einigen daraus resultierenden biologischen Effekten ist dargestellt. Die Angiotensin-Rezeptoren AT₁ und AT₂ haben gegenläufige Effekte. Angiotensin III und IV konnten ebenfalls als schwache Agonisten der AT₂- bzw. AT₁-Rezeptoren identifiziert werden [381,382]. Abbildung verändert nach [383].

Angiotensin II ist das Haupt-Effektor-Molekül des RAS [377], welches eine wichtige Rolle in der Regulation der kardiovaskulären Homöostase spielt und dessen Umwandlung das hier identifizierte Genprodukt APA katalysiert. Eine erhöhte Angiotensin II-Aktivität konnte bereits eng mit der Pathogenese von Bluthochdruck assoziiert werden, die durch den AT₁-Rezeptor vermittelt wird [384]. Dabei führt die dadurch resultierende Verengung der Blutgefäße zur Erhöhung des Blutdrucks, welche bei Rauchern beobachtet werden kann [385]. Somit wäre eine veränderte Methylierung von Genen, die den Blutdruck über das RAS regulieren, im Blut von Rauchern vorstellbar. Des Weiteren konnten Li *et al.* zeigen, dass Nikotin, eine Hauptkomponente des Zigarettenrauchs, direkt den AT₁-Rezeptor vermittelten Signalweg und die daraus resultierende Zellproliferation der glatten Gefäßmuskelzellen verstärkt [384]. Ebenfalls konnte eine durch chronisches Rauchen hervorgerufene erhöhte Aktivität des RAS demonstriert werden [386,387], die Laustiola *et al.* [386] letztendlich einer potentiellen erhöhten Bildung von Angiotensin II zuschrieben.

Zwei zwischen Rauchern und Nichtrauchern differentiell methylierte CpGs (cg15810415 und cg20773127) grenzen einen auf Chromosom 4 liegenden 22bp großen Bereich ein, der in der

TSS200 des *ENPEP*-Gens sowie im N-Shore einer 412bp großen CpG-Insel lokalisiert ist. Beide CpG-Loci zeigen bei Rauchern im Vergleich zu Nichtrauchern eine Hypomethylierung, wobei cg15810415 eine durchschnittliche Methylierungsdifferenz von 7,4% (mittlere Methylierung Raucher: 72,8%, mittlere Methylierung Nichtraucher: 80,2%; $p < 4,58 \times 10^{-5}$) und cg20773127 einen mittleren Methylierungsunterschied von 6,7% (mittlere Methylierung Raucher: 62,5%, mittlere Methylierung Nichtraucher: 69,2%; $p < 2,81 \times 10^{-4}$) aufwies. Eine durch die verminderte Methylierung in der TSS des Gens *ENPEP* bedingte erhöhte Expression von APA könnte in Rauchern zur gesteigerten Bildung von Angiotensin III führen, welches über den AT_1 -Rezeptor die Blutgefäßverengung und somit die Blutdruckerhöhung vermitteln kann.

Der mit dem Gen *ANPEP* assoziierte CpG-Locus cg23161492 liegt in der 5'UTR des Gens sowie im N-Shore einer etwa 1,2kb großen CpG-Insel. Der Locus zeigt bei Rauchern eine mittlere Methylierung von 21,4%, während diese bei Nichtrauchern bei 28,2% liegt (mittlere Methylierungsdifferenz: 6,8%; $q < 0,05$). Der hier zwischen Rauchern und Nichtrauchern identifizierte differentiell methylierte CpG-Locus konnte zuvor bereits in einer Studie von Zeilinger *et al.* [133] gezeigt werden, wo er in Rauchern im Vergleich zu den Nichtrauchern zwar auch eine Hypomethylierung zeigte, das Methylierungsniveau dabei aber insgesamt niedriger lag. Eine geringere Methylierung in der 5'UTR von *ANPEP* könnte bei Rauchern in einer erhöhten APN-Expression resultieren, die Folge der gesteigerten durch überexprimiertes APA verursachten Angiotensin III-Bildung sein könnte.

Insgesamt könnte die bei Rauchern durch Hypomethylierung vermittelte erhöhte Expression von APA und APN demnach wesentlich zum in Rauchern beobachteten Bluthochdruck beitragen.

5.4.3.3 Anreicherung der zwischen Rauchern und Nichtrauchern differentiell methylierten Gene im Krankheitsbild der Thrombose

Über die durch Rauchen differentiell methylierten Gene *F2RL3* und *F10 (FX)* konnte eine signifikante Anreicherung für die Assoziation mit Thrombosen identifiziert werden. Der CpG-Locus cg03636183 war im Genkörper des *F2RL3*-Gens im N-Shore einer 771bp großen CpG-Insel lokalisiert und lag bei Rauchern (mittlere Methylierung: 62,3%) im Vergleich zu Nichtrauchern (mittlere Methylierung: 76,6%) hypomethyliert vor ($q < 9,06 \times 10^{-9}$). Das CpG cg08143875 wurde im Genkörper von *FX* in einer 816bp großen CpG-Insel als aberrant

methyliert identifiziert, wobei es bei Rauchern mit einer mittleren Methylierung von 38,5% im Gegensatz zu 43,7% bei Nichtrauchern eine Hypomethylierung zeigte ($q < 0,05$).

F2RL3 gehört zur Familie der Protease-aktivierten Rezeptoren, die im kardiovaskulären System als Regulatoren der Hämostase und Thrombose eine wichtige Rolle spielen. F2RL3 hat beispielsweise eine Funktion in der Thrombozyten-Aggregation inne, wobei es u.a. durch Thrombin sowie die Koagulations-Protease FXa, welche die aktivierte Form von FX darstellt, aktiviert werden kann [388–390]. FX ist als Koagulationsfaktor X ein wichtiger Bestandteil der Gerinnungskaskade. Der aktivierte Faktor wandelt Prothrombin zum aktiven Thrombin um, welches wiederum die Polymerisation von Fibrin aus Fibrinogen katalysiert, wodurch letztendlich die Thrombus-Bildung initiiert wird [391,392].

Somit tragen die beiden hier identifizierten Gene zur Pathogenese der Thrombose bei, die durch den Effekt des Rauchens in Form einer aberranten Methylierung der besagten Gene gestört werden könnte. Diese Hypothese wird unterstützt durch die Tatsache, dass Rauchen mit einem erhöhten Thromboserisiko assoziiert ist [393–395]. Das Rauchen von Zigaretten scheint über verschiedene Mechanismen Auswirkungen auf die Hämostase zu haben, indem es die Funktionen von Endothelzellen, Thrombozyten, Fibrinogen und Gerinnungsfaktoren beeinträchtigt. Dies führt zu einem Ungleichgewicht von sowohl antithrombotischen und prothrombotischen als auch profibrinolytischen und antifibrinolytischen Faktoren, die folglich die Einleitung und Ausbreitung der Thrombose fördern [367,396,397].

5.4.4 Prüfung der über die FDR bei Rauchern identifizierten 22 CpG-Methylierungssignatur in Lungen-Proben

Das Rauchen von Zigaretten ist eine der Hauptursachen für Lungenkrebs [398]. In der hier durchgeführten Analyse ließ sich die Tendenz beobachten, dass sich unauffällige Lungenproben von Lungenkrebsproben über das Methylierungsmuster der hier im Rahmen dieser Arbeit in Zellen des peripheren Blut identifizierten mit dem Rauchen assoziierten 22 CpG-Signatur ($q < 0,05$) differenzieren lassen würden. Hierbei darf nicht außer Acht gelassen werden, dass sich die Tendenz einer Trennung abzeichnete, obwohl es sich um das DNA-Methylierungsmuster völlig unterschiedlicher Gewebe handelte und demnach eine gewebespezifische Methylierung zu erwarten war. Diese Tatsache fiel besonders bei zwei CpG-Loci ins Gewicht, an denen sich die Lungenproben von den Lungenkrebsproben

signifikant ($q < 0,05$) unterscheiden ließen. Die betroffenen Loci waren jeweils im Genkörper von *AHRR* und *F10* lokalisiert und zeigten in Lungenkrebsproben eine deutliche Hypermethylierung gegenüber den gesunden Lungenproben, während sich bei Rauchern im Vergleich zu Nichtrauchern eine klare Hypomethylierung manifestierte. Hierbei wurde ebenfalls deutlich, dass sich in beiden Genen die Methylierung im „Normal- bzw. Urzustand“ der beiden Gewebetypen (unauffälliges Lungengewebe vs. Zellen des peripheren Blutes von Nichtrauchern) mit mittleren Methylierungsdifferenzen von $14,4\% \pm 0,7\%$ in *F10* und $64,4\% \pm 0,8\%$ in *AHRR* jeweils signifikant unterschied ($p < 0,0001$; t-Test für unabhängige Stichproben). Im Folgenden werden die beiden betroffenen Gene separat diskutiert.

5.4.4.1 Differentielle Methylierung in *AHRR* (*aryl hydrocarbon receptor (AHR) repressor*)

Shenker *et al.* konnten für zwei CpG-Loci im *AHRR*-Gen ebenfalls eine signifikante Hypomethylierung in Zellen des Blutes von Rauchern verglichen mit denen der Nichtraucher identifizieren und ferner im Lungengewebe der Raucher eine identische Methylierung beobachten. Weiterhin konnten sie zeigen, dass das Methylierungsniveau dieser Loci im *AHRR*-Gen in den Lungenproben invers mit der Genexpression korrelierte [132]. Interessanterweise konnte beim Vergleich der hier erhobenen Daten für die Lungenkrebsproben sowohl im untersuchten Locus des *AHRR*- als auch *F10*-Gens eine ähnliche Methylierung ermittelt werden wie für die Blutproben der Raucher (siehe Tabelle 18 und Tabelle 19). In Anlehnung an die Ergebnisse der Studie von Shenker *et al.* könnte daraus geschlossen werden, dass zumindest der Locus im *AHRR*-Gen sowohl im Lungenkrebsgewebe als auch im Lungengewebe und Blut von Rauchern in Folge der Rauchexposition in vergleichbarer Weise aberrant methyliert zu sein scheint, was sich auch auf den hier untersuchten Locus im *F10*-Gen (siehe 5.4.4.2) übertragen lassen könnte.

Diskussion

Tabelle 18: Erhobene Methylierungswerte im *AHRR*-Locus von Rauchern, Nichtraucher, Lungenkrebs- und unauffälligem Lungengewebe.

<i>AHRR</i> (cg05575921)	Lungen-Ca	Raucher	<i>AHRR</i> (cg05575921)	Lunge	Nichtraucher
Anzahl der Proben	34	12	Anzahl der Proben	33	25
Minimum	11,3%	47,7%	Minimum	12,0%	72,0%
25. Perzentil	40,8%	51,2%	25. Perzenti	14,2%	79,1%
Median	54,2%	54,0%	Median	16,3%	80,2%
75. Perzentil	75,0%	55,9%	75. Perzenti	18,2%	82,5%
Maximum	79,6%	75,6%	Maximum	21,6%	89,9%
Mittelwert	53,7%	54,9%	Mittelwert	16,3%	80,7%
SD	20,8%	7,0%	SD	2,6%	3,8%
SEM	3,6%	2,0%	SEM	0,4%	0,8%
p-Wert	<0,8372		p-Wert	<0,0001	

SD: Standardabweichung. SEM: Standardfehler. p-Wert ermittelt über t-Test für unabhängige Stichproben.

Zudaire *et al.* konnten *AHRR* bereits als mögliches Tumorsuppressorgen in mehreren Tumorentitäten, darunter auch Lungenkrebs, identifizieren und eine Hypermethylierung an entsprechenden Loci im Gen mit einer Herunterregulation der *AHRR*-Expression in den Krebszellen assoziieren [399]. Ferner konnten Monick *et al.* für den in dieser Arbeit identifizierten durch Rauchen aberrant methylierten CpG-Locus cg05575921 eine signifikante Assoziation zwischen einer Zunahme der Methylierung und einer Abnahme der *AHRR*-Genexpression in Lymphoblasten zeigen, was auf eine funktionelle Rolle in der Regulation der Genexpression an diesem Locus hindeutet [130]. Demnach könnte hier geschlussfolgert werden, dass sowohl cg05575921 (*AHRR*) als auch cg08143875 (*F10*) regulatorische Funktionen innehaben, die im Blut von Rauchern mit einer Zunahme in der Expression der betroffenen Gene assoziiert sein könnten.

Das Gen *AHRR* ist Teil der Arylhydrokarbon-Rezeptor (AHR) Signalkaskade, welche Komponenten des Zigarettenrauchs, einschließlich der karzinogenen Dioxine und Dioxin-ähnlichen Stoffe sowie polyzyklische aromatische Kohlenwasserstoffe (PAHs), metabolisiert. Der AHR, auch Dioxin-Rezeptor genannt, ist dabei ein ligandenaktivierter Transkriptionsfaktor, der durch Noxen aktiviert wird und über eine Signalkaskade deren biologischen und toxischen Effekte vermittelt. Gerade im Atmungssystem kommt es somit durch eine von AHR induzierte Veränderung in der Expression von Zielgenen zu einigen pathologischen Effekten, beispielsweise chronischer Bronchitis und Lungenkrebs [132,400].

Diskussion

Studien mit transgenen Mäusen unterstreichen zudem das onkogene Potential des aktiven AHR, da eine konstitutive Überexpression des *AHRs* zu einem erhöhten Auftreten spontaner und induzierter Tumoren führt [401,402].

Der AHR reguliert die Expression des *AHRs*, welcher wiederum als transkriptioneller Repressor von AHR fungiert und somit dessen Funktion inhibiert. AHR und *AHR* bilden somit eine negative Rückkopplungsschleife [399,403,404]. *AHR* vermittelt die Entgiftung von PAHs, die als krebserregende Substanzen maßgeblich auf durch Rauchen verursachte Krebserkrankungen zurückzuführen sind [399,405].

Polymorphismen in *AHR* und *AHR* konnten bereits mit einigen Erkrankungen, darunter Lungenkrebs und im Zusammenhang mit Dioxin stehender männlichen Infertilität, in Verbindung gebracht werden [406–410].

Da *AHR* die transkriptionelle Aktivität und somit auch die Signaltransduktion des *AHRs* kontrolliert, ist mit einer Regulation der Expression von *AHR*-Zielgenen durchaus zu rechnen. Die Methylierung des hier untersuchten CpG-Locus lag für Lungenkrebs und Raucher zwar auf einem ähnlichen mittleren Niveau, jedoch verhielt sich dieses im Vergleich zu den jeweiligen korrespondierenden gesunden Kontrollen invers. Eine konträre aberrante Methylierung des *AHR*-Locus könnte in den beiden unterschiedlichen Kompartimenten somit entweder aufgrund der gewebespezifischen Methylierung bedingt oder Folge verschiedener *AHR* Funktionen in Lunge und Blut sein. Im Blut von Rauchern wäre dabei beispielsweise eine vor Krebs schützende Funktion des *AHRs* durch Inhibierung spezifischer Zielgene vorstellbar, die für toxische Enzyme kodieren könnten, was auf eine verstärkte Expression des *AHRs* infolge der hier identifizierten Hypomethylierung des Locus hindeutet. Folglich spricht die ermittelte Hypermethylierung des *AHR*-Locus im Blut der Nichtraucher für eine verminderte Expression oder Repression von *AHR*, da diese aufgrund einer geringen oder fehlenden Exposition gegenüber Noxen auf die AHR-Signalkaskade nicht angewiesen sind. Es wäre im Raucherblut aber auch eine durch Hypomethylierung bedingte Repression von *AHR* denkbar, da es somit zu keiner *AHR*-abhängigen Repression des AHR-vermittelten Signalwegs kommt, der durch die toxischen Substanzen im Zigarettenrauch aktiviert wird. Eine erhöhte Expression von *AHR* im Blut der Nichtraucher müsste demnach mit einer erhöhten *AHR*-Aktivität assoziiert sein, die als Art Schutzmechanismus wiederum den AHR-Signalweg inhibieren könnte, um so vor einer möglichen Expression „falscher“

Zielgene zu schützen. Die eben beschriebenen möglichen Szenarien wären auch auf die Lungenproben übertragbar, wobei das Methylierungsspektrum in den Lungenkrebsproben relativ variabel ist. Dies könnte entweder mit verschiedenen Formen des Lungenkrebs zusammenhängen oder aber ob der jeweilige Lungenkrebs durch das Rauchen von Zigaretten induziert wurde.

5.4.4.2 Differentielle Methylierung in *F10* (coagulation factor X)

Die Funktion von *F10* wurde in 5.4.3.3 bereits diskutiert, wo es im Blut von Rauchern signifikant weniger methyliert vorlag als im Blut von Nichtrauchern und mit einem erhöhten Thrombose-Risiko in Verbindung gebracht wurde. In den Lungenkrebsproben war der Grad der Methylierung des *F10*-Locus mit dem der Raucher vergleichbar (siehe Tabelle 19), auch wenn die Varianz der DNA-Methylierung im Gewebe des Lungenkarzinoms stärker variierte als in den peripheren Blutzellen der Raucher. Eine Vielzahl an Krebsentitäten, darunter auch Lungenkrebs, konnte bereits mit Hyperkoagulabilitäts-Störungen assoziiert werden, die wiederum mit einem erhöhten Thromboserisiko assoziiert sind [411,412]. Daher ist ein vergleichbares Methylierungsniveau des CpG-Locus cg08143875 im Genkörper von *F10* im Lungenkrebsgewebe und im Blut von Rauchern durchaus plausibel. Die Unterschiede in der DNA-Methylierung im Normalzustand (Lungengewebe vs. Blut von Nichtrauchern) verhält sich hingegen invers. Im unauffälligen Lungengewebe zeichnete sich eine Hypomethylierung im Vergleich zum Lungenkrebsgewebe ab, während sich in den Blutzellen der Nichtraucher verglichen mit denen der Raucher eine Hypermethylierung manifestierte. Dies lässt sich nur mit einer gewebespezifischen Methylierung vereinbaren.

Diskussion

Tabelle 19: Erhobene Methylierungswerte im *F10*-Locus von Rauchern, Nichtrauchern, Lungenkrebs- und unauffälligem Lungengewebe.

<i>F10</i> (cg08143875)	Lungen-Ca	Raucher	<i>F10</i> (cg08143875)	Lunge	Nichtraucher
Anzahl der Proben	34	12	Anzahl der Proben	33	25
Minimum	5,7%	32,1%	Minimum	23,3%	39,5%
25. Perzentil	29,9%	34,6%	25. Perzentil	27,9%	42,2%
Median	41,4%	39,8%	Median	29,0%	44,1%
75. Perzentil	56,4%	41,7%	75. Perzentil	31,2%	45,6%
Maximum	72,7%	42,5%	Maximum	34,4%	47,9%
Mittelwert	42,7%	38,5%	Mittelwert	29,3%	43,7%
SD	17,0%	3,7%	SD	2,7%	2,2%
SEM	2,9%	1,1%	SEM	0,5%	0,4%
p-Wert	<0,3974		p-Wert	<0,0001	

SD: Standardabweichung. SEM: Standardfehler. p-Wert ermittelt über t-Test für unabhängige Stichproben.

5.4.5 Stärke und Limitation des Vergleichs des genomweiten DNA-Methylierungsmusters von Rauchern und Nichtrauchern

Eine mögliche Stärke der hier durchgeführten Analyse liegt darin, dass zusätzlich Auskünfte über ehemaliges und Gelegenheits-Rauchen vorlagen. Aufgrund einer visuellen Darstellung des Methylierungsmusters der mit dem Rauchen assoziierten Loci in den ehemaligen Rauchern, konnte die tendenzielle Aussage getroffen werden, dass die differentielle DNA-Methylierung in den betroffenen Loci nach Aufgabe des Rauchens wieder in ihren ursprünglichen Zustand zurückkehren könnte. Dies haben Zeilinger *et al.* ebenfalls bereits postuliert [133]. Leider lag der Rauchstatus der GEMINAKAR-Proben erst nach Beendigung der durchgeführten Analysen vor, so dass eine Longitudinal-Analyse diese Vermutung bestätigen oder widerlegen hätte können.

Die wohl größte Limitation der Rauch-Studie ist die Tatsache, dass aufgrund des fehlenden Informationsgehalts sich unter den Nichtrauchern auch Passivraucher befinden könnten. Der damit aus der Umgebungsluft eingeatmete Tabakrauch könnte somit die Ergebnisse beeinträchtigen, da Passivrauchen ebenfalls einen Effekt auf die DNA-Methylierung ausübt. So verursacht Tabakrauch beispielsweise sowohl bei Rauchern als auch Nichtrauchern, die dem Passivrauchen ausgesetzt sind, Lungenkrebs. In einer Studie von Scesnaite *et al.* konnte in beiden Gruppen ein ähnliches DNA-Methylierungsmuster in Lungentumoren identifiziert

werden [413]. Ebenso zeigen Neugeborene, deren Mütter während der Schwangerschaft rauchten, signifikante Veränderungen in ihrem DNA-Methylierungsprofil [136,138].

5.4.6 Fazit der genomweiten DNA-Methylierungsanalyse von Rauchern und Nichtrauchern

Der Vergleich des genomweiten DNA-Methylierungsmusters von Rauchern und Nichtrauchern lieferte sowohl neue als auch bereits publizierte differentiell methylierte CpG-Loci und Gene. Insgesamt könnten diese zu den bekannten mit dem Rauchen verbundenen Risikofaktoren und Folgeerkrankungen beitragen. Dies würde die Rolle des Rauchens als einer der nachhaltigsten Umweltfaktoren für Modifikationen im DNA-Methylierungsprofil bestätigen.

5.5 Nacharbeit und der Einfluss auf das DNA-Methylom

Bisher gibt es nur wenige Studien, in denen die Auswirkungen von Schicht- und Nacharbeit auf das DNA-Methylom analysiert wurden. Bollati *et al.* untersuchten mittels Bisulfit-Pyrosequenzierung die globale (Alu und LINE-1) und genspezifische (*GCR*, *TNF- α* und *IFN- γ*) Methylierung in Zellen des peripheren Blutes von 100 männlichen Arbeitern im 3-Schicht-System und 50 Tagarbeitern und konnten dabei zwischen Nacht- und Tagarbeitern keine signifikanten Unterschiede feststellen [178]. Zhu *et al.* konnten in peripheren Lymphozyten mittels HM27K BC beobachten, dass sich die Methylierung in 5409/27578 CpG-Positionen, von denen viele in der Nähe von Promotoren krebsrelevanter Gene lokalisiert waren, zwischen zehn Langzeit-Schichtarbeiterinnen und zehn Tagesarbeiterinnen signifikant unterschied [176]. Jacobs *et al.* untersuchten mittels HM27K BC das Methylierungsniveau von CpG-Stellen in Promotorregionen imprinteter Gene und fanden signifikant differentielle CpG-Loci in peripheren Lymphozyten von zehn weiblichen Langzeit-Schichtarbeitern im Vergleich zu zehn Tagarbeiterinnen, welchen sie ein karzinogenes Potenzial für Schichtarbeit zuschrieben [177]. Shi *et al.* untersuchten die genomweite Methylierung von CpG-Inseln von 110 mikro-RNAs (miRNAs) in Zellen des peripheren Blutes von Langzeit-Schichtarbeitern und Tagarbeitern und identifizierten dabei 50 differentiell methylierte Loci, die mit 31 miRNAs assoziiert werden konnten, darunter auch eine zirkadian-relevante miRNA, der sie aufgrund der veränderten Methylierung eine Herunterregulierung immunvermittelter Antitumor-Aktivität in Langzeit-Schichtarbeitern zuschrieben [175]. Bhatti *et al.* konnten zwischen 59

Nachtschicht- und 65 Tagschichtarbeitern mittels HM450K BC 16135 signifikant differentiell methylierte Loci identifizieren, einschließlich 21 Loci in zirkadianen Genen, die allesamt in den peripheren Lymphozyten der Nachtschichtarbeiter eine Hypermethylierung aufwiesen und mit insgesamt 3769 Genen assoziiert werden konnten [414].

Insgesamt wurden DNA-Methylierungsanalysen spezifisch an Nachtarbeitern bisher lediglich von Bollati *et al.* [178] durchgeführt. Die Analyse beschränkte sich hierbei jedoch auf die DNA-Methylierung repetitiver Elemente und den Promotorregionen von drei Genen, die aufgrund ihrer Rollen in inflammatorischen und krebsrelevanten Signalwegen ausgewählt wurden.

In der im Rahmen dieser Arbeit vorgenommenen Analyse wurde der Effekt von Nachtarbeit auf die genomweite DNA-Methylierung von insgesamt sechs Nachtarbeitern untersucht. In Zellen des peripheren Blutes der Nachtarbeiter konnte im Vergleich zu den Tagarbeitern durchwegs eine Hypermethylierung in den identifizierten differentiell methylierten Loci (n=21) beobachtet werden, deren Signifikanz nach Korrektur für multiples Testen nicht Stand hielt. Von den 21 Positionen konnten insgesamt 13 CpG-Loci mit elf verschiedenen Genen assoziiert werden, zwei Loci waren dabei in der 5'-UTR, vier Loci in der Nähe der TSS, sechs CpG-Loci im Genkörper und ein Locus in der 3'-UTR lokalisiert.

Die Gene *CREB5* und *TRIM10* waren von jeweils zwei benachbarten CpG-Loci betroffen, die in den Nachtarbeitern eine mittlere Hypomethylierung von 10,4% bzw. 11,1% aufwiesen. *CREB5* (*cAMP responsive element binding protein 5*) konnte bereits mit dem metabolischen Syndrom assoziiert werden [415] und ist u.a. für die Fettzell-Differenzierung verantwortlich. Die zwei hier identifizierten betroffenen in Nachtarbeitern hypomethylierten Loci waren in der 5'-UTR lokalisiert. Da die Hypomethylierung in der Region vor Genen, wie in 5'-UTR und TSS, mit einer erhöhten Genexpression assoziiert ist [236], könnte demnach die beobachtete Hypomethylierung in der S-Shelf-Region einer 1312bp großen CpG-Insel in der 5'-UTR Nahe der TSS von *CREB5* zu einer erhöhten Expression des Gens in Nachtarbeitern führen und somit möglicherweise eine Verbindung zum metabolischen Syndrom herstellen. Eine erhöhte Expression von *CREB5* konnte bereits im Bauchfettgewebe von Frauen mit metabolischem Syndrom beobachtet werden [415].

Das Protein TRIM10 ist in der terminalen Differenzierung erythroider Zellen involviert [416], dem Prozess, bei dem aus unreifen Vorläuferzellen Erythrozyten gebildet werden [417].

Diskussion

Bisher konnte noch keine Assoziation zwischen Schicht- bzw. Nachtarbeit und Erythropoese erfasst werden. Die beobachtete Hypomethylierung im Genkörper von *TRIM10* könnte zu einer reduzierten Genexpression beitragen.

Insgesamt konnten drei differentiell methylierte CpG-Loci mit einer mittleren Methylierungsdifferenz >13% ermittelt werden. Während der CpG-Locus cg00604771 in einer 668bp großen CpG-Insel auf Chromosom 16 den größten Methylierungsunterschied aufwies (13,2%) und mit keinem Gen assoziiert werden konnte, wurde ein Methylierungsunterschied von jeweils 13,1% zwischen Nacht- und Tagarbeitern mit den Genen *MED1* und *TMEM132D* assoziiert. Hiervon war ein CpG-Locus in der 3'-UTR des *MED1*-Gens betroffen. *MED1* (*mediator complex subunit 1*) hat eine wichtige Rolle in der Adipogenese inne [418] und konnte im Mausmodell mit einer Beteiligung an der Ausbildung einer Fettleber assoziiert werden [419]. Ferner spielt *MED1* eine entscheidende Rolle in der transkriptionellen Kontrolle von Genen, die mit der Energiehomöostase in Verbindung gebracht wurden [420]. Die Hypomethylierung von *TMEM132D* (Transmembran-Protein 132D) trat im Genkörper in einem Enhancer-Element auf, was zu einer verminderten Genexpression in Nachtarbeitern führen könnte. Über die Funktion von *TMEM132D* ist in der Literatur bisher wenig beschrieben, es konnte jedoch eine Genvariante mit Angst- und Panikstörungen bei Individuen europäischer Abstammung assoziiert werden. Diese bedingte im frontalen Cortex eine gesteigerte Expression, was zu einer Störung der neuronalen Kommunikation führen könnte und somit die Auslösung von Panikattacken begünstigt [421,422]. Demnach wäre eine erhöhte Expression von *TMEM132D* mit Angststörungen assoziiert. Die identifizierte Risikovariante wurde in Intron 3 ermittelt, während der hier identifizierte hypomethylierte Locus im ersten Intron des *TMEM132D*-Gens lokalisiert ist. Größere Angstzustände konnten bereits mit Schichtarbeit assoziiert werden [423,424]. Darüber hinaus konnte unter den differentiell methylierten Loci ein weiterer Locus mit einer Assoziation zu Panikstörungen bestätigt werden. Dieser befand sich im N-Shore-Bereich einer 938bp großen CpG-Insel in der TSS1500 des Gens *SLC6A1*, welches für den GABA Transporter 1 (*GAT-1*) kodiert. Thoeringer *et al.* konnten hier eine Assoziation zwischen spezifischen Genvarianten in der 5'-UTR von *SLC6A1* mit der Ätiopathogenese von Angststörungen einschließlich Panikattacken herstellen [425]. Eine Hypomethylierung in der TSS1500 könnte zu einer erhöhten Genexpression führen. In *GAT-1*-Knockout-Mäusen konnte im Gegensatz zu Wildtyp-Mäusen ein niedrigeres Ausmaß an Depressions- und

Angst-ähnlichem Verhalten festgestellt werden [425]. Daher könnte eine erhöhte Expression von *GAT-1* mit Angststörungen einhergehen. Des Weiteren konnten Chiu *et al.* ebenfalls in *GAT-1*-Knockout-Mäusen zeigen, dass selbst geringe Unterschiede in der Expression von *GAT-1* innerhalb des ZNS starke Auswirkungen auf emotionales Verhalten wie Furcht und Angst haben könnten [427].

In der Studie von Bhatti *et al.* konnten nur etwa 9% der 16135 hypermethylierten Loci mit einer mittleren Methylierungsdifferenz von $\geq 5\%$ zwischen den beiden Gruppen identifiziert werden [414]. Unter den von Bhatti *et al.* in Nachtschichtarbeitern identifizierten signifikant hypermethylierten Loci und Genen konnte mit der hier durchgeführten Analyse eine Übereinstimmung von zwei Loci (cg03710889 und cg16912957) und drei Genen (*SLC11A1*, *TMEM132D*, *TRIM10*) gefunden werden, die hier jedoch in Nachtarbeitern eine Hypomethylierung zeigten. Im Gegensatz zur Bhatti-Studie, in der Tagschichtarbeiter mit Nachtschichtarbeitern, die ebenfalls auch partiell tagsüber beschäftigt waren, wurden in der hier vorliegenden Analyse ausschließliche Tagarbeiter und Nachtarbeiter untersucht. Der Locus im Genkörper von *SLC11A1* (*solute carrier family 11, member 1*) lag im S-Shore-Bereich einer 210bp großen CpG-Insel und könnte in Nachtarbeitern mit einer verminderten Expression assoziiert werden. *SLC11A1* wird von Makrophagen und Neutrophilen exprimiert und spielt bei der angeborenen Resistenz gegen Infektionen durch intrazelluläre Mikroben eine Schlüsselrolle [428]. Eine verminderte Expression könnte somit für eine erhöhte Anfälligkeit für Infektionen sprechen, die bereits in Schichtarbeitern beobachtet werden konnte [429]. Zudem konnte *SLC11A1* zusammen mit dem differentiell methylierten Genort *CLSTN2* (calsyntenin 2) mittels Anreicherungsanalyse in überrepräsentierten Erkrankungen, die auf einer Infektion mit einem Magen-Darm-Virus zurückzuführen sind, identifiziert werden. Ferner konnte eine reduzierte Genexpression von *SLC11A1* in Brustkrebszelllinien und humanem Brustkrebsgewebe beobachtet werden [430]. Bei Frauen konnte Nachtarbeit bereits mit einem erhöhten Risiko für Brust- und Eierstockkrebs [163,270,431–433] assoziiert werden, das von der Dauer der Beschäftigung und der Anzahl der geleisteten Nachtschichten abhängig zu sein scheint. Bei Männern hingegen konnte Nachtarbeit mit verschiedenen Krebsentitäten assoziiert werden, darunter beispielsweise Lungen-, Prostata-, Darm- und Pankreaskrebs [434].

Der Vergleich von Schichtarbeitern mit Nachtschicht und Tagarbeitern lieferte 14 differentiell methylierte CpG-Loci mit geringen Methylierungsdifferenzen. Hier konnte eine Übereinstimmung in einem CpG-Locus (cg08240847) identifiziert werden, der auf Chromosom 8 in einem Enhancer-Element lokalisiert war und keinem annotierten Gen zugeordnet werden konnte, und auch hier in Nachtschichtarbeitern eine Hypomethylierung aufwies. Ferner wurde eine Übereinstimmung von zwei assoziierten Genen beobachtet, die ebenfalls jeweils in der gleichen Genregion lokalisiert waren. Während der CpG-Locus im Genkörper von *C15orf27* bei Nachtschichtarbeitern ebenfalls hypermethyliert vorlag, zeigte der CpG-Locus in einer CpG-Insel in der 5'-UTR im Promotorbereich von *FAM55C* in diesen eine Hypomethylierung. Da die Methylierungsunterschiede zwischen Schichtarbeitern mit Nachtschicht und Tagarbeitern aber sehr gering ausfielen, sind die möglichen Auswirkungen auf die Genexpression ungewiss und die entsprechenden Gene wurden daher nicht weiter analysiert.

5.5.1 Limitationen der Nachtarbeits-Methylierungsanalyse

Die größte Limitation dieser Analyse lag in der sehr geringen Fallzahl in für Nachtarbeit diskordanter Zwillingspaare (n=6), weshalb die Signifikanz nach Korrektur für multiples Testen nicht bestehen blieb.

Eine weitere Limitation lag im Zeitpunkt der relevanten Blutentnahme der Probanden begründet, die in den Morgenstunden, gewöhnlich vor 12:00 Uhr, teilweise aber auch in den Nachmittagsstunden stattfand. Dabei wurde nicht berücksichtigt bzw. vermerkt, ob es sich bei den Nachtarbeitern um aktive Nachtarbeiter handelte, also um Personen, die zum Zeitpunkt der Blutentnahme gerade von einer Nachtschicht kamen. Bei einer Blutentnahme unmittelbar nach der durchgeführten Nachtarbeit wären aufgrund der Desynchronisation des zirkadianen Systems sicher die größten Effekte zu erwarten gewesen.

Weiterhin konnte in einer aktuellen Studie an humanen Gehirnproben gezeigt werden, dass das DNA-Methylierungsmuster einiger Gene im 24-Stunden-Rhythmus schwankt und die DNA-Methylierung besonders in der Nähe von TSS der Gene dabei mit der zirkadianen Expression korrelierte [435]. Da der beobachtete 24-Stunden-Rhythmus der DNA-Methylierung möglicherweise auch in anderen Gewebetypen, beispielsweise dem Blut, präsent sein könnte, müsste das Methylierungsmuster der (Nacht-)Schicht- und Tagarbeiter

demnach synchronisiert betrachtet und verglichen werden, um tatsächliche mit dem Schichtsystem assoziierte Abweichungen in diesen Phänotypen erfassen zu können.

Des Weiteren wurde zur Validierung ausgewählter Loci mittels Bisulfit-Pyrosequenzierung für die DNA-Proben ein anderes Konvertierungs-Kit als für den HM450K BC verwendet. Während für die arraybasierte Methylierungsanalyse das Kit von Zymo Research genutzt wurde, welches eine sehr aggressive Bisulfit-Konvertierung beinhaltet, wurde für die Bisulfit-Pyrosequenzierung das weniger aggressive Kit von Qiagen verwendet, da dieses für diese Methode im Institut für Humangenetik verwendet wird. Eine aggressive Bisulfit-Behandlung gewährleistet zwar eine vollständige Konversion von nicht methyliertem Cytosin zu Uracil, hierdurch werden aber auch höhere Spiegel degradierter DNA ermöglicht, was die Anzahl an verfügbaren DNA-Molekülen für eine anschließende Amplifikation mittels PCR reduziert. Die weniger aggressive Behandlung hingegen kann aufgrund der Detektion nicht vollständig konvertierter Cytosine zu einer möglichen Überschätzung des Methylierungsniveaus führen [436]. Dies stellte folglich eine potentielle Fehlerquelle bei der Validierung der zehn ausgewählten Loci dar. Da die DNA-Proben aller Zwillinge zum gleichen Zeitpunkt mit der gleichen Konvertierungs-Methode behandelt wurden, dürften die entsprechenden Methylierungsdifferenzen der für Nachtarbeit diskordanten Zwillingspaare hingegen nicht beeinflusst werden. Dennoch konnten die Methylierungsdifferenzen für die ausgewählten Loci der Nachtarbeits-Signatur nicht erfolgreich validiert werden. Mittels weiterer Regressionsanalysen konnte bestätigt werden, dass weder ein einzelner Locus noch ein einzelnes Zwillingsspaar für das Gesamtbild der gescheiterten Validierung verantwortlich war, welches somit vordergründig der mangelnden Signifikanz nach Korrektur für multiples Testen zugeschrieben wurde.

5.5.2 Fazit der zwillingsbasierten Analyse zur Untersuchung des Einflusses von Nachtarbeit auf die genomweite DNA-Methylierung

Mit dem für Nachtarbeit diskordanten Zwillingmodell konnten zwar einige differentiell methylierte Loci ermittelt werden, die nach Korrektur für multiples Testen jedoch nicht mehr signifikant waren. Die Methylierungsveränderung der korrespondierenden Gene könnte dabei mit bekannten Risikofaktoren von Schichtarbeit, wie dem metabolischen Syndrom, einer erhöhten Anfälligkeit für Infektionen sowie Angststörungen korrelieren. Die besagten Loci konnten mit einer unabhängigen Methode nicht validiert werden.

5.6 Synthese der Einzelbefunde und Ausblick der verschiedenen DNA-Methylierungsanalysen

In der vorliegenden Arbeit wurde das genomweite Methylierungsprofil von monozygoten und dizygoten Zwillingen entsprechend der zu untersuchenden Auswirkungen von Alter, BMI, Rauchen und Schicht- bzw. Nachtarbeit auf das DNA-Methylom analysiert, wobei zahlreiche Modifikationen in der DNA-Methylierung identifiziert werden konnten. Für alle hier durchgeführten Teilstudien konnten sowohl bereits publizierte mit dem jeweiligen Einflussfaktor korrelierende Methylierungsveränderungen von Genen bestätigt als auch von bisher damit nicht in Zusammenhang gebrachten neuen Genen identifiziert werden. Einige der hier neu identifiziert differentiell methylierten Gene konnten dabei in einen plausiblen biologischen Kontext zur jeweiligen Fragestellung gebracht werden. Darüber hinaus konnten einige Gene, die bisher nur über bestimmte Genvarianten oder Mutationen direkt oder indirekt (über die damit assoziierten Folgeerkrankungen) mit Alterung, Adipositas, Rauchen und Schicht- bzw. Nachtarbeit in Verbindung gebracht wurden, mit einer möglichen epigenetischen Beteiligung (in Form einer veränderten DNA-Methylierung) an diesen Prozessen assoziiert werden.

Insgesamt liefern die Ergebnisse dieser Arbeit mehrere Ansatzpunkte für mögliche Folgeprojekte und neue Studien.

Die Ergebnisse der einzelnen Teilstudien dieser Arbeit wurden im Rahmen des von der Europäischen Union geförderten Interreg 4A Syddanmark-Schleswig-K.E.R.N. Projektes "SAME" erhoben. Ein Vorteil von SAME bestand darin, dass auf das DNA-Material der Baselinestudie GEMINAKAR zugegriffen werden konnte, welches den SAME-Probanden bereits vor 10–15 Jahren entnommen wurde und somit theoretisch für longitudinale Analysen hätte dienen können. Generell werden Longitudinalanalysen für spezifische Fragestellungen, die beispielsweise ein bestimmtes Merkmal betreffen können, zu verschiedenen Zeitpunkten an demselben Studienkollektiv durchgeführt. Somit könnte eine longitudinale Analyse im Rahmen dieser Arbeit beispielsweise Aufschluss darüber geben, ob die durch das Rauchen bedingte veränderte DNA-Methylierung nach Aufgabe des Rauchens auf ihr ursprüngliches Niveau zurückkehren könnte wie es in den PCAs in Kapitel 4.2.5.1 (vgl. Abbildung 29B) bzw. 4.2.5.2 (vgl. Abbildung 32B) bereits angedeutet werden konnte. Diese Fragestellung konnte in bisher publizierten Arbeiten ebenfalls nur unzureichend beantwortet

werden, da es sich bei diesen Studien gleichermaßen um die Resultate von Querschnittstudien handelte [133,437,438]. Somit konnte die bei ehemaligen Rauchern beobachtete von der Anzahl der rauchfreien Jahre abhängige Näherung der Methylierung von definierten CpG-Loci an das für Nichtraucher charakteristische Methylierungsniveau bisher nur vorgeschlagen werden. Die Angaben über das Rauchverhalten der Probanden von vor 10–15 Jahren liegen nun, nach Fertigstellung aller für diese Arbeit durchgeführten Analysen, vor. Es wäre daher interessant, sich für diese Fragestellung zumindest die über die FDR ermittelten 22 differentiell methylierten CpG-Positionen in einem unsupervidierten Ansatz in den korrespondierenden GEMINAKAR-Proben anzusehen. Somit könnte über eine Longitudinalstudie nicht nur bestätigt werden, ob die ursprüngliche DNA-Methylierung an den besagten Loci nach Aufhören des Rauchens tatsächlich wiederhergestellt werden kann, sondern auch, ob die Dauer des regelmäßigen Rauchens zu einer progressiven oder relativ stabilen Methylierungsveränderung führt.

Des Weiteren lieferten die in dieser Arbeit durchgeführten DNA-Methylierungsanalysen Hinweise auf einige Gene, deren verändertes Methylierungsmuster zwar mit dem untersuchten Phänotyp korrelierte, die ermittelten Unterschiede zwischen den zu vergleichenden Gruppen aber nach Korrektur für multiples Testen (FDR-Methode nach Benjamini und Hochberg) nicht mehr signifikant waren. Um eine höhere statistische Power zu erreichen, sollten für zukünftige Arbeiten einerseits mehr Probanden rekrutiert und andererseits bevorzugt monozygote Zwillinge untersucht werden. Der große Vorteil des monozygoten Zwillingsmodells basiert auf der Tatsache, dass die Zwillinge eines Paares eine nahezu identische genetische Ausstattung aufweisen. Aufgrund der primären Fragestellung des SAME-Projektes wurden für diese Arbeit nur Zwillingspaare rekrutiert, die sich für Schichtarbeit diskordant verhielten bzw. Zwillinge, die einer Arbeit bei Nacht nachgingen. Folglich erfolgten die in dieser Arbeit durchgeführten Methylierungsanalysen, mit denen die Auswirkungen eines hohen BMIs und des Rauchens untersucht wurden, ausschließlich über gruppenbasierte Analysen (Statistik für unabhängige Stichproben). Für zukünftige Analysen wäre es daher sinnvoll, speziell für diese Merkmale diskordante monozygote Zwillingspaare zu rekrutieren, um mittels statistischer Verfahren für verbundene Stichproben eine höhere Teststärke gewährleisten zu können.

Ferner konnten in dieser Arbeit mittels verschiedener Anreicherungsanalysen signifikante Häufungen der differentiell methylierten Gene in funktionellen Kategorien, Stoffwechselwegen sowie Krankheitsbildern ermittelt werden, die sich teilweise in einen von der Fragestellung abhängig plausiblen biologischen Zusammenhang bringen lassen konnten. Um mögliche Assoziationen zwischen den beobachteten Veränderungen im DNA-Methylierungsmuster und dem Phänotyp mit den damit in dieser Arbeit in Verbindung gebrachten biologischen Prozessen zu bestätigen, wären ergänzende Studien bzw. funktionelle Analysen sinnvoll. So könnten z.B. Expressionsstudien Aufschluss darüber geben, ob die veränderte Methylierung von ausgewählten Kandidatengenen tatsächlich zu einer Änderung der Expression führt. Ebenso sollten in weiterführenden Analysen die öffentlich zugänglichen ENCODE-Daten miteinbezogen werden [439], beispielsweise ChIP-Seq, eine Technik zur Analyse von *in vivo* Protein-DNA-Interaktionen [440,441].

Neben der DNA-Methylierung, die im Fokus dieser Arbeit stand, gewinnt jüngst auch die DNA-Hydroxymethylierung an Bedeutung. So gibt es bereits Studien, die eine altersabhängige Veränderung in der Hydroxymethylierung in Gehirn [442] und Leberzellen [443] von Mäusen zeigen konnten. Ebenso konnte ein Zusammenhang von einigen Umwelteinflüssen und Erkrankungen mit gewebespezifischen Veränderungen in der DNA-Hydroxymethylierung gefunden werden [444]. In dieser Arbeit beruhte die Detektion des Methylierungsstatus der Cytosine auf der Bisulfitbehandlung der DNA. Mit dieser Methode kann jedoch nicht zwischen Methylierung und Hydroxymethylierung unterschieden werden [445], so dass möglicherweise an einigen Genloci nicht das Niveau der Methylierung, sondern das der Hydroxymethylierung gemessen wurde. Dies könnte durchaus von Bedeutung sein, zumal 5-Hydroxymethylcytosin mit einer aktiven DNA-Demethylierung in Verbindung steht [94]. Als nächste Forschungsarbeit könnte daher die Analyse von 5-Hydroxymethylcytosin an ausgewählten hier als differentiell methyliert identifizierten CpG-Positionen durchgeführt bzw. die durchgeführten Teilstudien um diese Analyse ergänzt werden.

6 Zusammenfassung

Umweltfaktoren können über epigenetische Mechanismen das Expressionsmuster von Genen verändern. So kann eine Veränderung des Lebensstils, z.B. Zigarettenkonsum, durch epigenetische Modifikationen, wie die der DNA-Methylierung, wesentlich zur Pathogenese einiger Erkrankungen beitragen. In der hier vorliegenden Arbeit sollte der Einfluss des Alters, eines hohen *Body Mass Index* (BMI), des Rauchens und der Schichtarbeit auf das DNA-Methylierungsmuster von Zwillingen mittels arraybasierter Technologie analysiert und charakterisiert werden. Die Untersuchung von monozygoten Zwillingen hat dabei den Vorteil, dass aufgrund des gleichen genetischen Hintergrundes der Zwillinge, Veränderungen im DNA-Methylierungsmuster den analysierten Parametern zugeschrieben werden können. Aufgrund des auf der primären Fragestellung dieser Arbeit basierenden Studiendesigns, konnte ein zwillingsbasierter Ansatz nur für die Analyse des Einflusses der Schichtarbeit gewählt werden. Für die Parameter Alter, BMI und Rauchen wurde für die Identifizierung einer veränderten DNA-Methylierung eine gruppenbasierte Analyse mittels der sog. *False Discovery Rate* (FDR, q) bzw. eines definierten Signifikanzwertes bei einer gleichzeitigen Differenz der mittleren Methylierungswerte zwischen zwei Gruppen von mindestens 5% ($|\Delta\beta| > 0,05$) durchgeführt. Zusammengefasst wurden folgende Befunde erzielt:

1. Der Infinium HumanMethylation450 BeadChip konnte in einem Vorexperiment erfolgreich etabliert und somit für die anschließenden genomweiten hypothesenfreien DNA-Methylierungsanalysen eingesetzt werden. Diese wurden an aus Zellen des peripheren Blutes isolierter genomischer DNA von monozygoten und dizygoten Zwillingen durchgeführt.
2. Für die Identifizierung altersabhängig methylierter CpG-Loci wurden insgesamt 53 mono- und 63 dizygoten DNA-Proben von Zwillingen im Alter von 20–63 Jahren einbezogen. Mittels linearer Regression konnten 111 CpGs bzw. 63 Gene identifiziert werden, deren Methylierung signifikant mit dem Alter korrelierte ($q < 1 \times 10^{-11}$). 47 Gene bzw. 58 Loci konnten dabei die Befunde altersabhängiger Methylierungsveränderungen aus anderen Studien bestätigen, während 16 erstmals neu identifizierte Gene ermittelt werden konnten. Über eine Genontologie-Analyse konnten unter den insgesamt 63 altersabhängig methylierten Genen statistisch signifikante Anreicherungen für Gene ermittelt werden, die an Prozessen des Blutdrucks, der Herzkontraktion und des Nervensystems beteiligt sind.

Zusammenfassung

3. Der Vergleich von fünf adipösen und 29 normalgewichtigen Zwillingen lieferte bei einer FDR $q < 0,05$ drei differentiell methylierte CpGs, die den Genen *TREML4* und *POM121C* zugeordnet werden konnten. Mit einem zweiten Ansatz wurden insgesamt 80 CpGs bzw. 43 Gene ermittelt, deren Methylierung mit einem hohen BMI korrelierte ($p < 3 \times 10^{-4}$; $|\Delta\beta| > 0,05$). 3/43 differentiell methylierte Gene (*ADORA1*, *LPGAT1*, *POM121C*) wurden bereits im Zusammenhang mit Adipositas publiziert, während zwei weitere Gene (*ADCY5*, *SYN2*) mit einer damit verbundenen Folgeerkrankung, dem Typ-2-Diabetes, assoziiert werden konnten.

4. Der Vergleich von 12 Rauchern und 25 Nichtrauchern resultierte bei einer FDR $q < 0,05$ in insgesamt 22 differentiell methylierten CpGs bzw. 14 Genen. Ein zweiter Ansatz lieferte insgesamt 131 bzw. 68 differentiell methylierte CpG-Loci bzw. Gene ($p < 4 \times 10^{-4}$; $|\Delta\beta| > 0,05$). Mit beiden Ansätzen konnten Loci und Gene bestätigt werden, für die andere Studien bereits eine mit dem Rauchen veränderte DNA-Methylierung gezeigt hatten. Insgesamt konnten unter den identifizierten Genen mittels verschiedener Anreicherungsanalysen Gene ermittelt werden, die z.B. in Herzerkrankungen (*CACNA1C*, *GATA4*, *EVC2*, *NODAL*) und Thrombose (*F2RL3*, *F10*), deren Entstehung teilweise durch das Rauchen begünstigt wird, involviert sind.

5. Bei sechs für Nachtarbeit diskordanten Zwillingspaaren konnten insgesamt 21 bzw. elf differentiell methylierte Loci bzw. Gene ermittelt werden ($p < 0,05$; $|\Delta\beta| \geq 0,05$). Von diesen konnten zehn ausgewählte Loci nicht erfolgreich validiert werden, was primär der geringen statistischen Power zugeschrieben wurde. Dennoch wurden 2/21 Loci bzw. 3/11 Gene (*SLC11A1*, *TMEM132D*, *TRIM10*) bereits in anderen Studien mit einer veränderten Methylierung bei Nachtschichtarbeitern identifiziert, was eine mögliche Auswirkung der Nachtarbeit auf den veränderten Methylierungsstatus dieser Gene unterstreicht. Weiterhin konnten für einige der von differentieller Methylierung betroffenen Gene Assoziationen zu bekannten Risikofaktoren von Schichtarbeit, z.B. dem metabolischen Syndrom (*CREB5*), erhöhter Anfälligkeit für Infektionen (*SLC11A1*) sowie Angststörungen (*SLC6A1*, *TMEM132D*) gefunden werden.

Insgesamt lässt sich schlussfolgern, dass die untersuchten Einflussfaktoren mit spezifischen Veränderungen in der DNA-Methylierung korrelieren. Dabei konnten Loci und Gene aus den einzelnen Teilstudien bestätigt und zumindest die Beteiligung einiger der betroffenen Gene in einen entsprechenden plausiblen biologischen Kontext gebracht werden.

7 Summary

Through epigenetic mechanisms environmental factors are able to change the expression pattern of genes. Thus, a change in lifestyle, such as cigarette consumption, may significantly contribute amongst others to the pathogenesis of some diseases through epigenetic modifications, such as DNA methylation. In the present thesis the influence of age as well as of a high body mass index (BMI), smoking and shiftwork on the DNA methylation pattern of twins should be analyzed and characterized through an array-based approach. Investigation of monozygotic twins offers the advantage that due to the identical genetic background of these twins, changes in DNA methylation patterns can be attributed to the analyzed parameters. Due to the study design based on the primary question of this thesis, a twin-based approach could be used only for the analysis of the influence of shiftwork. For the parameters age, BMI and smoking a group-based approach was used for the identification of changes in DNA methylation by applying the false discovery rate (FDR, q) and a defined level of significance with a simultaneously difference of the mean methylation values between two groups of at least 5% ($|\Delta\beta| > 0.05$), respectively. Summarized the following results were obtained:

1. The Infinium HumanMethylation450 BeadChip could be successfully validated in a pre-experiment and could therefore be used for subsequent genome-wide hypothesis free DNA methylation analyses. These were performed in whole blood genomic DNA from monozygotic and dizygotic twins.
2. For identification of age dependent methylated CpG loci a total of 53 monozygotic and 63 dizygotic DNA samples of twins at age of 20–63 years were included. By using the linear regression model 111 CpGs and 63 genes, respectively, could be identified whose methylation significantly correlated with age ($q < 1 \times 10^{-11}$). 47 genes and 58 loci, respectively, could be confirmed that have already been related to an age-dependent methylation by previous studies whereas 16 newly identified genes could be detected. Performing gene ontology analysis for the 63 age-dependent methylated genes, statistically significant enrichments could be found for genes involved in processes of blood pressure, heart contraction and nervous system.

Summary

3. The comparison of five obese and 29 normal weighted twins resulted in three differentially methylated CpGs associated with the genes *TREML4* and *POM121C* by applying a FDR $q < 0.05$. In a second approach a total of 80 CpGs and 43 genes, respectively, could be identified that correlated with a high BMI ($p < 3 \times 10^{-4}$; $|\Delta\beta| > 0.05$). 3/43 differentially methylated genes (*ADORA1*, *LPGAT1*, *POM121C*) were already published with association to adiposity whereas two further genes (*ADCY5*, *SYN2*) could be associated with a secondary disease, type 2 diabetes.

4. The comparison of 12 smokers and 25 nonsmokers resulted in a total of 22 differentially methylated CpGs and 14 genes, respectively, by applying a FDR $q < 0.05$. A second approach resulted in a total of 131 and 68 differentially methylated CpG loci and genes, respectively ($p < 4 \times 10^{-4}$; $|\Delta\beta| > 0.05$). With both approaches loci and genes could be confirmed of which other studies have already reported an altered DNA methylation due to tobacco smoking. All in all, with different enrichment analyses differentially methylated genes could be detected to be involved in e.g. heart diseases (*CACNA1C*, *GATA4*, *EVC2*, *NODAL*) and thrombosis (*F2RL3*, *F10*) whose development is partially promoted by smoking.

5. In six twin pairs discordant for night work a total of 21 and eleven differentially methylated loci and genes, respectively, could be detected ($p < 0.05$; $|\Delta\beta| \geq 0.05$). For ten selected loci validation failed which may be attributed to low statistical power. However, 2/21 loci and 3/11 genes (*SLC11A1*, *TMEM132D*, *TRIM10*), respectively, have already been identified in night shift workers with an altered DNA methylation by other studies, emphasizing a possible effect of night work on the methylation of these genes. Furthermore, for some of the differentially methylated genes associations to known risk factors of shiftwork such as the metabolic syndrome (*CREB5*), increased susceptibility to infections (*SLC11A1*) and anxiety disorders (*SLC6A1*, *TMEM132D*) could be found.

All in all it can be concluded that the investigated influence factors correlate with specific alterations in DNA methylation. Loci and genes from substudies could be confirmed and at least involvement of some of the affected genes may be plausible in the relevant biological context.

8 Literaturverzeichnis

1. Hattori M, Taylor TD. The human intestinal microbiome: a new frontier of human biology. *DNA Res.* 2009;16(1):1–12.
2. Doerfler W. In pursuit of the first recognized epigenetic signal – DNA methylation: a 1976 to 2008 synopsis. *Epigenetics.* 2008;3(3):125–33.
3. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature.* 2004;429(6990):457–63.
4. Luczak MW, Jagodziński PP. The role of DNA methylation in cancer development. *Folia Histochem Cytobiol.* 2006;44(3):143–54.
5. Ekström TJ. Epigenetic control of gene expression. *Biochim Biophys Acta.* 2009;1790(9):845–6.
6. Gundersen GG, Worman HJ. Nuclear positioning. *Cell.* 2013;152(6):1376–89.
7. Volle C, Dalal Y. Histone variants: the tricksters of the chromatin world. *Curr Opin Genet Dev.* 2014;25:8–14,138.
8. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol.* 2010;28(10):1057–68.
9. Matouk CC, Marsden PA. Epigenetic regulation of vascular endothelial gene expression. *Circ Res.* 2008;102(8):873–87.
10. Kristeleit R, Stimson L, Workman P, Aherne W. Histone modification enzymes: novel targets for cancer drugs. *Expert Opin Emerg Drugs.* 2004;9(1):135–54.
11. Rajender S, Avery K, Agarwal A. Epigenetics, spermatogenesis and male infertility. *Mutat Res.* 2011;727(3):62–71.
12. Grewal SI, Moazed D. Heterochromatin and epigenetic control of gene expression. *Science.* 2003;301(5634):798–802.
13. Luger K, Mäder AW, Richmond RK, et al. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* 1997;389(6648):251–60.
14. Luger K. Nucleosomes: structure and function. In: *Encyclopedia of Life Sciences.* 2001:1–8
15. Saraiva NZ, Oliveira CS, Garcia JM. Histone acetylation and its role in embryonic stem cell differentiation. *World J Stem Cells.* 2010;2(6):121–6.
16. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res.* 2011;21(3):381–95.
17. Kuroishi T, Rios-Avila L, Pestinger V, et al. Biotinylation is a natural, albeit rare, modification of human histones. *Mol Genet Metab.* 2011;104(4):537–45.
18. Schübeler D, MacAlpine D, Scalzo D, et al. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* 2004;18(11):1263–71.
19. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res.* 2011;21(4):564–78.
20. Liu X, Yang S, Zhao M, et al. Transcriptional repression by histone deacetylases in plants. *Mol Plant.* 2014;7(5):764–72.
21. Nesbitt AM, McCurdy RD, Bryant SM, Alter MD. Total levels of hippocampal histone acetylation predict normal variability in mouse behavior. *PLoS One.* 2014;9(5):e94224.
22. Kurdistani SK, Tavazoie S, Grunstein M. Mapping global histone acetylation patterns to gene expression. *Cell.* 2004;117(6):721–33.
23. Li F, Huarte M, Zaratiegui M, et al. Lid2 is required for coordinating H3K4 and H3K9 methylation of heterochromatin and euchromatin. *Cell.* 2008;135(2):272–83.

Literaturverzeichnis

24. Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol.* 2005;6(11):838–49.
25. Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006;15 Spec No 1:R17–29.
26. Khraiwesh B, Arif MA, Seumel GI, et al. Transcriptional control of gene expression by microRNAs. *Cell.* 2010;140(1):111–22.
27. Schmitz KM, Mayer C, Postepska A, Grummt I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* 2010;24(20):2264–9.
28. Kim YO, Park SJ, Balaban RS, et al. A functional genomic screen for cardiogenic genes using RNA interference in developing *Drosophila* embryos. *Proc Natl Acad Sci U S A.* 2004;101(1):159–64.
29. Collins LJ, Schönfeld B, Chen XS. The epigenetics of non-coding RNA. In: *Handbook of epigenetics: the new molecular and medical genetics.* 2011:49–61.
30. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci.* 2006;31(2):89–97.
31. Razin A, Riggs AD. DNA methylation and gene function. *Science.* 1980;210(4470):604–10.
32. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. *Science.* 2001;293(5532):1068–70.
33. Bestor TH. The DNA methyltransferases of mammals. *Hum Mol Genet.* 2000;9(16):2395–402.
34. Richardson B. Primer: epigenetics of autoimmunity. *Nat Clin Pract Rheumatol.* 2007;3(9):521–7.
35. Robertson KD. DNA methylation, methyltransferases, and cancer. *Oncogene.* 2001;20(24):3139–55.
36. Goll MG, Kirpekar F, Maggert KA, et al. Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2. *Science.* 2006;311(5759):395–8.
37. Subramaniam D, Thombre R, Dhar A, Anant S. DNA methyltransferases: a novel target for prevention and therapy. *Front Oncol.* 2014;4:80.
38. Deplus R, Brenner C, Burgers WA, et al. Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Res.* 2002;30(17):3831–8.
39. Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem.* 2005;74:481–514.
40. Leonhardt H, Page AW, Weier HU, Bestor TH. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell.* 1992;71(5):865–73.
41. Bartee L, Malagnac F, Bender J. Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev.* 2001;15(14):1753–8.
42. Ramsahoye BH, Biniszkiwicz D, Lyko F, et al. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A.* 2000;97(10):5237–42.
43. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature.* 2009;462(7271):315–22.
44. Tomizawa S, Kobayashi H, Watanabe T, et al. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development.* 2011;138(5):811–20.
45. Lister R, Mukamel EA, Nery JR, et al. Global epigenomic reconfiguration during mammalian brain development. *Science.* 2013;341(6146):1237905
46. Ehrlich M, Gama-Sosa MA, Huang LH, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* 1982;10(8):2709–21.
47. Razin A, Cedar H. DNA methylation in eukaryotic cells. *Int Rev Cytol.* 1984;92:159–85.

Literaturverzeichnis

48. Miranda TB, Jones PA. DNA Methylation: the nuts and bolts of repression. *J Cell Physiol.* 2007;213(2):384–90.
49. Caiafa P, Zampieri M. DNA methylation and chromatin structure: the puzzling CpG islands. *J Cell Biochem.* 2005;94(2):257–65.
50. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860–921.
51. Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. *Genomics.* 2011;98(4):288–95.
52. Sandoval J, Heyn H, Moran S, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics.* 2011;6(6):692–702.
53. Gundersen G, Kolstø AB, Larsen F, Prydz H. Tissue-specific methylation of a CpG island in transgenic mice. *Gene.* 1992;113(2):207–14.
54. Zhu Y, Pandya BJ, Choi HK. Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007–2008. *Arthritis Rheum.* 2011;63(10):3136–41.
55. Cheong J, Yamada Y, Yamashita R, et al. Diverse DNA methylation statuses at alternative promoters of human genes in various tissues. *DNA Res.* 2006;13(4):155–67.
56. Bird A. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* 1980;8(7):1499–504.
57. Sved J, Bird A. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc Natl Acad Sci U S A.* 1990;87(12):4692–6.
58. Shen J, Rideout WM 3rd, Jones PA. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* 1994;22(6):972–6.
59. Romeo G, Hassant HJ, Staempfli S, et al. Hereditary thrombophilia: Identification of nonsense and missense mutations in the protein C gene. *Proc Natl Acad Sci U S A.* 1987;84(9):2829–32.
60. Cooper DN, Youssoufian H. The CpG dinucleotide and human genetic disease. *Hum Genet.* 1988;78(2):151–5.
61. Ahmed RP, Ivaskevicius V, Kannan M, et al. Identification of 32 novel mutations in the factor VIII gene in Indian patients with hemophilia A. *Haematologica.* 2005;90(2):283–4.
62. Murphy BC, Scriver CR, Singh SM. CpG methylation accounts for a recurrent mutation (c.1222C>T) in the human PAH gene. *Hum Mutat.* 2006;27(9):975.
63. Noyer-Weidner M, Trautner TA. Methylation of DNA in prokaryotes. *EXS.* 1993;64:39–108.
64. Ratel D, Ravanat JL, Berger F, Wion D. N6-methyladenine: the other methylated base of DNA. *Bioessays.* 2006;28(3):309–15.
65. Ehrlich M, Gama-Sosa MA, Carreira LH. DNA methylation in thermophilic bacteria: N4-methylcytosine, 5-methylcytosine, and N6-methyladenine. *Nucleic Acids Res.* 1985;13(4):1399–412.
66. Bestor TH, Bourc'his D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb Symp Quant Biol.* 2004;69:381–7.
67. Comb M, Goodman HM. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Res.* 1990;18(13):3975–82.
68. Iguchi-Ariga SM, Schaffner W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev.* 1989;3(5):612–9.
69. Campanero MR, Armstrong MI, Flemington EK. CpG methylation as a mechanism for the regulation of E2F activity. *Proc Natl Acad Sci U S A.* 2000;97(12):6481–6.

Literaturverzeichnis

70. Bednarik DP, Duckett C, Kim SU, et al. DNA CpG methylation inhibits binding of NF-kappa B proteins to the HIV-1 long terminal repeat cognate DNA motifs. *New Biol.* 1991;3(10):969–76.
71. Ben-Hattar J, Beard P, Jiricny J. Cytosine methylation in CTF and Spl recognition sites of an HSV tk promoter: effects on transcription in vivo and on factor binding in vitro. *Nucleic Acids Res.* 1989;17(24):10179–90.
72. Holler M, Westin G, Jiricny J, Schaffner W. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.* 1988;2(9):1127–35.
73. Gaston K, Fried M. CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. *Nucleic Acids Res.* 1995;23(6):901–9.
74. Tate PH, Bird AP. Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr Opin Genet Dev.* 1993;3(2):226–31.
75. Newell-Price J, Clark AJ, King P. DNA methylation and silencing of gene expression. *Trends Endocrinol Metab.* 2000;11(4):142–8.
76. Fatemi M, Wade PA. MBD family proteins: reading the epigenetic code. *J Cell Sci.* 2006;119(Pt 15):3033–7.
77. Filion GJ, Zhenilo S, Salozhin S, et al. A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol Cell Biol.* 2006;26(1):169–81.
78. Antequera F. Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci.* 2003;60(8):1647–58.
79. Armstrong, 2014
80. Mostoslavsky R, Bergman Y. DNA methylation: regulation of gene expression and role in the immune system. *Biochim Biophys Acta.* 1997;1333(1):F29–50.
81. Bestor TH. The host defence function of genomic methylation patterns. *Novartis Found Symp.* 1998;214:187–95; discussion 195–9, 228–32.
82. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* 1997;13(8):335–40.
83. Smit AF, Riggs AD. Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci U S A.* 1996;93(4):1443–8.
84. Muotri AR, Marchetto MC, Coufal NG, Gage FH. The necessary junk: new functions for transposable elements. *Hum Mol Genet.* 2007;16 Spec No 2:R159–67.
85. Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science.* 2004;303(5664):1626–32.
86. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002;16(1):6–21.
87. Panning B, Jaenisch R. RNA and the epigenetic regulation of X chromosome inactivation. *Cell.* 1998;93(3):305–8.
88. Heard E, Distechi CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. 2006;20(14):1848–67.
89. Feil R, Khosla S. Genomic imprinting in mammals: an interplay between chromatin and DNA methylation? *Trends Genet.* 1999;15(11):431–5.
90. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature.* 2007;447(7143):425–32.
91. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet.* 2002;3(6):415–28.
92. Wyatt GR, Cohen SS. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochem J.* 1953;55(5):774–82.

93. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009;324(5929):930–5.
94. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324(5929):929–30.
95. Ito S, D'Alessio AC, Taranova OV, et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010;466(7310):1129–33.
96. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation *Nature*. 2013;502(7472):472–9.
97. Globisch D, Münzel M, Müller M, et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One*. 2010;5(12):e15367.
98. Münzel M, Globisch D, Brückl T, et al. Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem Int Ed Engl*. 2010;49(31):5375–7.
99. Wen L, Li X, Yan L, et al. Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol*. 2014;15(3):R49.
100. Stroud H, Feng S, Morey Kinney S, et al. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol*. 2011;12(6):R54.
101. Choi I, Kim R, Lim HW, et al. 5-hydroxymethylcytosine represses the activity of enhancers in embryonic stem cells: a new epigenetic signature for gene regulation. *BMC Genomics*. 2014;15:670.
102. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011;333(6047):1300–3.
103. He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 2011;333(6047):1303–7.
104. Spruijt CG, Gnerlich F, Smits AH, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*. 2013;152(5):1146–59.
105. Lee KW, Pausova Z. Cigarette smoking and DNA methylation. *Front Genet*. 2013;4:132.
106. Szyf M. Dynamisches Epigenom als Vermittler zwischen Umwelt und Genom. *Medizinische Genet*. 2009;21(1):7–13.
107. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet*. 1999;23(3):314–8.
108. Guerrero-Bosagna C, Skinner MK. Environmentally induced epigenetic transgenerational inheritance of male infertility. *Curr Opin Genet Dev*. 2014;26C:79–88.
109. Alegría-Torres J, Baccarelli A, Bollati V. Epigenetics and lifestyle. *Epigenomics*. 2011;3(3):267–77.
110. Campión J, Milagro FI, Martínez JA. Individuality and epigenetics in obesity. *Obes Rev*. 2009;10(4):383–92.
111. Lomba A, Milagro FI, García-Díaz DF, et al. Obesity induced by a pair-fed high fat sucrose diet: methylation and expression pattern of genes related to energy homeostasis. *Lipids Health Dis*. 2010;9:60.
112. Seki Y, Williams L, Vuguin PM, Charron MJ. Minireview: Epigenetic programming of diabetes and obesity: animal models. *Endocrinology*. 2012;153(3):1031–8.
113. Raychaudhuri N, Raychaudhuri S, Thamocharan M, Devaskar SU. Histone code modifications repress glucose transporter 4 expression in the intrauterine growth-restricted offspring. *J Biol Chem*. 2008;283(20):13611–26.

114. Zhang S, Rattanatravay L, MacLaughlin SM, et al. Periconceptional undernutrition in normal and overweight ewes leads to increased adrenal growth and epigenetic changes in adrenal IGF2/H19 gene in offspring. *FASEB J*. 2010;24(8):2772–82.
115. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol*. 2003;23(15):5293–300.
116. Barker DJ, Gluckman PD, Godfrey KM, et al. Fetal nutrition and cardiovascular disease in adult life. *Lancet*. 1993;341(8850):938–41.
117. Blackmore HL, Ozanne SE. Maternal diet-induced obesity and offspring cardiovascular health. *J Dev Orig Health Dis*. 2013;4(5):338–47.
118. Kyle UG, Pichard C. The Dutch Famine of 1944–1945: a pathophysiological model of long-term consequences of wasting disease. *Curr Opin Clin Nutr Metab Care*. 2006;9(4):388–94.
119. Ravelli AC, van Der Meulen JH, Osmond C, et al. Obesity at the age of 50 y in men and women exposed to famine prenatally. *Am J Clin Nutr*. 1999;70(5):811–6.
120. Roseboom TJ, van der Meulen JH, Osmond C, et al. Coronary heart disease after prenatal exposure to the Dutch famine, 1944–45. *Heart*. 2000;84(6):595–8.
121. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci U S A*. 2008;105(44):17046–9.
122. Howdeshell KL, Hotchkiss AK, Thayer KA, et al. Exposure to bisphenol A advances puberty. *Nature*. 1999;401(6755):763–4.
123. Rubin BS, Murray MK, Damassa DA, et al. Perinatal exposure to low doses of bisphenol A affects body weight, patterns of estrous cyclicity, and plasma LH levels. *Environ Health Perspect*. 2001;109(7):675–80.
124. Ohlstein JF, Strong AL, McLachlan JA, et al. Bisphenol A enhances adipogenic differentiation of human adipose stromal/stem cells. *J Mol Endocrinol*. 2014;53(3):345–53.
125. Singh S, Li SS. Epigenetic effects of environmental chemicals bisphenol a and phthalates. *Int J Mol Sci*. 2012;13(8):10143–53.
126. Fagerström K. The epidemiology of smoking: health consequences and benefits of cessation. *Drugs*. 2002;62 Suppl 2:1–9.
127. [No authors listed]. Cigarette smoking and health. American Thoracic Society. *Am J Respir Crit Care Med*. 1996;153(2):861–5.
128. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*. 1999;91(14):1194–210.
129. Breitling LP, Yang R, Korn B, et al. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet*. 2011;88(4):450–7.
130. Monick M, Beach S. Coordinated changes in AHRH methylation in lymphoblasts and pulmonary macrophages from smokers. *Am J Med Genet B Neuropsychiatr Genet*. 2012;159B(2):141–51.
131. Ostrow KL, Michailidi C, Guerrero-Preston R, et al. Cigarette smoke induces methylation of the tumor suppressor gene NISCH. *Epigenetics*. 2013;8(4):383–8.
132. Shenker NS, Polidoro S, van Veldhoven K, et al. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet*. 2013;22(5):843–51.
133. Zeilinger S, Kühnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*. 2013;8(5):e63812.

Literaturverzeichnis

134. Dogan MV, Shields B, Cutrona C, et al. The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. *BMC Genomics*. 2014;15:151.
135. Breton CV, Byun HM, Wenten M, et al. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. *Am J Respir Crit Care Med*. 2009;180(5):462–7.
136. Joubert BR, Håberg SE, Nilsen RM, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2012;120(10):1425–31.
137. Breton CV, Siegmund KD, Joubert BR, et al. Prenatal tobacco smoke exposure is associated with childhood DNA CpG methylation. *PLoS One*. 2014;9(6):e99716.
138. Markunas CA, Xu Z, Harlid S, et al. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2014;122(10):1147–53.
139. Simm A, Nass N, Bartling B, et al. Potential biomarkers of ageing. *Biol Chem*. 2008;389(3):257–65.
140. Garagnani P, Bacalini MG, Pirazzini C, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell*. 2012;11(6):1132–4.
141. Rakyan VK, Down TA, Maslau S, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res*. 2010;20(4):434–9.
142. Schraml E, Grillari J. From cellular senescence to age-associated diseases: the miRNA connection. *Longev Healthspan*. 2012;1(1):10.
143. Oberdoerffer P, Michan S, McVay M. SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. *Cell*. 2008;135(5):907–18.
144. Fraga MF, Agrelo R, Esteller M. Cross-talk between aging and cancer: the epigenetic language. *Ann N Y Acad Sci*. 2007;1100:60–74.
145. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. 2005;102(30):10604–9.
146. Christensen BC, Houseman EA, Marsit CJ, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet*. 2009;5(8):e1000602.
147. Blair SN, Kohl HW 3rd, Paffenbarger RS Jr, et al. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA*. 1989;262(17):2395–401.
148. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013;49(2):359–67.
149. Hernandez DG, Nalls MA, Gibbs JR, et al. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum Mol Genet*. 2011;20(6):1164–72.
150. Alisch RS, Barwick BG, Chopra P, et al. Age-associated DNA methylation in pediatric populations. *Genome Res*. 2012;22(4):623–32.
151. Heyn H, Li N, Ferreira HJ, et al. Distinct DNA methylomes of newborns and centenarians. *Proc Natl Acad Sci U S A*. 2012;109(26):10522–7.
152. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.
153. Johansson A, Enroth S, Gyllenstein U. Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan. *PLoS One*. 2013;8(6):e67378.
154. Weidner C, Lin Q, Koch CM, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol*. 2014;15(2):R24.
155. Bocklandt S, Lin W, Sehl ME, et al. Epigenetic Predictor of Age. *PLoS One*. 2011;6(6):e14821.

156. Koch C, Wagner W. Epigenetic-aging-signature to determine age in different tissues. *Aging* (Albany NY). 2011;3(10):1018–27.
157. Yi SH, Jia YS, Mei K, et al. Age-related DNA methylation changes for forensic age-prediction. *Int J Legal Med*. 2014;[Epub ahead of print].
158. Seibt A, Knauth P, Griefahn B: Arbeitsmedizinische Leitlinie der Deutschen Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V., Nacht- und Schichtarbeit, Arbeitsmedizin – Sozialmedizin – Umweltmedizin (ASU) 2006;41(8):390–7.
159. Harrington JM. Health effects of shift work and extended hours of work. *Occup Environ Med*. 2001;58(1):68–72.
160. Unite the union. Unite guide to shift work and night work. 2013:www.uniteunion.org
161. Åkerstedt T, Wright KP Jr. Sleep loss and fatigue in shift work and shift work disorder. *Sleep Med Clin*. 2009;4(2):257–271.
162. Erren TC, Falaturi P, Morfeld P, et al. Shift work and cancer: the evidence and the challenge. *Dtsch Arztebl Int*. 2010;107(38):657–62.
163. Knutsson A, Alfredsson L, Karlsson B, et al. Breast cancer among shift workers: results of the WOLF longitudinal cohort study. *Scand J Work Environ Health*. 2012;39(2):170–7.
164. Canuto R, Garcez AS, Olinto MT. Metabolic syndrome and shift work: A systematic review. *Sleep Med Rev*. 2013;17(6):425–31.
165. Caruso CC. Negative impacts of shiftwork and long work hours. *Rehabil Nurs*. 2014;39(1):16–25.
166. Bechtold DA, Gibbs JE, Loudon AS. Circadian dysfunction in disease. *Trends Pharmacol Sci*. 2010;31(5):191–8.
167. Golombek DA, Casiraghi LP, Agostino PV, et al. The times they're a-changing: effects of circadian desynchronization on physiology and disease. *J Physiol Paris*. 2013;107(4):310–22.
168. Gossan N, Boot-Handford R, Meng QJ. Ageing and osteoarthritis: a circadian rhythm connection. *Biogerontology*. 2014;[Epub ahead of print].
169. Oike H, Oishi K, Kobori M. Nutrients, Clock Genes, and Chrononutrition. *Curr Nutr Rep*. 2014;3:204–212.
170. Aguilar-Arnal L, Sassone-Corsi P. The circadian epigenome: how metabolism talks to chromatin remodeling. *Curr Opin Cell Biol*. 2013;25(2):170–6.
171. Joska TM, Zaman R, Belden WJ. Regulated DNA methylation and the circadian clock: implications in cancer. *Biology (Basel)*. 2014;3(3):560–77.
172. Orozco-Solis R, Sassone-Corsi P. Epigenetic control and the circadian clock: linking metabolism to neuronal responses. *Neuroscience*. 2014;264:76–87.
173. Feng D, Liu T, Sun Z, et al. A circadian rhythm orchestrated by histone deacetylase 3 controls hepatic lipid metabolism. *Science*. 2011;331(6022):1315–9.
174. Moore DD. Physiology. Crise de foie, redux? *Science*. 2011;331(6022):1275–6.
175. Shi F, Chen X, Fu A, et al. Aberrant DNA methylation of miR-219 promoter in long-term night shiftworkers. *Environ Mol Mutagen*. 2013;54(6):406–13.
176. Zhu Y, Stevens RG, Hoffman AE, et al. Epigenetic impact of long-term shiftwork: pilot evidence from circadian genes and whole-genome methylation analysis. *Chronobiol Int*. 2011;28(10):852–61.
177. Jacobs DI, Hansen J, Fu A, et al. Methylation alterations at imprinted genes detected among long-term shiftworkers. *Environ Mol Mutagen*. 2013;54(2):141–6.
178. Bollati V, Baccarelli A, Sartori S, et al. Epigenetic effects of shiftwork on blood DNA methylation. *Chronobiol Int*. 2010;27(5):1093–104.

179. Hogenson TL. Epigenetics as the Underlying Mechanism for Monozygotic Twin Discordance. *Medical Epigenetics*. 2013;1(1):3–18.
180. Bailey A, Le Couteur A, Gottesman I, et al. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med*. 1995;25(1):63–77.
181. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance - a population-based twin study. *Diabetologia*. 1999;42(2):139–45.
182. Kaminsky ZA, Tang T, Wang SC, et al. DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet*. 2009;41(2):240–5.
183. Poulsen P, Esteller M, Vaag A, Fraga MF. The epigenetic basis of twin discordance in age-related diseases. *Pediatr Res*. 2007;61(5 Pt 2):38R–42R.
184. Tan Q, Christiansen L, Thomassen M, et al. Twins for epigenetic studies of human aging and development. *Ageing Res Rev*. 2013;12(1):182–7.
185. Roessler J, Ammerpohl O, Gutwein J, et al. Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. *BMC Res Notes*. 2012;5:210.
186. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*. 1988;16(3):1215.
187. Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*. 1994;22(15):2990–7.
188. Krueger F, Kreck B, Franke A, Andrews SR. DNA methylome analysis using short bisulfite sequencing data. *Nat Methods*. 2012;9(2):145–51.
189. Illumina, Inc. Infinium® HumanMethylation450 BeadChip. Data Sheet: Epigenetics. 2012; Pub. No. 270-2010-001.
190. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*. 2012;13(6):R44.
191. Assenov Y, Müller F, Lutsik P, et al. Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods*. 2014;11(11):1138–1140.
192. Triche TJ Jr, Weisenberger DJ, Van Den Berg D, et al. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res*. 2013;41(7):e90.
193. Dedeurwaerder S, Defrance M, Calonne E, et al. Evaluation of the Infinium Methylation 450K technology. *Epigenomics*. 2011;3(6):771–84.
194. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J Roy Statist Soc Ser B (Methodological)*. 1995;57:289–300.
195. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GENE SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res*. 2013;41(Web Server issue):W77–83.
196. Mitchell PJ, Tjian R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*. 1989;245(4916):371–8.
197. He HH, Meyer CA, Chen MW, et al. Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. *Genome Res*. 2012;22(6):1015–25.
198. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. 2011;12(8):529–41.
199. Petrosino J, Highlander S, Luna RA, et al. Metagenomic pyrosequencing and microbial identification. *Clin Chem*. 2009;55(5):856–66.
200. Van der Auwera I, Yu W, Suo L, et al. Array-based DNA methylation profiling for breast cancer subtype discrimination. *PLoS One*. 2010;5(9):e12616.

Literaturverzeichnis

201. Ammerpohl O, Pratschke J, Schafmayer C, et al. Distinct DNA methylation patterns in cirrhotic liver and hepatocellular carcinoma. *Int J Cancer*. 2012;130(6):1319–28.
202. Sun YV, Turner ST, Smith JA, et al. Comparison of the DNA methylation profiles of human peripheral blood cells and transformed B-lymphocytes. *Hum Genet*. 2010;127(6):651–8.
203. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One*. 2012;7(7):e41361.
204. MacKinney AA Jr. Effect of aging on the peripheral blood lymphocyte count. *J Gerontol*. 1978;33(2):213–6.
205. Mahlkecht U, Kaiser S. Age-related changes in peripheral blood counts in humans. *Exp Ther Med*. 2010;1(6):1019–1025.
206. Jourquin J, Duncan D, Shi Z, Zhang B. GLAD4U: deriving and prioritizing gene lists from PubMed literature. *BMC Genomics*. 2012;13 Suppl 8:S20.
207. Ballas N, Mandel G. The many faces of REST oversee epigenetic programming of neuronal genes. *Curr Opin Neurobiol*. 2005;15(5):500–6.
208. Ohtaka-Maruyama C, Hirai S, Miwa A, et al. RP58 regulates the multipolar-bipolar transition of newborn neurons in the developing cerebral cortex. *Cell Rep*. 2013;3(2):458–71.
209. Davis SW, Castinetti F, Carvalho LR, et al. Molecular mechanisms of pituitary organogenesis: In search of novel regulatory genes. *Mol Cell Endocrinol*. 2010;323(1):4–19.
210. Lee S, Cuvillier JM, Lee B, et al. Fusion protein Isl1-Lhx3 specifies motor neuron fate by inducing motor neuron genes and concomitantly suppressing the interneuron programs. *Proc Natl Acad Sci U S A*. 2012;109(9):3383–8.
211. Mollard P, Hodson DJ, Lafont C, et al. A tridimensional view of pituitary development and function. *Trends Endocrinol Metab*. 2012;23(6):261–9.
212. Takagi M, Ishii T, Inokuchi M, et al. Gradual loss of ACTH due to a novel mutation in LHX4: comprehensive mutation screening in Japanese patients with congenital hypopituitarism. *PLoS One*. 2012;7(9):e46008.
213. Henriksson R, Bäckman CM, Harvey BK, et al. PDYN, a gene implicated in brain/mental disorders, is targeted by REST in the adult human brain. *Biochim Biophys Acta*. 2014 Nov;1839(11):1226–32.
214. Richter J, Ammerpohl O, Martín-Subero JI, et al. Array-based DNA methylation profiling of primary lymphomas of the central nervous system. *BMC Cancer*. 2009;9:455.
215. Friemel C, Ammerpohl O, Gutwein J, et al. Array-based DNA methylation profiling in male infertility reveals allele-specific DNA methylation in PIWIL1 and PIWIL2. *Fertil Steril*. 2014;101(4):1097–1103.e1.
216. LaNoue KF, Martin LF. Abnormal A1 adenosine receptor function in genetic obesity. *FASEB J*. 1994;8(1):72–80.
217. Dhalla AK, Chisholm JW, Reaven GM, Belardinelli L. A1 adenosine receptor: role in diabetes and obesity. *Handb Exp Pharmacol*. 2009;(193):271–95.
218. Traurig MT, Orczewska JI, Ortiz DJ, et al. Evidence for a Role of LPGAT1 in Influencing BMI and Percent Body Fat in Native Americans. *Obesity (Silver Spring)*. 2012;21(1):193–202.
219. Kannel WB, D'Agostino RB, Belanger AJ. Fibrinogen, cigarette smoking, and risk of cardiovascular disease: insights from the Framingham Study. *Am Heart J*. 1987;113(4):1006–10.
220. Grines CL, Topol EJ, O'Neill WW, et al. Effect of cigarette smoking on outcome after thrombolytic therapy for myocardial infarction. *Circulation*. 1995;91(2):298–303.

221. Weir MR, Dzau VJ. The renin-angiotensin-aldosterone system: a specific target for hypertension management. *Am J Hypertens.* 1999;12(12 Pt 3):205S–213S.
222. Omvik P. How smoking affects blood pressure. *Blood Press.* 1996;5(2):71–7.
223. Bartal M. Health effects of tobacco use and exposure. *Monaldi Arch Chest Dis.* 2001;56(6):545–54.
224. Rajaratnam SM, Howard ME, Grunstein RR. Sleep loss and circadian disruption in shift work: health burden and management. 2013;199(8):S11–5.
225. Bøggild H, Knutsson A. Shift work, risk factors and cardiovascular disease. *Scand J Work Environ Health.* 1999;25(2):85–99.
226. Nojkov B, Rubenstein JH, Chey WD, Hoogerwerf WA. The impact of rotating shift work on the prevalence of irritable bowel syndrome in nurses. *Am J Gastroenterol.* 2010;105(4):842–7.
227. Konturek PC, Brzozowski T, Konturek SJ. Gut clock: implication of circadian rhythms in the gastrointestinal tract. *J Physiol Pharmacol.* 2011;62(2):139–150.
228. Sancini A, Ciarrocca M, Capozzella A, et al. [Shift and night work and mental health]. *G Ital Med Lav Ergon.* 2012;34(1):76–84.
229. Morrissette DA. Twisting the night away: a review of the neurobiology, genetics, diagnosis, and treatment of shift work disorder. *CNS Spectr.* 2013;18 Suppl 1:45–53; quiz 54.
230. Vanyushin BF, Tkacheva SG, Belozersky AN. Rare bases in animal DNA. *Nature.* 1970;225(5236):948–9.
231. Vanyushin BF, Mazin AL, Vasilyev VK, Belozersky AN. The content of 5-methylcytosine in animal DNA: the species and tissue specificity. *Biochim Biophys Acta.* 1973;299(3):397–403.
232. Wilson V, Smith R, Ma S, Cutler R. Genomic 5-methyldeoxycytidine decreases with age. *J Biol Chem.* 1987;262(21):9948–51.
233. Bollati V, Schwartz J, Wright R, et al. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. *Mech Ageing Dev.* 2009;130(4):234–9.
234. Jintaridh P, Mutirangura A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. *Physiol Genomics.* 2010;41(2):194–200.
235. Jintaridh P, Tungtrongchitr R, Preutthipan S, Mutirangura A. Hypomethylation of Alu elements in post-menopausal women with osteoporosis. *PLoS One.* 2013;8(8):e70386.
236. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet.* 2012;13(7):484–92.
237. Kreck B, Richter J, Ammerpohl O, et al. Base-pair resolution DNA methylome of the EBV-positive Endemic Burkitt lymphoma cell line DAUDI determined by SOLiD bisulfite-sequencing. *Leukemia.* 2013;27(8):1751–3.
238. Florath I, Butterbach K, Müller H, et al. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Hum Mol Genet.* 2014;23(5):1186–201.
239. Brenet F, Moh M, Funk P, et al. DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One.* 2011;6(1):e14524.
240. Bell JT, Tsai PC, Yang TP, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* 2012;8(4):e1002629.
241. Ranganathan S, Noyes NC, Migliorini M, et al. LRAD3, a novel low-density lipoprotein receptor family member that modulates amyloid precursor protein trafficking. *J Neurosci.* 2011;31(30):10836–46.

242. Suzu S, Hayashi Y, Harumi T, et al. Molecular cloning of a novel immunoglobulin superfamily gene preferentially expressed by brain and testis. *Biochem Biophys Res Commun.* 2002;296(5):1215–21.
243. Katoh M, Katoh M. IGSF11 gene, frequently up-regulated in intestinal-type gastric cancer, encodes adhesion molecule homologous to CXADR, FLJ22415 and ESAM. *Int J Oncol.* 2003;23(2):525–31.
244. Watanabe T, Suda T, Tsunoda T, et al. Identification of immunoglobulin superfamily 11 (IGSF11) as a novel target for cancer immunotherapy of gastrointestinal and hepatocellular carcinomas. *Cancer Sci.* 2005;96(8):498–506.
245. Henderson LJ, Coe BP, Lee EH, et al. Genomic and gene expression profiling of minute alterations of chromosome arm 1p in small-cell lung carcinoma cells. *Br J Cancer.* 2005;92(8):1553–60.
246. Bavarva JH, Tae H, Settlage RE, Garner HR. Characterizing the Genetic Basis for Nicotine Induced Cancer Development: A Transcriptome Sequencing Study. *PLoS One.* 2013;8(6):e67252.
247. Bensaïd M, Hmani-Aifa M, Hammami B, et al. DFNB66 and DFNB67 loci are non allelic and rarely contribute to autosomal recessive nonsyndromic hearing loss. *Eur J Med Genet.* 2011;54(6):e565–9.
248. Wang SS, Smiraglia DJ, Wu YZ, et al. Identification of novel methylation markers in cervical cancer using restriction landmark genomic scanning. *Cancer Res.* 2008;68(7):2489–97.
249. Rajab A, Kelberman D, de Castro SC, et al. Novel mutations in LHX3 are associated with hypopituitarism and sensorineural hearing loss. *Hum Mol Genet.* 2008;17(14):2150–9.
250. Kriström B, Zdunek AM, Rydh A, et al. A novel mutation in the LIM homeobox 3 gene is responsible for combined pituitary hormone deficiency, hearing impairment, and vertebral malformations. *J Clin Endocrinol Metab.* 2009;94(4):1154–61.
251. Giamas G, Filipović A, Jacob J, et al. Kinome screening for regulators of the estrogen receptor identifies LMTK3 as a new therapeutic target in breast cancer. *Nat Med.* 2011;17(6):715–9.
252. Stebbing J, Filipovic A, Ellis IO, et al. LMTK3 expression in breast cancer: association with tumor phenotype and clinical outcome. *Breast Cancer Res Treat.* 2012;132(2):537–44.
253. Xu Z, Qi X, Zhang X, Yu L. Preoperative serum LMTK3 as a novel biomarker in non-small cell lung cancer. *Tumour Biol.* 2014;35(5):5007–11.
254. Li Z, Wu J, Ji M, et al. Prognostic role of lemur tyrosine kinase 3 in postoperative gastric cancer. *Mol Clin Oncol.* 2014;2(5):756–760.
255. Shi H, Li Q, Ji M, et al. Lemur tyrosine kinase-3 is a significant prognostic marker for patients with colorectal cancer. *Int J Clin Exp Pathol.* 2014;7(3):1101–7.
256. Han A, Pan F, Stroud JC, et al. Sequence-specific recruitment of transcriptional co-repressor Cabin1 by myocyte enhancer factor-2. *Nature.* 2003;422(6933):730–4.
257. Zhang CL, McKinsey TA, Chang S, et al. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell.* 2002;110(4):479–88.
258. Satoh J, Kawana N, Yamamoto Y. Pathway Analysis of ChIP-Seq-Based NRF1 Target Genes Suggests a Logical Hypothesis of their Involvement in the Pathogenesis of Neurodegenerative Diseases. *Gene Regul Syst Bio.* 2013;7:139–52.
259. Faryna M, Konermann C, Aulmann S, et al. Genome-wide methylation screen in low-grade breast cancer identifies novel epigenetically altered genes as potential biomarkers for tumor diagnosis. *FASEB J.* 2012;26(12):4937–50.
260. Hong SM, Omura N, Vincent A, et al. Genome-wide CpG island profiling of intraductal papillary mucinous neoplasms of the pancreas. *Clin Cancer Res.* 2012;18(3):700–12.

261. Hu Y, Xing J, Wang L, et al. RGS22, a novel cancer/testis antigen, inhibits epithelial cell invasion and metastasis. *Clin Exp Metastasis*. 2011;28(6):541–9.
262. Richardson B. Impact of aging on DNA methylation. *Ageing Res Rev*. 2003;2(3):245–61.
263. Barodka VM, Joshi BL, Berkowitz DE, et al. Review article: implications of vascular aging. *Anesth Analg*. 2011;112(5):1048–60.
264. Ahmed A. Myocardial beta-1 adrenoceptor down-regulation in aging and heart failure: implications for beta-blocker use in older adults with heart failure. *Eur J Heart Fail*. 2003;5(6):709–15.
265. Vågnes ØB, Hansen FH, Christiansen REF, et al. Age-dependent regulation of vasopressin V 1a receptors in preglomerular vessels from the spontaneously hypertensive rat. *Am J Physiol Renal Physiol*. 2004;286(5):F997–1003.
266. Li H, Ma XQ, Ye F, et al. Expressions of cardiac sympathetic norepinephrine transporter and beta1-adrenergic receptor decreased in aged rats. *J Zhejiang Univ Sci B*. 2009;10(3):203–10.
267. Rice RA, Berchtold NC, Cotman CW, Green KN. Age-related downregulation of the CaV3.1 T-type calcium channel as a mediator of amyloid beta production. *Neurobiol Aging*. 2014;35(5):1002–11.
268. Lieber SC, Qiu H, Chen L, et al. Cardiac dysfunction in aging conscious rats: altered cardiac cytoskeletal proteins as a potential mechanism. *Am J Physiol Heart Circ Physiol*. 2008;295(2):H860–6.
269. Russ DW, Grandy JS. Increased desmin expression in hindlimb muscles of aging rats. *J Cachexia Sarcopenia Muscle*. 2011;2(3):175–180.
270. Li X, Chan TO, Myers V, et al. Controlled and cardiac-restricted overexpression of the arginine vasopressin V1A receptor causes reversible left ventricular dysfunction through Gαq-mediated cell signaling. *Circulation*. 2011;124(5):572–81.
271. Dorfman LJ, Bosley TM. Age-related changes in peripheral and central nerve conduction in man. *Neurology*. 1979;29(1):38–44.
272. Deary IJ, Corley J, Gow AJ, et al. Age-associated cognitive decline. *Br Med Bull*. 2009;92:135–52.
273. Weiskopf D, Weinberger B, Grubeck-Loebenstien B. The aging of the immune system. *Transpl Int*. 2009;22(11):1041–50.
274. Adalsteinsson BT, Gudnason H, Aspelund T, et al. Heterogeneity in white blood cells has potential to confound DNA methylation measurements. *PLoS One*. 2012;7(10):e46705.
275. Talens RP, Boomsma DI, Tobi EW, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J*. 2010;24(9):3135–44.
276. Lavebratt C, Almgren M, Ekström TJ. Epigenetic regulation in obesity. *Int J Obes (Lond)*. 2012;36(6):757–65.
277. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 2014;383(9933):1990–8.
278. Renehan AG, Tyson M, Egger M, et al. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008;371(9612):569–78.
279. In-lw S, Biro FM. Adolescent women and obesity. *J Pediatr Adolesc Gynecol*. 2011;24(2):58–61.
280. Canoy D, Cairns BJ, Balkwill A, et al. Coronary heart disease incidence in women by waist circumference within categories of body mass index. *Eur J Prev Cardiol*. 2013;20(5):759–62.
281. Yokomori N, Tawata M, Onaya T. DNA demethylation during the differentiation of 3T3-L1 cells affects the expression of the mouse GLUT4 gene. *Diabetes*. 1999;48(4):685–90.

282. Fujiki K, Kano F, Shiota K, Murata M. Expression of the peroxisome proliferator activated receptor gamma gene is repressed by DNA methylation in visceral adipose tissue of mouse models of diabetes. *BMC Biol.* 2009;7:38.
283. Kuroda A, Rauch TA, Todorov I, et al. Insulin gene expression is regulated by DNA methylation. *PLoS One.* 2009;4(9):e6953.
284. Feinberg AP, Irizarry RA. Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc Natl Acad Sci U S A.* 2010;107 Suppl 1:1757–64.
285. Keller M, Kralisch S, Rohde K, et al. Global DNA methylation levels in human adipose tissue are related to fat distribution and glucose homeostasis. *Diabetologia.* 2014;57(11):2374–83.
286. Almén MS, Jacobsson JA, Moschonis G, et al. Genome wide analysis reveals association of a FTO gene variant with epigenetic changes. *Genomics.* 2012;99(3):132–7.
287. Wang X, Zhu H, Snieder H, et al. Obesity related methylation changes in DNA of peripheral blood leukocytes. *BMC Med.* 2010;8:87.
288. Xu X, Su S, Barnes VA, et al. A genome-wide methylation study on obesity. differential variability and differential methylation. *Epigenetics.* 2013;8(5):522–33.
289. Velloso LA, Torsoni MA, Araujo EP. Hypothalamic dysfunction in obesity. *Rev Neurosci.* 2009;20(5–6):441–9.
290. Dodt C, Lönnroth P, Wellhöner JP, et al. Sympathetic control of white adipose tissue in lean and obese humans. *Acta Physiol Scand.* 2003;177(3):351–7.
291. Baum P, Petroff D, Classen J, et al. Dysfunction of autonomic nervous system in childhood obesity: a cross-sectional study. *PLoS One.* 2013;8(1):e54546.
292. Milagro FI, Campión J, García-Díaz DF, et al. High fat diet-induced obesity modifies the methylation pattern of leptin promoter in rats. *J Physiol Biochem.* 2009;65(1):1–9.
293. Campión J, Milagro FI, Goyenechea E, Martínez JA. TNF-alpha promoter methylation as a predictive biomarker for weight-loss response. *Obesity (Silver Spring).* 2009;17(6):1293–7.
294. Brøns C, Jacobsen S, Nilsson E, et al. Deoxyribonucleic acid methylation and gene expression of PPARGC1A in human muscle is influenced by high-fat overfeeding in a birth-weight-dependent manner. *J Clin Endocrinol Metab.* 2010;95(6):3048–56.
295. Saggerson ED, Jamal Z. Differences in the properties of A1-type adenosine receptors in rat white and brown adipocytes. *Biochem J.* 1990;269(1):157–61.
296. Reppert SM, Weaver DR, Stehle JH, Rivkees SA. Molecular cloning and characterization of a rat A1-adenosine receptor that is widely expressed in brain and spinal cord. *Mol Endocrinol.* 1991;5(8):1037–48.
297. Yang M, Chu R, Chisholm JW, et al. Adenosine A₁ receptors do not play a major role in the regulation of lipogenic gene expression in hepatocytes. *Eur J Pharmacol.* 2012;683(1–3):332–9.
298. Schoelch C, Kuhlmann J, Gossel M, et al. Characterization of adenosine-A1 receptor-mediated antilipolysis in rats by tissue microdialysis, 1H-spectroscopy, and glucose clamp studies. *Diabetes.* 2004;53(7):1920–6.
299. Malumbres M, Pérez de Castro I, Santos J, et al. Hypermethylation of the cell cycle inhibitor p15INK4b 3'-untranslated region interferes with its transcriptional regulation in primary lymphomas. *Oncogene.* 1999;18(2):385–96.
300. Smith JF, Mahmood S, Song F, et al. Identification of DNA methylation in 3' genomic regions that are associated with upregulation of gene expression in colorectal cancer. *Epigenetics.* 2007;2(3):161–72.

Literaturverzeichnis

301. Mayol G, Martín-Subero JI, Ríos J, et al. DNA hypomethylation affects cancer-related biological functions and genes relevant in neuroblastoma pathogenesis. *PLoS One*. 2012;7(11):e48401.
302. Maussion G, Yang J, Suderman M, et al. Functional DNA methylation in a transcript specific 3'UTR region of TrkB associates with suicide. *Epigenetics*. 2014;9(8):1061–70.
303. Engfeldt P, Arner P. Lipolysis in human adipocytes, effects of cell size, age and of regional differences. *Horm Metab Res Suppl*. 1988;19:26–9.
304. Arner P. Human fat cell lipolysis: biochemistry, regulation and clinical role. *Best Pract Res Clin Endocrinol Metab*. 2005;19(4):471–82.
305. Gharibi B, Abraham AA, Ham J, Evans BA. Contrasting effects of A1 and A2b adenosine receptors on adipogenesis. *Int J Obes (Lond)*. 2012;36(3):397–406.
306. Eisenstein A, Ravid K. G protein-coupled receptors and adipogenesis: a focus on adenosine receptors. *J Cell Physiol*. 2014;229(4):414–21.
307. Shi Y, Cheng D. Beyond triglyceride synthesis: the dynamic functional roles of MGAT and DGAT enzymes in energy metabolism. *Am J Physiol Endocrinol Metab*. 2009;297(1):E10–8.
308. Turkish AR, Sturley SL. The genetics of neutral lipid biosynthesis: an evolutionary perspective. *Am J Physiol Endocrinol Metab*. 2009;297(1):E19–27.
309. Hiramane Y, Emoto H, Takasuga S, Hiramatsu R. Novel acyl-coenzyme A:monoacylglycerol acyltransferase plays an important role in hepatic triacylglycerol secretion. *J Lipid Res*. 2010;51(6):1424–31.
310. Comuzzie AG, Cole SA, Laston SL, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One*. 2012;7(12):e51954.
311. Greengard P, Valtorta F, Czernik AJ, Benfenati F. Synaptic Vesicle Phosphoproteins and Regulation of Synaptic Function. *Science*. 1993;259(5096):780–5.
312. Li L, Chin LS, Greengard P, et al. Localization of the synapsin II (SYN2) gene to human chromosome 3 and mouse chromosome 6. *Genomics*. 1995;28(2):365–6.
313. Sharma NK, Langberg KA, Mondal AK, et al. Type 2 diabetes (T2D) associated polymorphisms regulate expression of adjacent transcripts in transformed lymphocytes, adipose, and muscle from Caucasian and African-American subjects. *J Clin Endocrinol Metab*. 2011;96(2):E394–403.
314. Kuppaswamy U, Ananthasubramanian S, Wang Y, et al. Predicting gene ontology annotations of orphan GWAS genes using protein-protein interactions. *Algorithms Mol Biol*. 2014;9(1):10.
315. Andersson EA, Pilgaard K, Pisinger C, et al. Type 2 diabetes risk alleles near ADCY5, CDKAL1 and HHEX-IDE are associated with reduced birthweight. *Diabetologia*. 2010;53(9):1908–16.
316. Freathy RM, Mook-Kanamori DO, Sovio U, et al. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nat Genet*. 2010;42(5):430–5.
317. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42(2):105–16.
318. Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 1985;28(7):412–9.
319. Zhao J, Goldberg J, Bremner JD, Vaccarino V. Global DNA methylation is associated with insulin resistance: a monozygotic twin study. *Diabetes*. 2012;61(2):542–6.
320. Kim JW, Oh MM, Yoon CY, et al. The effect of diet-induced insulin resistance on DNA methylation of the androgen receptor promoter in the penile cavernosal smooth muscle of mice. *Asian J Androl*. 2013;15(4):487–91.

Literaturverzeichnis

321. Ting W, Yanyan Q, Jian H, et al. The relationship between insulin resistance and CpG island methylation of LMNA gene in polycystic ovary syndrome. *Cell Biochem Biophys.* 2013;67(3):1041–7.
322. Lokk K, Modhukur V, Rajashekar B, et al. DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol.* 2014;15(4):r54.
323. Ahrens M, Ammerpohl O, von Schönfels W, et al. DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab.* 2013;18(2):296–302.
324. Claussen M, Suter B. BicD-dependent localization processes: from *Drosophila* development to human cell biology. *Ann Anat.* 2005;187(5–6):539–53.
325. Mangino M, Brouillette S, Braund P, et al. A regulatory SNP of the BICD1 gene contributes to telomere length variation in humans. *Hum Mol Genet.* 2008;17(16):2518–23.
326. Slagboom PE, Droog S, Boomsma DI. Genetic determination of telomere size in humans: a twin study of three age groups. *Am J Hum Genet.* 1994;55(5):876–82.
327. Ju Z, Rudolph KL. Telomeres and telomerase in cancer stem cells. *Eur J Cancer.* 2006;42(9):1197–203.
328. Epel ES, Lin J, Wilhelm FH, et al. Cell aging in relation to stress arousal and cardiovascular disease risk factors. *Psychoneuroendocrinology.* 2006;31(3):277–87.
329. Fitzpatrick AL, Kronmal RA, Gardner JP, et al. Leukocyte telomere length and cardiovascular disease in the cardiovascular health study. *Am J Epidemiol.* 2007;165(1):14–21.
330. Tsuji T, Aoshiba K, Nagai A. Alveolar cell senescence in patients with pulmonary emphysema. *Am J Respir Crit Care Med.* 2006;174(8):886–93.
331. Adaikalakoteswari A, Balasubramanyam M, Ravikumar R, et al. Association of telomere shortening with impaired glucose tolerance and diabetic macroangiopathy. *Atherosclerosis.* 2007;195(1):83–9.
332. Cui Y, Gao YT, Cai Q, et al. Associations of leukocyte telomere length with body anthropometric indices and weight change in Chinese women. *Obesity (Silver Spring).* 2013;21(12):2582–8.
333. Formichi C, Cantara S, Ciuoli C, et al. Weight loss associated with bariatric surgery does not restore short telomere length of severe obese patients after 1 year. 2014:2089–93.
334. Valdes AM, Andrew T, Gardner JP, et al. Obesity, cigarette smoking, and telomere length in women. *Lancet.* 2005;366(9486):662–4.
335. Tomita K, Caramori G, Ito K, et al. Telomere Shortening in Alveolar Macrophages of Smokers and COPD Patients. *The Open Pathology Journal.* 2010;4: 23–29.
336. Kong X, Cho MH, Anderson W, et al. Genome-wide association study identifies BICD1 as a susceptibility gene for emphysema. *Am J Respir Crit Care Med.* 2011;183(1):43–9.
337. Maritz GS, Mutemwa M. Tobacco smoking: patterns, health consequences for adults, and the long-term health of the offspring. *Glob J Health Sci.* 2012;4(4):62–75.
338. Rooke SE, Norberg MM, Copeland J, Swift W. Health outcomes associated with long-term regular cannabis and tobacco smoking. *Addict Behav.* 2013;38(6):2207–13.
339. Macha MA, Matta A, Sriram U, et al. Clinical significance of TC21 overexpression in oral cancer. *J Oral Pathol Med.* 2010;39(6):477–85.
340. Moreno-López LA, Esparza-Gómez GC, González-Navarro A, et al. Risk of oral cancer associated with tobacco smoking, alcohol consumption and oral hygiene: a case-control study in Madrid, Spain. *Oral Oncol.* 2000;36(2):170–4.

341. Berna-Erro A, Braun A, Kraft R, et al. STIM2 regulates capacitive Ca²⁺ entry in neurons and plays a key role in hypoxic neuronal cell death. *Sci Signal*. 2009;2(93):ra67.
342. Aytes A, Molleví DG, Martínez-Iniesta M, et al. Stromal interaction molecule 2 (STIM2) is frequently overexpressed in colorectal tumors and confers a tumor cell growth suppressor phenotype. *Mol Carcinog*. 2012;51(9):746–53.
343. Li Z, Lai G, Deng L, et al. Association of SLC38A4 and system A with abnormal fetal birth weight. *Exp Ther Med*. 2012;3(2):309–313.
344. Januchowski R, Zawierucha P, Ruciński M, et al. Drug transporter expression profiling in chemoresistant variants of the A2780 ovarian cancer cell line. *Biomed Pharmacother*. 2014;68(4):447–53.
345. Zhang J, Kamdar O, Le W, et al. Nicotine induces resistance to chemotherapy by modulating mitochondrial signaling in lung cancer. *Am J Respir Cell Mol Biol*. 2009;40(2):135–46.
346. Videtic GM, Stitt LW, Dar AR, et al. Continued cigarette smoking by patients receiving concurrent chemoradiotherapy for limited-stage small-cell lung cancer is associated with decreased survival. *J Clin Oncol*. 2003;21(8):1544–9.
347. Tsao AS, Liu D, Lee JJ, et al. Smoking affects treatment outcome in patients with advanced nonsmall cell lung cancer. *Cancer*. 2006;106(11):2428–36.
348. Trexler M, Bányai L, Patthy L. Distinct expression pattern of two related human proteins containing multiple types of protease-inhibitory modules. *Biol Chem*. 2002;383(1):223–8.
349. Hinman LM, Stevens CA, Matthay RA, Gee JB. Elastase and lysozyme activities in human alveolar macrophages. Effects of cigarette smoking. *Am Rev Respir Dis*. 1980;121(2):263–71.
350. Bridges RB, Fu MC, Rehm SR. Increased neutrophil myeloperoxidase activity associated with cigarette smoking. *Eur J Respir Dis*. 1985;67(2):84–93.
351. Numanami H, Koyama S, Nelson DK, et al. Serine protease inhibitors modulate smoke-induced chemokine release from human lung fibroblasts. *Am J Respir Cell Mol Biol*. 2003;29(5):613–9.
352. Tetley TD, Smith SF, Burton GH, et al. Effects of cigarette smoking and drugs on respiratory tract proteases and antiproteases. *Eur J Respir Dis Suppl*. 1987;153:93–102.
353. Bridges RB, Wyatt RJ, Rehm SR. Effect of smoking on peripheral blood leukocytes and serum antiproteases. *Eur J Respir Dis Suppl*. 1985;139:24–33.
354. Shapiro SD. Evolving concepts in the pathogenesis of chronic obstructive pulmonary disease. *Clin Chest Med*. 2000;21(4):621–32.
355. Yao I, Iida J, Nishimura W, Hata Y. Synaptic and nuclear localization of brain-enriched guanylate kinase-associated protein. *J Neurosci*. 2002;22(13):5354–64.
356. Nagy C, Suderman M, Yang J, et al. Astrocytic abnormalities and global DNA methylation patterns in depression and suicide. *Mol Psychiatry*. 2014.
357. Dierker L, Rose J, Selya A, et al. Depression and nicotine dependence from adolescence to young adulthood. *Addict Behav*. 2014;41C:124–128.
358. Sanchez-Villegas A, Schlatter J, Ortuno F, et al. Validity of a self-reported diagnosis of depression among participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-I). *BMC Psychiatry*. 2008;8:43.
359. Taylor G, McNeill A, Girling A, et al. Change in mental health after smoking cessation: systematic review and meta-analysis. *BMJ*. 2014;348:g1151.
360. Puig-Vilanova E, Aguiló R, Rodríguez-Fuster A, et al. Epigenetic mechanisms in respiratory muscle dysfunction of patients with chronic obstructive pulmonary disease. *PLoS One*. 2014;9(11):e111514.

361. Cohen AL, Piccolo SR, Cheng L, et al. Genomic pathway analysis reveals that EZH2 and HDAC4 represent mutually exclusive epigenetic pathways across human cancers. *BMC Med Genomics*. 2013;6:35.
362. Bakhru A, Erlinger TP. Smoking cessation and cardiovascular disease risk factors: results from the Third National Health and Nutrition Examination Survey. *PLoS Med*. 2005;2(6):e160.
363. Soyombo AA, Hofmann SL. Molecular cloning and expression of palmitoyl-protein thioesterase 2 (PPT2), a homolog of lysosomal palmitoyl-protein thioesterase with a distinct substrate specificity. *J Biol Chem*. 1997;272(43):27456–63.
364. Hancock DB, Eijgelsheim M, Wilk JB. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet*. 2010;42(1):45–52.
365. Castaldi PJ, Cho MH, Litonjua AA, et al. The association of genome-wide significant spirometric loci with chronic obstructive pulmonary disease susceptibility. *Am J Respir Cell Mol Biol*. 2011;45(6):1147–53.
366. Kim WJ, Lee MK, Shin C, et al. Genome-wide association studies identify locus on 6p21 influencing lung function in the Korean population. *Respirology*. 2014;19(3):360–8.
367. Delgado G, Siekmeier R, Grammer TB, et al. Alterations in the Coagulation System of Active Smokers from the Ludwigshafen Risk and Cardiovascular Health (LURIC) Study. *Adv Exp Med Biol*. 2015;832:9–14.
368. Bartecchi CE, MacKenzie TD, Schrier RW. The human costs of tobacco use (1). *N Engl J Med*. 1994;330(13):907–12.
369. Ilgenli TF, Akpınar O. Acute effects of smoking on right ventricular function. A tissue Doppler imaging study on healthy subjects. *Swiss Med Wkly*. 2007;137(5–6):91–6.
370. Barutcu I, Esen AM, Kaya D, et al. Effect of acute cigarette smoking on left and right ventricle filling parameters: a conventional and tissue Doppler echocardiographic study in healthy participants. *Angiology*. 2008;59(3):312–6.
371. Giacomini E, Palmerini E, Ballo P, et al. Acute effects of caffeine and cigarette smoking on ventricular long-axis function in healthy subjects. *Cardiovasc Ultrasound*. 2008;6:9.
372. Kasikcioglu E, Elitok A, Onur I, et al. Acute effects of smoking on coronary flow velocity reserve and ventricular diastolic functions. *Int J Cardiol*. 2008;129(1):e18–20.
373. Alshehri AM, Azoz AM, Shaheen HA, et al. Acute effects of cigarette smoking on the cardiac diastolic functions. *J Saudi Hear Assoc*. 2013;25(3):173–9.
374. Dogan A, Yarlioglu M, Gul I, et al. Acute effects of passive smoking on left ventricular systolic and diastolic function in healthy volunteers. *J Am Soc Echocardiogr*. 2011;24(2):185–91.
375. Yarlioglu M, Kaya MG, Ardic I, et al. Dose-dependent acute effects of passive smoking on left ventricular cardiac functions in healthy volunteers. *J Investig Med*. 2012;60(2):517–22.
376. Wang Y, Moreira Mda C, Heringer-Walther S, et al. Does the aminopeptidase A have prognostic and diagnostic value in Chagas disease and other dilated cardiomyopathies? *J Cardiovasc Pharmacol*. 2011;58(4):374–9.
377. Le TH, Coffman TM. Targeting genes in the renin-angiotensin system. *Curr Opin Nephrol Hypertens*. 2008;17(1):57–63.
378. Goto Y, Hattori A, Ishii Y, et al. Enzymatic properties of human aminopeptidase A. Regulation of its enzymatic activity by calcium and angiotensin IV. *J Biol Chem*. 2006;281(33):23503–13.
379. Padia SH, Kemp BA, Howell NL, et al. Conversion of renal angiotensin II to angiotensin III is critical for AT2 receptor-mediated natriuresis in rats. *Hypertension*. 2008;51(2):460–5.
380. Wright JW, Harding JW. Important roles for angiotensin III and IV in the brain renin-angiotensin system. *Brain Res Brain Res Rev*. 1997;25(1):96–124.

Literaturverzeichnis

381. Chai SY, Fernando R, Peck G, et al. The angiotensin IV/AT4 receptor. *Cell Mol Life Sci.* 2004;61(21):2728–37.
382. Yatabe J, Yoneda M, Yatabe MS, et al. Angiotensin III stimulates aldosterone secretion from adrenal gland partially via angiotensin II type 2 receptor but not angiotensin II type 1 receptor. *Endocrinology.* 2011;152(4):1582–8.
383. Yang Y, Liu C, Lin YL, Li F. Structural insights into central hypertension regulation by human aminopeptidase A. *J Biol Chem.* 2013;288(35):25638–45.
384. Li JM, Cui TX, Shiuchi T, et al. Nicotine enhances angiotensin II-induced mitogenic response in vascular smooth muscle cells and fibroblasts. *Arterioscler Thromb Vasc Biol.* 2004;24(1):80–4.
385. Gropelli A, Giorgi DM, Omboni S, et al. Persistent blood pressure increase induced by heavy smoking. *J Hypertens.* 1992;10(5):495–9.
386. Laustiola KE, Lassila R, Nurmi AK. Enhanced activation of the renin-angiotensin-aldosterone system in chronic cigarette smokers: a study of monozygotic twin pairs discordant for smoking. *Clin Pharmacol Ther.* 1988;44(4):426–30.
387. Kotamäki M. Smoking induced differences in autonomic responses in military pilot candidates. *Clin Auton Res.* 1995;5(1):31–6.
388. Kahn ML, Nakanishi-Matsui M, Shapiro MJ, et al. Protease-activated receptors 1 and 4 mediate activation of human platelets by thrombin. *J Clin Invest.* 1999;103(6):879–87.
389. Leger AJ, Covic L, Kuliopulos A. Protease-activated receptors in cardiovascular diseases. *Circulation.* 2006;114(10):1070–7.
390. Fu Q, Cheng J, Gao Y, et al. Protease-Activated Receptor 4: A Critical Participator in Inflammatory Response. *Inflammation.* 2014.
391. Daubie V, Pochet R, Houard S, Philippart P. Tissue factor: a mini-review. *J Tissue Eng Regen Med.* 2007;1(3):161–9.
392. Chapin JC, Hajjar KA. Fibrinolysis and the control of blood coagulation. *Blood Rev.* 2014.
393. Pomp ER, Rosendaal FR, Doggen CJ. Smoking increases the risk of venous thrombosis and acts synergistically with oral contraceptive use. *Am J Hematol.* 2008;83(2):97–102.
394. Campbell RA, Machlus KR, Wolberg AS. Smoking out the cause of thrombosis. *Arterioscler Thromb Vasc Biol.* 2010;30(1):7–8.
395. Zhang G, Xu X, Su W, Xu Q. Smoking and risk of venous thromboembolism: a systematic review. *Southeast Asian J Trop Med Public Health.* 2014;45(3):736–45.
396. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *J Am Coll Cardiol.* 2004;43(10):1731–7.
397. Barua RS, Ambrose JA. Mechanisms of coronary thrombosis in cigarette smoke exposure. *Arterioscler Thromb Vasc Biol.* 2013;33(7):1460–7.
398. Yu Y, Liu H, Zheng S, et al. Gender susceptibility for cigarette smoking-attributable lung cancer: a systematic review and meta-analysis. *Lung Cancer.* 2014;85(3):351–60.
399. Zudaire E, Cuesta N, Murty V, et al. The aryl hydrocarbon receptor repressor is a putative tumor suppressor gene in multiple human cancers. *J Clin Invest.* 2008;118(2): 640–50.
400. Chiba T, Uchi H, Yasukawa F, Furue M. Role of the arylhydrocarbon receptor in lung disease. *Int Arch Allergy Immunol.* 2011;155 Suppl 1:129–34.
401. Andersson P, Mcguire J, Rubio C, et al. A constitutively active dioxin/aryl hydrocarbon receptor induces stomach tumors. *Proc Natl Acad Sci U S A.* 2002;99(15):9990–5.
402. Moennikes O, Loeppen S, Buchmann A, et al. A constitutively active dioxin/aryl hydrocarbon receptor promotes hepatocarcinogenesis in mice. *Cancer Res.* 2004;64(14):4707–10.

403. Mimura J, Ema M, Sogawa K, Fujii-Kuriyama Y. Identification of a novel mechanism of regulation of Ah (dioxin) receptor function. *Genes Dev.* 1999;13(1):20–5.
404. Karchner SI, Franks DG, Powell WH, Hahn ME. Regulatory interactions among three members of the vertebrate aryl hydrocarbon receptor family: AHR repressor, AHR1, and AHR2. *J Biol Chem.* 2002;277(9):6949–59.
405. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer.* 2003;3(10):733–44.
406. Cauchi S, Stücker I, Cénée S, et al. Structure and polymorphisms of human aryl hydrocarbon receptor repressor (AhRR) gene in a French population: relationship with CYP1A1 inducibility and lung cancer. *Pharmacogenetics.* 2003;13(6):339–47.
407. Watanabe M, Sueoka K, Sasagawa I, et al. Association of male infertility with Pro185Ala polymorphism in the aryl hydrocarbon receptor repressor gene: implication for the susceptibility to dioxins. *Fertil Steril.* 2004;82 Suppl 3:1067–71.
408. Merisalu A, Punab M, Altmäe S, et al. The contribution of genetic variations of aryl hydrocarbon receptor pathway genes to male factor infertility. *Fertil Steril.* 2007;88(4):854–9.
409. Ruosaari ST, Nymark PE, Aavikko MM, et al. Aberrations of chromosome 19 in asbestos-associated lung cancer and in asbestos-induced micronuclei of bronchial epithelial cells in vitro. *Carcinogenesis.* 2008;29(5):913–7.
410. Chen D, Tian T, Wang H, et al. Association of human aryl hydrocarbon receptor gene polymorphisms with risk of lung cancer among cigarette smokers in a Chinese population. *Pharmacogenet Genomics.* 2009;19(1):25–34.
411. Falanga A, Ofosu FA, Delaini F, et al. The hypercoagulable state in cancer patients: evidence for impaired thrombin inhibitions. *Blood Coagul Fibrinolysis.* 1994;5 Suppl 1:S19–23; discussion 59–64.
412. Lee JM, Lim JH, Kim JS, et al. Multiple hypercoagulability disorders at presentation of non-small-cell lung cancer. *Tuberc Respir Dis (Seoul).* 2014;77(1):34–7.
413. Scesnaite A, Jarmalaite S, Mutanen P, et al. Similar DNA methylation pattern in lung tumours from smokers and never-smokers with second-hand tobacco smoke exposure. *Mutagenesis.* 2012;27(4):423–9.
414. Bhatti P, Zhang Y, Song X, et al. Nightshift work and genome-wide DNA methylation. *Chronobiol Int.* 2014:1–10.
415. Bouchard L, Tchernof A, Deshaies Y, et al. ZFP36: a promising candidate gene for obesity-related metabolic complications identified by converging genomics. *Obes Surg.* 2007;17(3):372–82.
416. Harada H, Harada Y, Brien DP, et al. HERF1, a novel hematopoiesis-specific RING finger protein, is required for terminal differentiation of erythroid cells. *Mol Cell Biol.* 1999;19(5):3808–15.
417. Koury S, Yarlagadda S, Moskalik-Liermo K, et al. Differential gene expression during terminal erythroid differentiation. *Genomics.* 2007;90(5):574–82.
418. Ge K, Guermah M, Yuan CX, et al. Transcription coactivator TRAP220 is required for PPAR gamma 2-stimulated adipogenesis. *Nature.* 2002;417(6888):563–7.
419. Bai L, Jia Y, Viswakarma N, et al. Transcription coactivator mediator subunit MED1 is required for the development of fatty liver in the mouse. *Hepatology.* 2011;53(4):1164–74.
420. Becerril S, Rodríguez A, Catalán V, et al. Transcriptional analysis of brown adipose tissue in leptin-deficient mice lacking inducible nitric oxide synthase: evidence of the role of Med1 in energy balance. *Physiol Genomics.* 2012;44(13):678–88.

Literaturverzeichnis

421. Erhardt A, Czibere L, Roeske D, et al. TMEM132D, a new candidate for anxiety phenotypes: evidence from human and mouse studies. *Mol Psychiatry*. 2011;16(6):647–63.
422. Erhardt A, Akula N, Schumacher J, et al. Replication and meta-analysis of TMEM132D gene variants in panic disorder. *Transl Psychiatry*. 2012;2:e156.
423. de Almondes KM, Araújo JF. The impact of different shift work schedules on the levels of anxiety and stress in workers in a petrochemicals company. *Estudos de Psicologia*. 2009;26(1):15–23
424. Carrillo J, Becerra M. Anxiety and depression associated with sleep disturbances in shift workers. *Sleep Med*. 2013;14 Suppl 1:e91.
425. Thoeringer CK, Ripke S, Unschuld PG, et al. The GABA transporter 1 (SLC6A1): a novel candidate gene for anxiety disorders. *J Neural Transm*. 2009;116(6):649–57.
426. Liu GX, Cai GQ, Cai YQ, et al. Reduced anxiety and depression-like behaviors in mice lacking GABA transporter subtype 1. *Neuropsychopharmacology*. 2007;32(7):1531–9.
427. Chiu CS, Brickley S, Jensen K, et al. GABA transporter deficiency causes tremor, ataxia, nervousness, and increased GABA-induced tonic conductance in cerebellum. *J Neurosci*. 2005;25(12):3234–45.
428. Montalbetti N, Simonin A, Kovacs G, Hediger MA. Mammalian iron transporters: families SLC11 and SLC40. *Mol Aspects Med*. 2013;34(2–3):270–87.
429. Mohren DC, Jansen NW, Kant IJ, et al. Prevalence of common infections among employees in different work schedules. *J Occup Environ Med*. 2002;44(11):1003–11.
430. Pinnix ZK, Miller LD, Wang W, et al. Ferroportin and iron regulation in breast cancer progression and prognosis. *Sci Transl Med*. 2010;2(43):43ra56.
431. Hansen J. Increased breast cancer risk among women who work predominantly at night. *Epidemiology*. 2001;12(1):74–7.
432. Schernhammer ES, Kroenke CH, Laden F, Hankinson SE. Night work and risk of breast cancer. *Epidemiology*. 2006;17(1):108–11.
433. Menegaux F, Truong T, Anger A, et al. Night work and breast cancer: a population-based case-control study in France (the CECILE study). *Int J Cancer*. 2013;132(4):924–31.
434. Parent MÉ, El-Zein M, Rousseau MC, et al. Night work and the risk of cancer among men. *Am J Epidemiol*. 2012;176(9):751–9.
435. Lim AS, Srivastava GP, Yu L, et al. 24-hour rhythms of DNA methylation and their relation with rhythms of RNA expression in the human dorsolateral prefrontal cortex. *PLoS Genet*. 2014;10(11):e1004792.
436. Ehrich M, Zoll S, Sur S, van den Boom D. A new method for accurate assessment of DNA quality after bisulfite treatment. *Nucleic Acids Res*. 2007;35(5):e29.
437. Wan ES, Qiu W, Baccarelli A, et al. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet*. 2012;21(13):3073–82.
438. Zhang Y, Yang R, Burwinkel B, et al. F2RL3 methylation as a biomarker of current and lifetime smoking exposures. *Environ Health Perspect*. 2014;122(2):131–7.
439. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.
440. Mundade R, Ozer HG, Wei H, et al. Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. *Cell Cycle*. 2014;13(18):2847–52.

Literaturverzeichnis

441. Griffon A, Barbier Q, Dalino J, et al. Integrative analysis of public ChIP-seq experiments reveals a complex multi-cell regulatory landscape. *Nucleic Acids Res.* 2014; doi: 10.1093/nar/gku1280.
442. Chen H, Dzitoyeva S, Manev H. Effect of aging on 5-hydroxymethylcytosine in the mouse hippocampus. *Restor Neurol Neurosci.* 2012;30(3):237–45.
443. Tammen SA, Dolnikowski GG, Ausman LM, et al. Aging alters hepatic DNA hydroxymethylation, as measured by liquid chromatography/mass spectrometry. *J Cancer Prev.* 2014;19(4):301–8.
444. Dao T, Cheng RY, Revelo MP, et al. Hydroxymethylation as a Novel Environmental Biosensor. *Curr Environ Health Rep.* 2014;1(1):1–10.
445. Nestor C, Ruzov A, Meehan R, Dunican D. Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *Biotechniques.* 2010;48(4):317–9.

9 Anhang

Tabelle A1: Verteilung der Zwillings-Proben über die HM450K BeadChips, die detektierten CpGs ($p < 0,01$) und die korrespondierende loci detection rate pro Probe.
Die für die DNA-Methylierungsanalysen verwendeten Merkmale der Zwillinge sind angegeben.

Paar Nr.	Zwilling ID	Studie	Geschlecht	Alter [Jahre]	Rauchstatus	Arbeitszeit	BMI [kg/m ²]	Charge	Array	Array Position	Detektierte CpGs ^a	LDR [%] _b
1	MZM1_ZP1	GEMINAKAR	M	36	n.b.	n.b.	n.b.	2	7959804097	R01C01	485378	99,96
1	MZM1_ZP1	SAME	M	46	Nichtraucher	mit Nachtschicht	28,7	2	7959804097	R02C01	485420	99,97
1	MZM2_ZP1	GEMINAKAR	M	36	n.b.	n.b.	n.b.	2	7959804097	R03C01	485367	99,96
1	MZM2_ZP1	SAME	M	46	Nichtraucher	Tag	27,3	2	7959804097	R04C01	485362	99,96
2	DZF2_ZP2	GEMINAKAR	W	31	n.b.	n.b.	n.b.	2	7959804097	R01C02	485198	99,92
2	DZF2_ZP2	SAME	W	41	Nichtraucher	mit Nachtschicht	21,3	2	7959804097	R02C02	485137	99,91
2	DZF1_ZP2	GEMINAKAR	W	31	n.b.	n.b.	n.b.	2	7959804097	R05C01	485062	99,89
2	DZF1_ZP2	SAME	W	41	Nichtraucher	Tag	24,1	2	7959804097	R06C01	485071	99,90
3	MZM1_ZP3	GEMINAKAR	M	26	n.b.	n.b.	n.b.	2	7959804097	R03C02	485327	99,95
3	MZM1_ZP3	SAME	M	36	Gelegenheitsraucher	Tag	33,0	2	7959804097	R04C02	485324	99,95
3	MZM2_ZP3	GEMINAKAR	M	26	n.b.	n.b.	n.b.	2	7959804097	R05C02	485257	99,93
3	MZM2_ZP3	SAME	M	36	ehemaliger Raucher	mit Nachtschicht	27,9	2	7959804097	R06C02	485267	99,94
4	DZM_Z4	GEMINAKAR	M	34	n.b.	n.b.	n.b.	3	7959804114	R01C01	485443	99,97
4	DZM_Z4	SAME	M	44	Raucher	mit Nachtschicht	21,9	3	7959804114	R02C01	485419	99,97
5	DZF1_ZP5	SAME	W	40	Nichtraucher	mit Nachtschicht	25,5	1	7970368126	R01C01	485246	99,93
5	DZF2_ZP5	SAME	W	40	n.b.	Tag	29,6	1	7970368126	R02C01	485238	99,93
5	DZF1_ZP5	GEMINAKAR	W	30	n.b.	n.b.	n.b.	1	7973201129	R01C02	485150	99,91
5	DZF2_ZP5	GEMINAKAR	W	30	n.b.	n.b.	n.b.	1	7973201129	R02C02	485159	99,91
6	DZM1_ZP6	GEMINAKAR	M	46	n.b.	n.b.	n.b.	1	7970368126	R01C02	485442	99,97
6	DZM2_ZP6	GEMINAKAR	M	46	n.b.	n.b.	n.b.	1	7970368126	R02C02	485451	99,97
6	DZM1_ZP6	SAME	M	56	Nichtraucher	mit Nachtschicht	30	1	7970376025	R01C01	485287	99,94
6	DZM2_ZP6	SAME	M	56	Raucher	Abendarbeit	27,3	1	7970376025	R02C01	485261	99,93
7	DZF1_ZP7	SAME	W	35	Nichtraucher	mit Nachtschicht	24,6	1	7970368126	R03C01	483560	99,58
7	DZF2_ZP7	SAME	W	35	Raucher	Tag	28,7	1	7970368126	R04C01	481720	99,21

Anhang

Paar Nr.	Zwilling ID	Studie	Geschlecht	Alter [Jahre]	Rauchstatus	Arbeitszeit	BMI [kg/m ²]	Charge	Array	Array Position	Detektierte CpGs ^a	LDR [%] ^b
7	DZF1_3020421	GEMINAKAR	W	25	n.b.	n.b.	n.b.	1	7973201129	R03C02	485116	99,91
7	DZF2_3020421	GEMINAKAR	W	25	n.b.	n.b.	n.b.	1	7973201129	R04C02	485109	99,90
8	DZF1_ZP8	GEMINAKAR	W	45	n.b.	n.b.	n.b.	1	7970368126	R03C02	485090	99,90
8	DZF2_ZP8	GEMINAKAR	W	45	n.b.	n.b.	n.b.	1	7970368126	R04C02	485118	99,91
8	DZF1_ZP8	SAME	W	55	Nichtraucher	Abendarbeit	21,1	1	7970376025	R03C01	485078	99,90
8	DZF2_ZP8	SAME	W	55	Nichtraucher	Tag	23,3	1	7970376025	R04C01	485135	99,91
9	MZM1_ZP9	SAME	M	40	Nichtraucher	mit Nachtschicht	24,6	1	7970368126	R05C01	485371	99,96
9	MZM2_ZP9	SAME	M	40	Nichtraucher	Tag	23,0	1	7970368126	R06C01	484354	99,75
9	MZM1_ZP9	GEMINAKAR	M	30	n.b.	n.b.	n.b.	1	7973201129	R05C02	485340	99,95
9	MZM2_ZP9	GEMINAKAR	M	30	n.b.	n.b.	n.b.	1	7973201129	R06C02	485229	99,93
10	MZF1_ZP10	GEMINAKAR	W	45	n.b.	n.b.	n.b.	1	7970368126	R05C02	485148	99,91
10	MZF2_ZP10	GEMINAKAR	W	45	n.b.	n.b.	n.b.	1	7970368126	R06C02	485027	99,89
10	MZF1_ZP10	SAME	W	55	Nichtraucher	mit Nachtschicht	23,4	1	7970376025	R05C01	484978	99,88
10	MZF2_ZP10	SAME	W	55	ehemaliger Raucher	Tag	20,9	1	7970376025	R06C01	482695	99,41
11	MZM1_ZP11	GEMINAKAR	M	48	n.b.	n.b.	n.b.	2	7970368137	R01C01	485490	99,98
11	MZM1_ZP11	SAME	M	58	Nichtraucher	Tag	23,7	2	7970368137	R02C01	485456	99,98
11	MZM2_ZP11	GEMINAKAR	M	47	n.b.	n.b.	n.b.	2	7970368137	R03C01	485433	99,97
11	MZM2_ZP11	SAME	M	57	ehemaliger Raucher	mit Nachtschicht	26,0	2	7970368137	R04C01	485459	99,98
12	MZM2_ZP12	GEMINAKAR	M	43	n.b.	n.b.	n.b.	2	7970368137	R01C02	485439	99,97
12	MZM2_ZP12	SAME	M	53	Nichtraucher	mit Nachtschicht	25,8	2	7970368137	R02C02	485470	99,98
12	MZM1_ZP12	GEMINAKAR	M	43	n.b.	n.b.	n.b.	2	7970368137	R05C01	485434	99,97
12	MZM1_ZP12	SAME	M	53	Nichtraucher	Tag	24,1	2	7970368137	R06C01	485392	99,96
13	DZM1_ZP13	GEMINAKAR	M	40	n.b.	n.b.	n.b.	5	7970368137	R03C02	485445	99,97
13	DZM1_ZP13	SAME	M	50	ehemaliger Raucher	mit Nachtschicht	24,5	5	7970368137	R04C02	485449	99,97
13	DZM2_ZP13	GEMINAKAR	M	40	n.b.	n.b.	n.b.	5	7970368137	R05C02	485394	99,96
13	DZM2_ZP13	SAME	M	50	Nichtraucher	Tag	25,5	5	7970368137	R06C02	485140	99,91

Anhang

Paar Nr.	Zwilling ID	Studie	Geschlecht	Alter [Jahre]	Rauchstatus	Arbeitszeit	BMI [kg/m ²]	Charge	Array	Array Position	Detektierte CpGs ^a	LDR [%] ^b
14	DZF1_ZP14	SAME	W	52	ehemaliger Raucher	Tag	39,3	1	7970376025	R01C02	484655	99,81
14	DZF2_ZP14	SAME	W	52	Raucher	Abendarbeit	41,7	1	7970376025	R02C02	484739	99,83
14	DZF1_ZP14	GEMINAKAR	W	42	n.b.	n.b.	n.b.	1	7973201129	R01C01	485163	99,91
14	DZF2_ZP14	GEMINAKAR	W	42	n.b.	n.b.	n.b.	1	7973201129	R02C01	485103	99,90
15	DZF1_ZP15	SAME	W	34	n.b.	Tag	n.b.	1	7970376025	R03C02	481000	99,06
15	DZF2_ZP15	SAME	W	34	Nichtraucher	mit Nachtschicht	30,3	1	7970376025	R04C02	484284	99,73
15	DZF1_ZP15	GEMINAKAR	W	24	n.b.	n.b.	n.b.	1	7973201129	R03C01	485027	99,89
15	DZF2_ZP15	GEMINAKAR	W	24	n.b.	n.b.	n.b.	1	7973201129	R04C01	485082	99,90
16	MZM1_ZP16	SAME	W	47	ehemaliger Raucher	mit Nachtschicht	20,8	1	7970376025	R05C02	483957	99,67
16	MZM2_ZP16	SAME	W	47	Raucher	Tag	21,5	1	7970376025	R06C02	483774	99,63
16	MZM1_ZP16	GEMINAKAR	M	37	n.b.	n.b.	n.b.	1	7973201129	R05C01	485387	99,96
16	MZM2_ZP16	GEMINAKAR	M	37	n.b.	n.b.	n.b.	1	7973201129	R06C01	485265	99,94
17	MZF2_ZP17	SAME	W	48	ehemaliger Raucher	mit Nachtschicht	30,7	2	8454787063	R01C01	485061	99,89
17	MZF2_ZP17	GEMINAKAR	W	38	n.b.	n.b.	n.b.	2	8454787063	R02C01	484907	99,86
17	MZF1_ZP17	SAME	W	48	Nichtraucher	Tag	26,8	2	8454787063	R03C01	484980	99,88
17	MZF1_ZP17	GEMINAKAR	W	38	n.b.	n.b.	n.b.	2	8454787063	R04C01	484966	99,87
18	DZM_Z18	GEMINAKAR	M	46	n.b.	n.b.	n.b.	3	8454787063	R01C02	485319	99,95
18	DZM_Z18	SAME	M	56	Raucher	mit Nachtschicht	24,0	3	8454787063	R02C02	485354	99,95
19	DZF1_ZP19	GEMINAKAR	W	29	n.b.	n.b.	n.b.	2	8454787063	R03C02	484973	99,88
19	DZF1_ZP19	SAME	W	39	Nichtraucher	Tag	28,7	2	8454787063	R04C02	484983	99,88
19	DZM2_ZP19	GEMINAKAR	M	28	n.b.	n.b.	n.b.	2	8454787063	R05C02	485272	99,94
19	DZM2_ZP19	SAME	M	38	Gelegenheitsraucher	Nachtdienst	30,2	2	8454787063	R06C02	485307	99,94
20	MZF2_ZP20	GEMINAKAR	W	28	n.b.	n.b.	n.b.	2	8454787063	R05C01	484925	99,87
20	MZF2_ZP20	SAME	W	38	Nichtraucher	Tag	19,5	2	8454787063	R06C01	484939	99,87
20	MZF1_ZP20	SAME	W	38	Nichtraucher	mit Nachtschicht	19,1	2	8667045098	R03C02	485131	99,91
20	MZF1_ZP20	GEMINAKAR	W	28	n.b.	n.b.	n.b.	2	8667045098	R04C02	485082	99,90

Anhang

Paar Nr.	Zwilling ID	Studie	Geschlecht	Alter [Jahre]	Rauchstatus	Arbeitszeit	BMI [kg/m ²]	Charge	Array	Array Position	Detektierte CpGs ^a	LDR [%] ^b
21	MZF1_ZP21	SAME	W	57	Raucher	Tag	21,4	4	8454787086	R01C01	485077	99,90
21	MZF1_ZP21	GEMINAKAR	W	47	n.b.	n.b.	n.b.	4	8454787086	R01C02	485182	99,92
21	MZF2_ZP21	SAME	W	57	ehemaliger Raucher	mit Nachtschicht	21,8	4	8454787086	R02C01	485011	99,88
21	MZF2_ZP21	GEMINAKAR	W	47	n.b.	n.b.	n.b.	4	8454787086	R02C02	485065	99,89
22	DZF1_ZP22	SAME	W	33	Nichtraucher	mit Nachtschicht	23,0	4	8454787086	R03C01	485045	99,89
22	DZF1_ZP22	GEMINAKAR	W	23	n.b.	n.b.	n.b.	4	8454787086	R03C02	485054	99,89
22	DZM2_ZP22	SAME	M	34	Raucher	ohne Nachtschicht	21,6	4	8454787086	R04C01	485177	99,92
22	DZM2_ZP22	GEMINAKAR	M	24	n.b.	n.b.	n.b.	4	8454787086	R04C02	485240	99,93
23	MZF1_ZP23	SAME	W	32	Nichtraucher	ohne Nachtschicht	19,4	4	8454787086	R05C01	484684	99,82
23	MZF1_ZP23	GEMINAKAR	W	22	n.b.	n.b.	n.b.	4	8454787086	R05C02	484854	99,85
23	MZF2_ZP23	SAME	W	32	Nichtraucher	mit Nachtschicht	20,6	4	8454787086	R06C01	484667	99,81
23	MZF2_ZP23	GEMINAKAR	W	22	n.b.	n.b.	n.b.	4	8454787086	R06C02	484840	99,85
24	DZM1_ZP24	SAME	M	31	Raucher	Tag	29,1	4	8454787087	R01C01	485457	99,98
24	DZM2_ZP24	GEMINAKAR	M	21	n.b.	n.b.	n.b.	4	8454787087	R01C02	485422	99,97
24	DZM2_ZP24	SAME	M	31	ehemaliger Raucher	mit Nachtschicht	25,5	4	8454787087	R02C01	485375	99,96
24	DZM1_ZP24	GEMINAKAR	M	21	n.b.	n.b.	n.b.	4	8454787087	R06C01	485165	99,92
25	DZF1_Z40	GEMINAKAR	M	20	n.b.	n.b.	n.b.	4	8454787087	R02C02	485043	99,89
25	DZF1_Z40	SAME	W	30	ehemaliger Raucher	mit Nachtschicht	25,9	4	8454787087	R03C01	485032	99,89
26	DZM1_ZP26	GEMINAKAR	M	39	n.b.	n.b.	n.b.	4	8454787087	R03C02	485414	99,97
26	DZM1_ZP26	SAME	M	49	ehemaliger Raucher	Tag	21,9	4	8454787087	R04C01	485342	99,95
26	DZM2_ZP26	GEMINAKAR	M	39	n.b.	n.b.	n.b.	4	8454787087	R04C02	485335	99,95
26	DZM2_ZP26	SAME	M	49	Nichtraucher	Nachtdienst	26,1	4	8454787087	R05C01	485269	99,94
27	DZF1_ZP27	GEMINAKAR	W	24	n.b.	n.b.	n.b.	2	8454787111	R01C01	485009	99,88
27	DZF1_ZP27	SAME	W	34	Nichtraucher	mit Nachtschicht	21,8	2	8454787111	R02C01	485046	99,89
27	DZF2_ZP27	GEMINAKAR	W	23	n.b.	n.b.	n.b.	2	8454787111	R03C01	484930	99,87
27	DZF2_ZP27	SAME	W	33	Gelegenheitsraucher	Tag	21,9	2	8454787111	R04C01	484949	99,87

Anhang

Paar Nr.	Zwilling ID	Studie	Geschlecht	Alter [Jahre]	Rauchstatus	Arbeitszeit	BMI [kg/m ²]	Charge	Array	Array Position	Detektierte CpGs ^a	LDR [%] ^b
28	DZF2_ZP28	GEMINAKAR	W	53	n.b.	n.b.	n.b.	3	8454787111	R01C02	485062	99,89
28	DZF2_ZP28	SAME	W	63	Nichtraucher	Tag	23,0	3	8454787111	R02C02	485152	99,91
28	DZF1_ZP28	GEMINAKAR	W	53	n.b.	n.b.	n.b.	3	8454787111	R05C01	484913	99,86
28	DZF1_ZP28	SAME	W	63	Nichtraucher	mit Nachtschicht	20,6	3	8454787111	R06C01	484988	99,88
29	DZF1_ZP29	GEMINAKAR	W	37	n.b.	n.b.	n.b.	2	8454787111	R03C02	485046	99,89
29	DZF1_ZP29	SAME	W	47	Nichtraucher	Tag	27,0	2	8454787111	R04C02	484982	99,88
29	DZM2_ZP29	GEMINAKAR	M	37	n.b.	n.b.	n.b.	2	8454787111	R05C02	485305	99,94
29	DZM2_ZP29	SAME	M	47	Nichtraucher	Nachtdienst	21,7	2	8454787111	R06C02	485335	99,95
30	DZF1_ZP30	GEMINAKAR	W	37	n.b.	n.b.	n.b.	2	8667045098	R01C01	485070	99,90
30	DZF1_ZP30	SAME	W	47	Nichtraucher	mit Nachtschicht	25,8	2	8667045098	R02C01	485055	99,89
30	DZF2_ZP30	GEMINAKAR	W	37	n.b.	n.b.	n.b.	2	8667045098	R03C01	484996	99,88
30	DZF2_ZP30	SAME	W	47	Nichtraucher	Tag	30,4	2	8667045098	R04C01	484951	99,87
31	MZF1_ZP31	SAME	W	47	Nichtraucher	mit Nachtschicht	21,1	3	8667045098	R05C01	484986	99,88
31	MZF2_ZP31	SAME	W	46	Nichtraucher	mit Nachtschicht	20,2	3	8667045098	R06C01	485004	99,88
32	MZM1_ZP32	GEMINAKAR	M	36	n.b.	n.b.	n.b.	2	8667045146	R01C01	485573	100,00
32	MZM1_ZP32	SAME	M	46	Nichtraucher	Tag	24,6	2	8667045146	R02C01	485488	99,98
32	MZM2_ZP32	GEMINAKAR	M	36	n.b.	n.b.	n.b.	2	8667045146	R03C01	485459	99,98
32	MZM2_ZP32	SAME	M	46	Nichtraucher	Nachtdienst	24,0	2	8667045146	R04C01	485465	99,98
33	DZF1_ZP33	GEMINAKAR	W	30	n.b.	n.b.	n.b.	2	8667045146	R05C01	485056	99,89
33	DZF1_ZP33	SAME	W	40	Raucher	mit Nachtschicht	25,7	2	8667045146	R06C01	485024	99,89
33	DZM2_ZP33	GEMINAKAR	W	30	n.b.	n.b.	n.b.	2	8667045146	R01C02	485476	99,98
33	DZM2_ZP33	SAME	W	40	Nichtraucher	Tag	19,8	2	8667045146	R02C02	485547	99,99
34	MZF1_ZP34	GEMINAKAR	W	33	n.b.	n.b.	n.b.	2	8667045146	R03C02	485536	99,99
34	MZF1_ZP34	SAME	W	43	Nichtraucher	mit Nachtschicht	27,5	2	8667045146	R04C02	485522	99,99
34	MZF2_ZP34	GEMINAKAR	W	33	n.b.	n.b.	n.b.	2	8667045146	R05C02	485023	99,89
34	MZF2_ZP34	SAME	W	43	Nichtraucher	Tag	22,6	2	8667045146	R06C02	485095	99,90

Anhang

Paar Nr.	Zwilling ID	Studie	Geschlecht	Alter [Jahre]	Rauchstatus	Arbeitszeit	BMI [kg/m ²]	Charge	Array	Array Position	Detektierte CpGs ^a	LDR [%] ^b
35	DZM1_ZP35	GEMINAKAR	M	42	n.b.	n.b.	n.b.	2	8691803007	R01C01	485525	99,99
35	DZM1_ZP35	SAME	M	52	Raucher	Tag	24,1	2	8691803007	R02C01	485511	99,99
35	DZM2_ZP35	GEMINAKAR	M	42	n.b.	n.b.	n.b.	2	8691803007	R03C01	485438	99,97
35	DZM2_ZP35	SAME	M	52	Raucher	mit Nachtschicht	23,4	2	8691803007	R04C01	485418	99,97
36	MZM1_ZP36	GEMINAKAR	M	39	n.b.	n.b.	n.b.	2	8691803007	R01C02	485483	99,98
36	MZM2_ZP36	SAME	M	49	ehemaliger Raucher	Tag	28,0	2	8691803007	R02C02	485479	99,98
36	MZM1_ZP36	GEMINAKAR	M	39	n.b.	n.b.	n.b.	2	8691803007	R05C01	485346	99,95
36	MZM1_ZP36	SAME	M	49	Gelegenheitsraucher	mit Nachtschicht	27,9	2	8691803007	R06C01	484981	99,88
37	MZF1_ZP37	GEMINAKAR	W	29	n.b.	n.b.	n.b.	2	8691803007	R03C02	485128	99,91
37	MZF1_ZP37	SAME	W	39	ehemaliger Raucher	Tag	21,9	2	8691803007	R04C02	485053	99,89
37	MZF2_ZP37	GEMINAKAR	W	29	n.b.	n.b.	n.b.	3	8691803007	R05C02	485010	99,88
37	MZF2_ZP37	SAME	W	39	Raucher	Nachtdienst	22,8	3	8691803007	R06C02	484687	99,82
38	DZF1_ZP38	SAME	W	47	Raucher	Nachtdienst	37,5	6	9373551053	R01C01	485555	100,00
38	DZF2_ZP38	SAME	W	47	Raucher	Tag	28,7	6	9373551053	R02C01	485568	100,00
39	MZF1_ZP39	SAME	W	52	Nichtraucher	Nachtdienst	26,4	6	9373551053	R03C01	485503	99,98
39	MZF2_ZP39	SAME	W	52	Nichtraucher	Tag	22,8	6	9373551053	R04C01	485099	99,90
40	DZF1_ZP25	SAME	W	42	Raucher	Abendarbeit	25,1	6	9373551053	R05C01	485040	99,89
40	DZF2_ZP25	SAME	W	42	Raucher	Nachtdienst	25,0	6	9373551053	R06C01	485527	99,99

DZ: Dizygot; MZ: Monozygot; GEMINAKAR: Studie 1997–2000; SAME: Studie 2010–2013; Alter: entspricht dem Alter zum Zeitpunkt der Blutentnahme; n.b.: Information nicht bekannt; R=Reihe, C=Spalte; ^a Anzahl der CpGs, die mit p<0,01 detektiert wurden; ^b LDR=*loci detection rate* [%] basierend auf der Gesamtanzahl der Loci auf dem HM450K BeadChip (n=485577).

Anhang

Tabelle A2: Liste der mittels Infinium HumanMethylation450 BeadChip in 116 DNA-Proben (20–63 Jahre) identifizierten 111 altersabhängig methylierten CpG Regionen ($q < 1 \times 10^{-11}$, $\sigma/\sigma_{\max} > 0,1$). Die CpG-Loci sind nach aufsteigender Position auf den jeweiligen Chromosomen geordnet.

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	q-Wert	p-Wert
cg25410668	1	28241577	RPA2	TSS1500	S Shore	$3,86 \times 10^{-19}$	$6,19 \times 10^{-23}$
cg24433287	1	38261057	MANEAL	Body	S Shore	$4,19 \times 10^{-14}$	$2,51 \times 10^{-17}$
cg22851420	1	40149650	HPCAL4	Body	Island	$1,05 \times 10^{-12}$	$9,58 \times 10^{-16}$
cg24466241	1	53308908	ZYG11A	Body	Island	$1,94 \times 10^{-16}$	$5,58 \times 10^{-20}$
cg18933331	1	110186418	N/A	N/A	S Shore	$3,76 \times 10^{-14}$	$2,17 \times 10^{-17}$
cg16008966	1	114761794	N/A	N/A	N/A	$1,41 \times 10^{-12}$	$1,31 \times 10^{-15}$
cg17326555	1	119535693	N/A	N/A	Island	$1,39 \times 10^{-13}$	$9,23 \times 10^{-17}$
cg10501210	1	207997020	N/A	N/A	N/A	$2,06 \times 10^{-24}$	$8,81 \times 10^{-29}$
cg01557798	1	228400628	OBSCN	Body	Island	$3,74 \times 10^{-14}$	$2,08 \times 10^{-17}$
cg22303418	2	63284784	N/A	N/A	N Shore	$2,31 \times 10^{-14}$	$1,23 \times 10^{-17}$
cg22943590	2	66648797	N/A	N/A	N Shelf	$1,80 \times 10^{-12}$	$1,77 \times 10^{-15}$
cg06639320	2	106015739	FHL2	TSS200	Island	$1,71 \times 10^{-28}$	$3,65 \times 10^{-33}$
cg22454769	2	106015767	FHL2	TSS200	Island	$6,71 \times 10^{-23}$	$5,74 \times 10^{-27}$
cg24079702	2	106015771	FHL2	TSS200	Island	$6,78 \times 10^{-17}$	$1,67 \times 10^{-20}$
cg23606718	2	131513927	FAM123C	5'UTR	Island	$2,06 \times 10^{-13}$	$1,43 \times 10^{-16}$
cg06942701	2	162280009	TBR1	Body	Island	$4,85 \times 10^{-13}$	$3,78 \times 10^{-16}$
cg20065463	2	220283156	DES	1stExon	N Shore	$1,82 \times 10^{-13}$	$1,24 \times 10^{-16}$
cg24436906	2	242498081	BOK	TSS200	Island	$8,76 \times 10^{-17}$	$2,25 \times 10^{-20}$
cg11084334	3	9594264	LHFPL4	Body	Island	$1,30 \times 10^{-13}$	$8,35 \times 10^{-17}$
cg12899747	3	25391527	N/A	N/A	N/A	$7,85 \times 10^{-12}$	$9,14 \times 10^{-15}$
cg00664406	3	51740875	GRM2	TSS1500	Island	$3,66 \times 10^{-16}$	$1,13 \times 10^{-19}$
cg04453050	3	51740896	GRM2	TSS200	Island	$4,85 \times 10^{-13}$	$3,77 \times 10^{-16}$
cg12934382	3	51741135	GRM2	1stExon	Island	$2,04 \times 10^{-19}$	$2,94 \times 10^{-23}$
cg00329615	3	118706648	IGSF11	Body	N/A	$3,53 \times 10^{-12}$	$3,78 \times 10^{-15}$
cg05371578	3	147128157	ZIC1	1stExon	Island	$5,47 \times 10^{-12}$	$6,14 \times 10^{-15}$
cg19029181	3	147130536	ZIC1	Body	Island	$7,61 \times 10^{-13}$	$6,26 \times 10^{-16}$
cg06570224	3	157812475	N/A	N/A	Island	$1,53 \times 10^{-14}$	$7,53 \times 10^{-18}$
cg04521004	3	157821262	SHOX2	Body	Island	$5,14 \times 10^{-12}$	$5,71 \times 10^{-15}$
cg07553761	3	160167977	TRIM59	TSS1500	Island	$1,96 \times 10^{-17}$	$3,77 \times 10^{-21}$
cg15618978	3	160167990	TRIM59	TSS1500	Island	$8,47 \times 10^{-15}$	$3,82 \times 10^{-18}$
cg00481951	3	187387650	SST	Body	N Shore	$6,63 \times 10^{-12}$	$7,51 \times 10^{-15}$
cg23995914	4	10459228	ZNF518B	TSS200	Island	$1,67 \times 10^{-12}$	$1,57 \times 10^{-15}$
cg17621438	5	63461216	RNF180	TSS1500	N Shore	$1,97 \times 10^{-19}$	$2,53 \times 10^{-23}$
cg07850154	5	63461232	RNF180	TSS1500	N Shore	$1,95 \times 10^{-15}$	$7,08 \times 10^{-19}$
cg02699218	5	132150128	ANKRD43	1stExon	Island	$8,69 \times 10^{-12}$	$1,03 \times 10^{-14}$
cg23500537	5	140419819	N/A	N/A	N/A	$8,97 \times 10^{-21}$	$9,59 \times 10^{-25}$
cg00059225	5	151304357	GLRA1	1stExon	Island	$8,27 \times 10^{-12}$	$9,72 \times 10^{-15}$
cg23476877	5	176543964	N/A	N/A	Island	$3,70 \times 10^{-12}$	$3,99 \times 10^{-15}$
cg16867657	6	11044877	ELOVL2	TSS1500	Island	$2,54 \times 10^{-36}$	$2,71 \times 10^{-41}$
cg24724428	6	11044888	ELOVL2	TSS1500	Island	$1,10 \times 10^{-22}$	$1,06 \times 10^{-26}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	q-Wert	p-Wert
cg21572722	6	11044894	ELOVL2	TSS1500	Island	2,86x10 ⁻²⁶	9,18x10 ⁻³¹
cg22736354	6	18122719	NHLRC1	1stExon	Island	2,40x10 ⁻¹⁷	5,13x10 ⁻²¹
cg03236802	6	32055316	TNXB	Body	Island	7,35x10 ⁻¹³	5,97x10 ⁻¹⁶
cg13782301	6	32116875	PRRT1	3'UTR	Island	3,93x10 ⁻¹²	4,28x10 ⁻¹⁵
cg24125828	6	32117049	PRRT1	Body	Island	7,10x10 ⁻¹²	8,19x10 ⁻¹⁵
cg26350754	6	33043868	HLA-DPB1	1stExon	N/A	8,56x10 ⁻¹³	7,50x10 ⁻¹⁶
cg12939283	6	33044012	HLA-DPB1	Body	N/A	2,82x10 ⁻¹³	1,99x10 ⁻¹⁶
cg03738025	6	105388694	N/A	N/A	Island	3,46x10 ⁻¹³	2,55x10 ⁻¹⁶
cg17885226	6	105388731	N/A	N/A	Island	4,47x10 ⁻¹⁵	1,75x10 ⁻¹⁸
cg01047555	6	127835990	C6orf174	Body	Island	8,47x10 ⁻¹⁵	3,89x10 ⁻¹⁸
cg12597389	7	8482235	NXP1	Body	Island	2,40x10 ⁻¹²	2,47x10 ⁻¹⁵
cg04427498	7	20830657	N/A	N/A	Island	1,47x10 ⁻¹⁶	4,06x10 ⁻²⁰
cg26153045	7	73441834	ELN	TSS1500	N/A	2,94x10 ⁻¹²	3,11x10 ⁻¹⁵
cg14361627	7	130419116	KLF14	TSS1500	Island	2,90x10 ⁻²⁴	1,55x10 ⁻²⁸
cg17372101	7	147500722	CNTNAP2	Body	N/A	2,95x10 ⁻¹⁶	8,82x10 ⁻²⁰
cg16419235	8	57360613	PENK	TSS1500	Island	1,47x10 ⁻¹⁹	1,73x10 ⁻²³
cg20219381	8	101118083	RGS22	Body	Island	3,76x10 ⁻¹⁴	2,21x10 ⁻¹⁷
cg00688297	8	145752292	LRRC24	5'UTR	Island	1,47x10 ⁻¹⁶	4,08x10 ⁻²⁰
cg05419198	8	145753896	MGC70857	Body	Island	1,80x10 ⁻¹²	1,75x10 ⁻¹⁵
cg08809361	9	138192215	N/A	N/A	N Shore	2,90x10 ⁻¹²	3,04x10 ⁻¹⁵
cg20300246	9	139096747	LHX3	Body	Island	7,80x10 ⁻¹³	6,58x10 ⁻¹⁶
cg13823169	9	139776893	N/A	N/A	N Shelf	1,76x10 ⁻¹²	1,69x10 ⁻¹⁵
cg09099868	10	22622793	N/A	N/A	N Shore	2,40x10 ⁻¹⁷	4,90x10 ⁻²¹
cg10804656	10	22623460	N/A	N/A	Island	5,18x10 ⁻¹⁷	1,22x10 ⁻²⁰
cg13327545	10	22623548	N/A	N/A	Island	2,31x10 ⁻¹⁴	1,26x10 ⁻¹⁷
cg22796704	10	49673534	ARHGAP22	Body	N Shore	7,12x10 ⁻¹⁴	4,41x10 ⁻¹⁷
cg03354992	10	88149475	N/A	N/A	Island	4,59x10 ⁻¹³	3,48x10 ⁻¹⁶
cg13848598	10	115804578	ADRB1	1stExon	Island	3,62x10 ⁻¹³	2,71x10 ⁻¹⁶
cg11705975	10	120354248	PRLHR	Body	Island	1,92x10 ⁻¹⁴	9,85x10 ⁻¹⁸
cg19389884	11	1358515	N/A	N/A	Island	8,63x10 ⁻¹³	7,75x10 ⁻¹⁶
cg01097384	11	1358675	N/A	N/A	Island	1,95x10 ⁻¹⁵	6,98x10 ⁻¹⁹
cg15948836	11	36114126	LDLRAD3	Body	N/A	4,71x10 ⁻¹⁵	1,91x10 ⁻¹⁸
cg10915263	12	7592647	CD163L1	Body	Island	1,76x10 ⁻¹²	1,68x10 ⁻¹⁵
cg00748589	12	11653486	N/A	N/A	Island	4,47x10 ⁻¹⁵	1,73x10 ⁻¹⁸
cg18473521	12	54448265	HOXC4	Body	S Shore	1,18x10 ⁻¹²	1,09x10 ⁻¹⁵
cg10906284	12	63544430	AVPR1A	1stExon	Island	5,08x10 ⁻¹⁵	2,12x10 ⁻¹⁸
cg11649376	12	81473234	ACSS3	Body	S Shore	1,24x10 ⁻¹⁷	2,25x10 ⁻²¹
cg16519487	13	20876063	N/A	N/A	Island	2,21x10 ⁻¹²	2,24x10 ⁻¹⁵
cg05404236	13	110437093	IRS2	1stExon	Island	3,79x10 ⁻²³	2,83x10 ⁻²⁷
cg02547394	13	112721789	SOX1	TSS200	Island	7,65x10 ⁻¹³	6,38x10 ⁻¹⁶
cg02287710	14	102027660	DIO3	TSS200	Island	4,25x10 ⁻¹²	4,68x10 ⁻¹⁵
cg15341124	14	102027734	DIO3	5'UTR	Island	5,49x10 ⁻¹⁵	2,35x10 ⁻¹⁸
cg04875128	15	31775895	OTUD7A	Body	Island	2,04x10 ⁻¹⁹	3,05x10 ⁻²³

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	q-Wert	p-Wert
cg03467555	16	54404408	N/A	N/A	N/A	$2,59 \times 10^{-12}$	$2,69 \times 10^{-15}$
cg06279276	16	67184164	B3GNT9	Body	Island	$1,09 \times 10^{-18}$	$1,86 \times 10^{-22}$
cg12317815	17	3379283	ASPA	TSS200	N/A	$6,91 \times 10^{-12}$	$7,90 \times 10^{-15}$
cg10137837	17	6926742	BCL6B	5'UTR	Island	$5,93 \times 10^{-13}$	$4,75 \times 10^{-16}$
cg20273670	17	21356245	N/A	N/A	Island	$3,31 \times 10^{-13}$	$2,37 \times 10^{-16}$
cg11071401	17	48637194	CACNA1G	TSS1500	Island	$1,08 \times 10^{-15}$	$3,45 \times 10^{-19}$
cg19283806	18	66389420	CCDC102B	5'UTR	N/A	$4,83 \times 10^{-14}$	$2,94 \times 10^{-17}$
cg17471102	19	5851255	FUT3	5'UTR	N/A	$8,06 \times 10^{-13}$	$6,98 \times 10^{-16}$
cg07788092	19	9473691	ZNF177	TSS200	Island	$7,99 \times 10^{-13}$	$6,83 \times 10^{-16}$
cg14556683	19	15342982	EPHX3	1stExon	Island	$2,95 \times 10^{-23}$	$1,89 \times 10^{-27}$
cg26504305	19	19281019	LOC729991	Body	Island	$1,06 \times 10^{-13}$	$6,66 \times 10^{-17}$
cg05368740	19	36246882	HSPB6	Body	Island	$3,76 \times 10^{-14}$	$2,21 \times 10^{-17}$
cg00286125	19	36246906	HSPB6	Body	Island	$1,34 \times 10^{-14}$	$6,43 \times 10^{-18}$
cg17110586	19	36454623	N/A	N/A	S Shelf	$1,45 \times 10^{-15}$	$4,95 \times 10^{-19}$
cg26746936	19	42503392	GRIK5	Body	Island	$2,09 \times 10^{-12}$	$2,07 \times 10^{-15}$
cg04731544	19	49004834	LMTK3	Body	Island	$3,41 \times 10^{-13}$	$2,48 \times 10^{-16}$
cg05213896	19	50393653	IL4I1	Body	Island	$2,54 \times 10^{-17}$	$5,70 \times 10^{-21}$
cg06878361	19	52390810	ZNF577	Body	N Shore	$1,66 \times 10^{-14}$	$8,34 \times 10^{-18}$
cg03562414	19	52391078	ZNF577	Body	Island	$1,26 \times 10^{-14}$	$5,93 \times 10^{-18}$
cg10635122	19	52391090	ZNF577	Body	Island	$1,45 \times 10^{-15}$	$4,82 \times 10^{-19}$
cg16731240	19	52391250	ZNF577	TSS200	Island	$2,17 \times 10^{-12}$	$2,18 \times 10^{-15}$
cg23010048	19	52391257	ZNF577	TSS200	Island	$2,13 \times 10^{-14}$	$1,12 \times 10^{-17}$
cg11269599	19	52391304	ZNF577	TSS200	Island	$5,41 \times 10^{-13}$	$4,28 \times 10^{-16}$
cg22331349	19	52391350	ZNF577	TSS200	Island	$8,59 \times 10^{-13}$	$7,62 \times 10^{-16}$
cg09547119	19	52391367	ZNF577	TSS200	Island	$5,92 \times 10^{-15}$	$2,59 \times 10^{-18}$
cg08153621	19	53561441	N/A	N/A	Island	$1,42 \times 10^{-13}$	$9,55 \times 10^{-17}$
cg22682811	20	21372462	N/A	N/A	Island	$1,36 \times 10^{-13}$	$8,86 \times 10^{-17}$
cg07547549	20	44658225	SLC12A5	Body	Island	$4,47 \times 10^{-15}$	$1,77 \times 10^{-18}$

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position des analysierten CpG-Locus nach UCSC Genome Browser. q-Wert: p-Wert korrigiert nach Benjamini-Hochberg. p-Wert: t-Test für unabhängige Stichproben. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

Anhang

Tabelle A3: Krankheitsassoziations-Analyse der identifizierten altersabhängig methylierten Gene. Die assoziierten Erkrankungen sind entsprechend ihrer Signifikanz angeordnet.

Krankheits-Assoziation	beobachtet	erwartet	OR	p-Wert	q-Wert
Rigor (Muskelstarre)	3	0,13	23,33	$3,00 \times 10^{-04}$	0,022
Muskelhypertonie	2	0,04	47,85	$8,00 \times 10^{-04}$	0,022
Stimmungsschwankungen	5	0,74	6,76	0,001	0,022
Exfoliationssyndrom	2	0,05	41,47	0,001	0,022
Epilepsie	4	0,47	8,52	0,001	0,022
Uterusprolaps	2	0,05	36,59	0,001	0,022
Zwangsstörungen	3	0,23	12,96	0,002	0,022
Anfälle	4	0,52	7,73	0,002	0,022
Erkrankung des Nervensystems	7	1,76	3,99	0,002	0,022
Status Epilepticus	2	0,07	29,62	0,002	0,023
Kardiomyopathie NOS	3	0,28	10,73	0,003	0,026
Kardiomyopathien	3	0,28	10,73	0,003	0,026
Geistige Störung	6	1,50	4,00	0,004	0,032
Aspirin-induziertes Asthma	2	0,10	20,07	0,005	0,035
Erkrankungen des Zwölffingerdarms	2	0,12	16,81	0,006	0,045
Myoklonus (Schüttelkrampf)	2	0,12	16,37	0,007	0,045
Primitive neuroektodermale Tumoren	2	0,13	15,95	0,007	0,045

OR: *odds ratio*. p-Wert ermittelt über hypergeometrischenTest. q-Wert: mittels Benjamini-Hochberg-Methode korrigierter p-Wert.

Tabelle A4: Liste der mittels Infinium HumanMethylation450 BeadChip identifizierten 80 differentiell methylierten CpG Regionen ($p < 0,0003$; t-Test; $|\Delta\beta| > 0,05$) im Vergleich von adipösen ($n=5$) und normalgewichtigen ($n=29$) Zwillingen. Die CpG-Loci sind nach aufsteigender Position auf den jeweiligen Chromosomen geordnet.

CpG ID	Chr	Position (hg19)	Gen	Gen - Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (adipös – normal)	Mittlere Methylierung [β] adipös	Mittlere Methylierung [β] normal	q-Wert	p-Wert
cg11019172	1	6296721	ICMT	TSS1500	S Shore	-0,158	0,674	0,832	0,91	$6,73 \times 10^{-05}$
cg11019172	1	6296721	ICMT	TSS1500	S Shore	-0,070	0,786	0,856	0,93	$2,16 \times 10^{-04}$
cg03812309	1	46938281	N/A	N/A	N Shore	-0,058	0,344	0,401	0,72	$2,22 \times 10^{-05}$
cg00597687	1	63785003	N/A	N/A	Island	-0,056	0,262	0,318	0,93	$2,15 \times 10^{-04}$
cg05160197	1	203135890	ADORA1	3'UTR	N/A	0,070	0,140	0,070	0,91	$7,01 \times 10^{-05}$
cg13064658	1	212003989	LPGAT1	5'UTR	Island	-0,128	0,345	0,473	0,93	$2,95 \times 10^{-04}$
cg11518608	1	214782434	CENPF	5'UTR	Island	-0,156	0,710	0,865	0,82	$3,42 \times 10^{-05}$
cg22796025	1	227509201	N/A	N/A	S Shelf	-0,193	0,718	0,911	0,93	$1,82 \times 10^{-04}$
cg21665850	2	731073	N/A	N/A	N Shore	-0,234	0,604	0,838	0,92	$1,36 \times 10^{-04}$
cg01267068	2	731158	N/A	N/A	N Shore	-0,068	0,534	0,603	0,92	$1,32 \times 10^{-04}$
cg16437385	2	74943219	N/A	N/A	S Shore	-0,052	0,212	0,264	0,93	$2,00 \times 10^{-04}$
cg27560561	2	105484823	N/A	N/A	S Shore	-0,183	0,333	0,516	0,92	$1,30 \times 10^{-04}$
cg23337647	2	114359736	MGC13005	TSS1500	N/A	-0,064	0,809	0,873	0,92	$1,37 \times 10^{-04}$
cg13523468	3	12047002	SYN2	Body	S Shore	-0,080	0,328	0,408	0,92	$1,15 \times 10^{-04}$
cg04908625	3	123166882	ADCY5	1stExon	Island	-0,132	0,810	0,942	0,91	$8,21 \times 10^{-05}$
cg00796374	3	155748954	N/A	N/A	N/A	-0,074	0,711	0,785	0,62	$1,64 \times 10^{-05}$
cg24564096	4	7449698	SORCS2	Body	N/A	-0,051	0,907	0,958	0,93	$1,46 \times 10^{-04}$
cg24797456	4	76622044	N/A	N/A	N/A	-0,064	0,798	0,861	0,93	$2,34 \times 10^{-04}$
cg00737846	4	183214573	N/A	N/A	N/A	-0,161	0,653	0,814	0,91	$8,07 \times 10^{-05}$
cg14655345	4	185659507	N/A	N/A	S Shelf	-0,065	0,813	0,878	0,92	$1,17 \times 10^{-04}$
cg02109704	5	1721452	N/A	N/A	N Shelf	-0,087	0,775	0,862	0,93	$2,81 \times 10^{-04}$
cg15813594	5	38445563	EGFLAM	TSS200	N/A	-0,129	0,697	0,826	0,48	$7,37 \times 10^{-06}$
cg17382841	5	38445614	EGFLAM	TSS200	N/A	-0,176	0,213	0,389	0,93	$2,84 \times 10^{-04}$
cg25259944	5	110062473	TMEM232	TSS200	N/A	-0,197	0,301	0,497	0,93	$1,72 \times 10^{-04}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (adipös – normal)	Mittlere Methylierung [β] adipös	Mittlere Methylierung [β] normal	q-Wert	p-Wert
cg06414816	5	110062837	TMEM232	TSS1500	N/A	-0,153	0,433	0,587	0,93	$2,10 \times 10^{-04}$
cg11182018	5	140474062	PCDHB2	TSS200	N Shore	-0,379	0,409	0,788	0,11	$9,65 \times 10^{-07}$
cg17857094	6	30907280	DPCR1	TSS1500	N/A	-0,104	0,787	0,891	0,93	$2,85 \times 10^{-04}$
cg25857471	6	30910113	DPCR1	Body	N/A	-0,058	0,834	0,893	0,92	$1,03 \times 10^{-04}$
cg20161227	6	31980836	TNXB	5'UTR	N/A	-0,179	0,783	0,962	0,06	$4,12 \times 10^{-07}$
cg03849834	6	41195891	TREM14	TSS200	N/A	-0,076	0,814	0,889	0,86	$4,93 \times 10^{-05}$
cg21020855	6	123100189	FABP7	TSS1500	N/A	-0,057	0,740	0,798	0,92	$1,16 \times 10^{-04}$
cg02451831	7	26578098	KIAA0087	Body	N/A	-0,067	0,691	0,758	0,61	$1,35 \times 10^{-05}$
cg22664425	7	51226655	COBL	Body	N/A	-0,058	0,841	0,899	0,93	$2,42 \times 10^{-04}$
cg14449317	7	62198343	N/A	N/A	N/A	0,098	0,125	0,027	0,93	$2,15 \times 10^{-04}$
cg11791427	7	73704194	CLIP2	5'UTR	S Shore	-0,052	0,909	0,961	0,06	$3,47 \times 10^{-07}$
cg10351692	7	75053196	POM121C	Body	S Shore	-0,060	0,814	0,875	0,93	$1,74 \times 10^{-04}$
cg04290826	7	157929653	PTPRN2	Body	N/A	-0,157	0,807	0,964	0,93	$2,30 \times 10^{-04}$
cg03377073	7	157980803	PTPRN2	Body	Island	-0,071	0,828	0,900	0,93	$2,87 \times 10^{-04}$
cg06934974	7	158294666	PTPRN2	Body	N/A	-0,084	0,764	0,848	0,93	$2,67 \times 10^{-04}$
cg06238661	8	136409523	N/A	N/A	N/A	-0,062	0,852	0,915	0,85	$4,54 \times 10^{-05}$
cg02395917	8	139613712	COL22A1	Body	N/A	-0,080	0,781	0,862	0,93	$1,66 \times 10^{-04}$
cg21782089	8	143214455	N/A	N/A	N/A	-0,066	0,851	0,917	0,93	$2,17 \times 10^{-04}$
cg03946168	10	2357461	N/A	N/A	Island	-0,143	0,638	0,781	0,93	$2,91 \times 10^{-04}$
cg18912160	10	3281074	N/A	N/A	N/A	-0,078	0,859	0,937	0,92	$9,75 \times 10^{-05}$
cg06473608	10	134630642	N/A	N/A	N/A	-0,204	0,664	0,869	0,93	$2,29 \times 10^{-04}$
cg23693569	10	134724823	N/A	N/A	S Shelf	-0,052	0,787	0,840	0,93	$1,73 \times 10^{-04}$
cg20813240	11	1717973	HCCA2	Body	S Shore	-0,169	0,624	0,793	0,85	$4,24 \times 10^{-05}$
cg13170468	11	10531665	N/A	N/A	N/A	-0,065	0,855	0,919	0,93	$2,85 \times 10^{-04}$
cg27475923	11	63530997	C11orf95	Body	Island	-0,144	0,746	0,890	0,61	$1,41 \times 10^{-05}$
cg14759565	11	65360123	N/A	N/A	Island	-0,094	0,719	0,812	0,92	$9,29 \times 10^{-05}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (adipös – normal)	Mittlere Methylierung [β] adipös	Mittlere Methylierung [β] normal	q-Wert	p-Wert
cg13654525	11	65360509	KCNK7	Body	S Shore	-0,096	0,489	0,585	0,91	$8,90 \times 10^{-05}$
cg10054980	11	89477675	N/A	N/A	N/A	-0,076	0,675	0,752	0,93	$1,82 \times 10^{-04}$
cg12018163	11	119456270	N/A	N/A	S Shore	-0,072	0,859	0,931	0,85	$4,55 \times 10^{-05}$
cg16490778	11	120038654	N/A	N/A	N Shore	-0,058	0,907	0,965	0,91	$5,56 \times 10^{-05}$
cg18057887	11	133800913	IGSF9B	Body	Island	-0,061	0,858	0,919	0,93	$2,93 \times 10^{-04}$
cg01161055	12	32437813	BICD1	Body	N/A	-0,140	0,831	0,971	0,93	$2,15 \times 10^{-04}$
cg14751369	12	117042332	N/A	N/A	N/A	0,137	0,169	0,032	0,92	$1,21 \times 10^{-04}$
cg02291365	13	25317267	N/A	N/A	N Shelf	-0,068	0,784	0,852	0,93	$2,03 \times 10^{-04}$
cg09415720	13	76244804	LMO7	Body	N/A	-0,171	0,567	0,738	0,93	$2,36 \times 10^{-04}$
cg04646842	13	94107224	GPC6	Body	N/A	-0,097	0,534	0,630	0,91	$7,69 \times 10^{-05}$
cg09970967	14	73602752	PSEN1	TSS1500	N Shore	-0,068	0,846	0,914	0,16	$1,79 \times 10^{-06}$
cg21561672	14	102101551	N/A	N/A	Island	-0,059	0,854	0,913	0,93	$2,72 \times 10^{-04}$
cg07685326	15	25380111	PAR1	TSS1500	N/A	-0,058	0,857	0,916	0,93	$2,84 \times 10^{-04}$
cg21404878	15	72578605	BRUNOL6	3'UTR	N Shelf	-0,147	0,663	0,810	0,93	$2,69 \times 10^{-04}$
cg08288433	17	20799511	CCDC144NL	TSS200	Island	-0,079	0,406	0,485	0,93	$2,33 \times 10^{-04}$
cg11633473	17	20799714	CCDC144NL	TSS1500	Island	-0,094	0,258	0,352	0,65	$1,87 \times 10^{-05}$
cg21692620	17	40835849	CNTNAP1	Body	Island	-0,098	0,441	0,540	0,91	$6,59 \times 10^{-05}$
cg09776463	17	40836087	CNTNAP1	Body	Island	-0,088	0,621	0,709	0,92	$1,29 \times 10^{-04}$
cg03989507	17	40836125	CNTNAP1	Body	Island	-0,061	0,502	0,563	0,93	$1,95 \times 10^{-04}$
cg22434594	17	47072339	N/A	N/A	N Shore	-0,062	0,450	0,512	0,93	$2,81 \times 10^{-04}$
cg14004678	17	77001516	CANT1	5'UTR	N/A	0,092	0,238	0,146	0,82	$3,41 \times 10^{-05}$
cg16779839	17	79374327	BAHCC1	Body	Island	-0,056	0,661	0,717	0,93	$2,43 \times 10^{-04}$
cg18415772	19	1503042	N/A	N/A	Island	-0,071	0,361	0,431	0,92	$1,40 \times 10^{-04}$
cg00706441	19	12958340	MAST1	Body	Island	-0,063	0,543	0,606	0,92	$1,23 \times 10^{-04}$
cg23055064	19	49235655	RASIP1	Body	N Shelf	-0,113	0,594	0,707	0,91	$9,01 \times 10^{-05}$
cg08113002	19	51017947	ASPDH	TSS1500	N Shelf	-0,051	0,850	0,901	0,93	$1,95 \times 10^{-04}$

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (adipös – normal)	Mittlere Methylierung [β] adipös	Mittlere Methylierung [β] normal	q-Wert	p-Wert
cg11128808	19	54495970	CACNG6	1stExon	N Shore	0,110	0,291	0,181	0,47	$6,21 \times 10^{-06}$
cg07455279	19	54605703	NDUFA3	TSS1500	N Shore	-0,118	0,559	0,676	0,93	$1,88 \times 10^{-04}$
cg08927738	20	52687394	BCAS1	TSS200	N/A	-0,151	0,800	0,951	0,92	$1,27 \times 10^{-04}$
cg14723284	21	46370511	C21orf70	Body	S Shelf	-0,091	0,826	0,917	0,92	$1,41 \times 10^{-04}$
cg23702117	22	31897078	SF11	5'UTR	S Shelf	-0,070	0,786	0,856	0,93	$2,16 \times 10^{-04}$

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position des analysierten CpG-Locus nach UCSC Genome Browser. q-Wert: p-Wert korrigiert nach Benjamini-Hochberg. p-Wert: t-Test für unabhängige Stichproben. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

Tabelle A5: Liste der mittels Infinium HumanMethylation450 BeadChip identifizierten 131 differentiell methylierten CpG Regionen ($p < 0,0004$; t-Test; $|\Delta\beta| > 0,05$, $\sigma/\sigma_{\max} > 0,1$) im Vergleich von Rauchern (n=12) und Nichtrauchern (n=25). Die CpG-Loci sind nach aufsteigender Position auf den jeweiligen Chromosomen geordnet.

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert	p-Wert
cg04121631	1	875345	SAMD11	Body	N Shore	-0,055	0,112	0,167	0,07	$1,42 \times 10^{-05}$
cg05776630	1	20015027	TMCO4	Body	N/A	0,095	0,399	0,303	0,14	$6,40 \times 10^{-05}$
cg24253618	1	32054460	N/A	N/A	Island	-0,096	0,509	0,605	0,23	$2,82 \times 10^{-04}$
cg13935394	1	65733223	DNAJC6	Body	S Shore	-0,062	0,688	0,750	0,10	$2,90 \times 10^{-05}$
cg10827434	1	84944976	RPF1	1stExon	Island	-0,061	0,182	0,243	0,14	$6,41 \times 10^{-05}$
cg02345146	1	85512957	MCOLN3	5'UTR	N Shore	-0,096	0,577	0,673	0,26	$3,67 \times 10^{-04}$
cg11161837	1	110010384	SYPL2	Body	S Shore	-0,059	0,606	0,665	0,08	$2,04 \times 10^{-05}$
cg06936779	1	151171405	PIP5K1A	5'UTR	Island	-0,066	0,303	0,369	0,18	$1,18 \times 10^{-04}$
cg27310092	1	151945663	N/A	N/A	N/A	-0,089	0,345	0,434	0,16	$9,36 \times 10^{-05}$
cg27509823	1	173930183	RC3H1	Body	N/A	-0,070	0,784	0,854	0,18	$1,35 \times 10^{-04}$
cg07476963	1	174127462	RABGAP1L	TSS1500	N Shore	-0,059	0,677	0,736	0,23	$2,82 \times 10^{-04}$
cg15915602	2	114737459	N/A	N/A	Island	-0,084	0,555	0,639	0,18	$1,21 \times 10^{-04}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert	p-Wert
cg03225444	2	129494558	N/A	N/A	Island	-0,062	0,547	0,609	0,14	$5,93 \times 10^{-05}$
cg24964215	2	161826714	N/A	N/A	N/A	-0,084	0,237	0,321	0,16	$9,42 \times 10^{-05}$
cg27241845	2	233250370	N/A	N/A	N Shore	-0,067	0,664	0,731	0,07	$1,66 \times 10^{-05}$
cg03329539	2	233283329	N/A	N/A	N Shore	-0,060	0,347	0,407	$2,32 \times 10^{-04}$	$1,32 \times 10^{-08}$
cg05951221	2	233284402	N/A	N/A	Island	-0,154	0,305	0,459	$1,62 \times 10^{-05}$	$7,38 \times 10^{-10}$
cg21566642	2	233284661	N/A	N/A	Island	-0,186	0,294	0,480	$4,04 \times 10^{-07}$	$1,38 \times 10^{-11}$
cg01940273	2	233284934	N/A	N/A	Island	-0,160	0,544	0,704	0,19	$1,72 \times 10^{-04}$
cg03568507	2	240153791	HDAC4	Body	Island	0,125	0,635	0,510	0,14	$6,62 \times 10^{-05}$
cg12547218	2	240153820	HDAC4	Body	S Shore	0,118	0,592	0,473	0,03	$3,56 \times 10^{-06}$
cg12207248	3	59038598	N/A	N/A	S Shelf	-0,076	0,396	0,471	0,26	$3,86 \times 10^{-04}$
cg27192606	3	61728802	PTPRG	Body	N/A	0,078	0,668	0,590	0,10	$2,94 \times 10^{-05}$
cg08439880	3	133502540	N/A	N/A	N Shore	-0,147	0,725	0,872	0,21	$2,04 \times 10^{-04}$
cg08991643	3	194014928	N/A	N/A	Island	-0,092	0,648	0,740	0,18	$1,40 \times 10^{-04}$
cg12155036	3	194015171	N/A	N/A	S Shore	-0,074	0,548	0,622	0,26	$3,93 \times 10^{-04}$
cg05834845	3	195489306	MUC4	Body	Island	0,267	0,848	0,581	0,23	$2,65 \times 10^{-04}$
cg15305511	4	5712676	EVC2	TSS1500	N Shore	-0,054	0,320	0,375	0,26	$3,33 \times 10^{-04}$
cg19786733	4	39449130	KLB	Body	S Shore	-0,056	0,244	0,301	0,14	$6,83 \times 10^{-05}$
cg07321536	4	39459897	LIAS	TSS1500	N Shore	-0,084	0,261	0,345	0,18	$1,25 \times 10^{-04}$
cg26383193	4	47488375	ATP10D	5'UTR	N/A	-0,069	0,736	0,805	0,06	$1,25 \times 10^{-05}$
cg24814707	4	57035957	KIAA1211	TSS1500	N/A	-0,063	0,842	0,905	0,18	$1,48 \times 10^{-04}$
cg11939399	4	57773300	REST	TSS1500	N Shore	-0,067	0,223	0,290	0,14	$5,67 \times 10^{-05}$
cg02214188	4	104020949	BDH2	1stExon	N/A	-0,070	0,157	0,227	0,16	$8,27 \times 10^{-05}$
cg20073050	4	106395643	PPA2	TSS1500	S Shore	-0,052	0,394	0,447	0,14	$4,58 \times 10^{-05}$
cg15810415	4	111397134	ENPEP	TSS200	N Shore	-0,074	0,728	0,802	0,23	$2,81 \times 10^{-04}$
cg20773127	4	111397155	ENPEP	TSS200	N Shore	-0,067	0,625	0,692	0,23	$2,67 \times 10^{-04}$
cg04760493	4	128981486	LARP1B	TSS1500	N Shore	-0,057	0,767	0,824	0,23	$2,62 \times 10^{-04}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert	p-Wert
cg17310600	4	165305334	MARCH1	TSS1500	S Shore	-0,059	0,405	0,463	0,21	$2,24 \times 10^{-04}$
cg17837517	4	189541174	N/A	N/A	N/A	-0,108	0,292	0,401	0,18	$1,42 \times 10^{-04}$
cg17924476	5	323794	AHRR	Body	S Shore	0,066	0,505	0,440	0,12	$3,66 \times 10^{-05}$
cg01899089	5	369969	AHRR	Body	N Shore	-0,058	0,488	0,546	$1,51 \times 10^{-11}$	$1,72 \times 10^{-16}$
cg05575921	5	373378	AHRR	Body	N Shore	-0,258	0,549	0,807	0,25	$3,13 \times 10^{-04}$
cg05575921	5	373378	AHRR	Body	N Shore	-0,078	0,629	0,707	$2,45 \times 10^{-04}$	$1,68 \times 10^{-08}$
cg26703534	5	377358	AHRR	Body	S Shelf	-0,084	0,201	0,285	$4,36 \times 10^{-03}$	$3,95 \times 10^{-07}$
cg14817490	5	392920	AHRR	Body	N/A	-0,071	0,790	0,861	0,02	$2,22 \times 10^{-06}$
cg25648203	5	395444	AHRR	Body	N/A	-0,050	0,808	0,858	0,18	$1,19 \times 10^{-04}$
cg24090911	5	400732	AHRR	Body	N/A	-0,092	0,297	0,388	0,10	$2,96 \times 10^{-05}$
cg16578742	5	35617687	SPEF2	TSS1500	N Shore	-0,051	0,491	0,542	0,19	$1,55 \times 10^{-04}$
cg00712841	5	37835548	GDNF	5'UTR	N Shore	-0,084	0,690	0,774	0,14	$5,84 \times 10^{-05}$
cg08952407	5	42756860	CCDC152	TSS200	N/A	-0,069	0,712	0,780	0,18	$1,48 \times 10^{-04}$
cg13006681	5	158476743	EBF1	Body	N Shore	-0,062	0,316	0,379	0,21	$2,24 \times 10^{-04}$
cg26775866	5	159849193	PTTG1	5'UTR	S Shore	-0,052	0,207	0,258	0,14	$7,08 \times 10^{-05}$
cg10199913	6	26045987	HIST1H3C	1stExon	Island	-0,057	0,298	0,355	0,19	$1,75 \times 10^{-04}$
cg12828313	6	28744086	N/A	N/A	N/A	-0,146	0,635	0,780	$1,56 \times 10^{-03}$	$1,25 \times 10^{-07}$
cg06126421	6	30720080	N/A	N/A	N/A	-0,105	0,357	0,463	0,14	$5,31 \times 10^{-05}$
cg14753356	6	30720108	N/A	N/A	N/A	-0,074	0,625	0,699	0,14	$5,52 \times 10^{-05}$
cg22880122	6	31027005	HCG22	Body	N/A	-0,062	0,870	0,932	0,06	$1,02 \times 10^{-05}$
cg02956248	6	32120901	PPT2	TSS1500	N Shore	-0,105	0,470	0,576	0,21	$2,24 \times 10^{-04}$
cg05133205	6	32121249	PPT2	TSS200	N Shore	-0,064	0,782	0,847	0,18	$1,44 \times 10^{-04}$
cg13598232	6	33284748	ZBTB22	5'UTR	N Shore	-0,060	0,263	0,322	0,26	$3,75 \times 10^{-04}$
cg11348249	6	151560831	AKAP12	TSS1500	Island	0,071	0,860	0,789	0,26	$3,68 \times 10^{-04}$
cg12854815	6	160024144	N/A	N/A	N/A	-0,055	0,819	0,874	0,26	$3,83 \times 10^{-04}$
cg11263074	6	167727700	UNC93A	Body	N/A	0,133	0,877	0,743	0,25	$3,07 \times 10^{-04}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert	p-Wert
cg08877521	6	170453256	N/A	N/A	N Shore	0,068	0,222	0,154	0,21	$2,09 \times 10^{-04}$
cg04852348	6	170463167	N/A	N/A	N/A	-0,050	0,819	0,869	0,26	$3,87 \times 10^{-04}$
cg27003309	7	5343752	N/A	N/A	Island	-0,059	0,287	0,346	0,16	$8,63 \times 10^{-05}$
cg13604553	7	34361111	N/A	N/A	N/A	-0,098	0,470	0,567	0,18	$1,40 \times 10^{-04}$
cg16652817	7	54956598	N/A	N/A	Island	-0,136	0,778	0,915	0,18	$1,33 \times 10^{-04}$
cg02145213	7	76609946	PMS2L11	TSS200	N/A	-0,060	0,785	0,845	0,18	$1,44 \times 10^{-04}$
cg10443315	8	7947282	MIR54813	TSS1500	N/A	-0,110	0,300	0,409	0,22	$2,46 \times 10^{-04}$
cg16810626	8	11594003	GATA4	Body	N/A	-0,071	0,842	0,913	0,19	$1,54 \times 10^{-04}$
cg23303108	8	23083578	TNFRSF10A	TSS1500	S Shore	0,076	0,821	0,745	0,14	$6,08 \times 10^{-05}$
cg26126879	8	103916295	N/A	N/A	N/A	-0,127	0,322	0,450	0,16	$9,00 \times 10^{-05}$
ch.8.2353618R	8	119282796	SAMD12	Body	N/A	-0,093	0,325	0,417	0,18	$1,14 \times 10^{-04}$
cg01692968	9	108005349	N/A	N/A	N Shore	-0,072	0,461	0,533	0,21	$2,19 \times 10^{-04}$
cg13418576	9	126105738	N/A	N/A	S Shelf	-0,058	0,718	0,776	0,07	$1,63 \times 10^{-05}$
cg03068346	10	3113919	PFKF	Body	S Shelf	-0,070	0,174	0,244	0,23	$2,64 \times 10^{-04}$
cg00524374	10	3894968	N/A	N/A	N/A	-0,088	0,272	0,360	0,26	$3,41 \times 10^{-04}$
cg26229043	10	24407477	KIAA1217	5'UTR	N/A	-0,059	0,804	0,863	0,12	$3,85 \times 10^{-05}$
cg19495650	10	65226682	JMJD1C	TSS1500	S Shore	-0,081	0,525	0,606	0,22	$2,34 \times 10^{-04}$
cg17107540	10	72197470	NODAL	Body	N Shelf	-0,177	0,556	0,733	0,05	$8,04 \times 10^{-06}$
cg19336497	11	14380999	RRAS2	TSS1500	S Shore	-0,098	0,741	0,839	0,22	$2,50 \times 10^{-04}$
cg03274456	11	62273116	AHNAK	Body	N/A	-0,072	0,516	0,588	0,21	$2,18 \times 10^{-04}$
cg11660018	11	86510915	PRSS23	TSS1500	N Shore	-0,113	0,445	0,559	0,14	$4,87 \times 10^{-05}$
cg10635895	12	2339439	CACNA1C	Body	Island	-0,064	0,530	0,594	0,26	$3,98 \times 10^{-04}$
cg23390118	12	2339513	CACNA1C	Body	Island	-0,089	0,433	0,522	0,19	$1,69 \times 10^{-04}$
cg09576209	12	2339614	CACNA1C	Body	Island	-0,087	0,326	0,413	0,26	$3,54 \times 10^{-04}$
cg15584813	12	47219626	SLC38A4	5'UTR	N/A	-0,118	0,308	0,426	0,22	$2,42 \times 10^{-04}$
cg04053798	12	47219705	SLC38A4	1stExon	N/A	-0,189	0,305	0,495	0,20	$1,93 \times 10^{-04}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert	p-Wert
cg11146691	12	47219737	SLC38A4	1stExon	N/A	-0,138	0,475	0,614	0,19	$1,60 \times 10^{-04}$
cg21428710	12	47219797	SLC38A4	TSS200	N/A	-0,113	0,211	0,323	0,26	$3,60 \times 10^{-04}$
cg22049899	12	47219821	SLC38A4	TSS200	N/A	-0,153	0,332	0,485	0,19	$1,59 \times 10^{-04}$
cg07601320	12	47219841	SLC38A4	TSS200	N/A	-0,128	0,429	0,556	0,03	$3,98 \times 10^{-06}$
cg14671384	12	47219920	SLC38A4	TSS200	N/A	-0,078	0,631	0,709	0,19	$1,75 \times 10^{-04}$
cg18064171	12	47219957	SLC38A4	TSS200	N/A	-0,061	0,335	0,396	0,05	$8,51 \times 10^{-06}$
cg11146034	12	96590670	ELK3	5'UTR	S Shore	-0,080	0,631	0,710	0,16	$9,44 \times 10^{-05}$
cg21518287	13	51744561	N/A	N/A	N/A	-0,054	0,724	0,778	0,16	$8,70 \times 10^{-05}$
cg05078075	13	76250138	LMO7	Body	N/A	-0,070	0,469	0,539	0,18	$1,18 \times 10^{-04}$
cg05600496	13	80357187	N/A	N/A	N/A	-0,054	0,769	0,823	0,25	$3,25 \times 10^{-04}$
cg26923084	14	52536893	NID2	TSS1500	S Shore	-0,052	0,808	0,859	0,18	$1,27 \times 10^{-04}$
cg19512582	14	61083516	N/A	N/A	N/A	-0,057	0,495	0,552	0,18	$1,21 \times 10^{-04}$
cg13038618	14	77467391	N/A	N/A	N/A	-0,121	0,330	0,451	0,26	$3,42 \times 10^{-04}$
cg13521842	14	101036202	BEGAIN	TSS200	S Shore	-0,089	0,486	0,574	0,25	$3,06 \times 10^{-04}$
cg21193577	14	101036322	BEGAIN	TSS200	S Shore	-0,068	0,776	0,845	0,20	$1,86 \times 10^{-04}$
cg14879065	14	104662468	N/A	N/A	N/A	-0,059	0,839	0,898	0,26	$3,84 \times 10^{-04}$
cg13213799	14	104834222	N/A	N/A	N/A	-0,076	0,562	0,638	0,18	$1,37 \times 10^{-04}$
cg16540262	15	72883661	N/A	N/A	N/A	0,055	0,677	0,622	0,20	$1,91 \times 10^{-04}$
cg22734451	15	78219748	LOC645752	TSS1500	N/A	-0,068	0,214	0,282	0,04	$5,89 \times 10^{-06}$
cg23161492	15	90357202	ANPEP	5'UTR	N Shore	-0,055	0,767	0,823	0,18	$1,14 \times 10^{-04}$
cg09866598	16	4361610	N/A	N/A	N Shelf	-0,066	0,725	0,791	0,14	$4,94 \times 10^{-05}$
cg04890713	16	58022504	N/A	N/A	S Shelf	-0,081	0,740	0,822	0,14	$7,00 \times 10^{-05}$
cg05828191	16	75019382	WDR59	TSS1500	S Shore	0,081	0,909	0,828	0,16	$8,47 \times 10^{-05}$
cg00837290	16	89350068	ANKRD11	Body	S Shore	0,063	0,235	0,172	0,23	$2,81 \times 10^{-04}$
cg08726900	16	89550474	ANKRD11	5'UTR	N/A	0,090	0,600	0,510	0,23	$2,81 \times 10^{-04}$
cg20021790	17	181288	RPH3AL	5'UTR	N Shelf	0,087	0,134	0,047	0,26	$3,70 \times 10^{-04}$

Anhang

CpG ID	Chr	Position (hg19)	Gen	Gen-Region	CpG-Lokalisation	Mittlere Methylierungsdifferenz [$\Delta\beta$] (Raucher – Nichtraucher)	Mittlere Methylierung [β] Raucher	Mittlere Methylierung [β] Nichtraucher	q-Wert	p-Wert
cg18633784	17	18965580	N/A	N/A	Island	-0,063	0,595	0,658	0,21	$2,11 \times 10^{-04}$
cg05341260	17	37730450	N/A	N/A	S Shore	-0,058	0,849	0,907	0,12	$3,94 \times 10^{-05}$
cg10408284	17	48912327	WFIKK2	TSS1500	N/A	-0,066	0,799	0,865	0,20	$1,97 \times 10^{-04}$
cg07168556	17	48912483	WFIKK2	TSS200	N/A	-0,072	0,790	0,863	0,26	$3,31 \times 10^{-04}$
cg00003625	17	48912543	WFIKK2	TSS200	N/A	-0,081	0,642	0,723	0,26	$3,88 \times 10^{-04}$
cg00475235	17	54236118	ANKFN1	Body	N/A	-0,070	0,768	0,838	0,26	$3,77 \times 10^{-04}$
cg10377414	17	80832836	TBCD	Body	S Shelf	-0,068	0,331	0,400	0,21	$2,08 \times 10^{-04}$
cg23345004	18	44526430	KATNAL2	TSS1500	N Shore	-0,059	0,342	0,401	0,19	$1,61 \times 10^{-04}$
cg00073090	19	1265879	N/A	N/A	Island	-0,143	0,623	0,766	$6,50 \times 10^{-09}$	$1,48 \times 10^{-13}$
cg03636183	19	17000585	F2RL3	Body	N Shore	0,078	0,622	0,544	0,20	$1,96 \times 10^{-04}$
cg26377653	19	54191706	MIR1283-1	TSS200	N/A	-0,062	0,184	0,245	0,03	$4,56 \times 10^{-06}$
cg13972460	20	1165703	C20orf46	TSS1500	Island	0,051	0,612	0,561	0,26	$3,69 \times 10^{-04}$
cg26929161	20	50701266	ZFP64	Body	Island	-0,091	0,534	0,625	0,26	$3,62 \times 10^{-04}$
cg24322131	21	45971503	KRTAP10-2	TSS200	N/A	0,101	0,470	0,369	0,08	$1,90 \times 10^{-05}$
cg07018760	21	46074616	KRTAP12-4	TSS200	N/A	-0,112	0,382	0,494	0,14	$5,94 \times 10^{-05}$
cg07519229	22	41486835	N/A	N/A	N Shore	-0,055	0,112	0,167	0,25	$3,13 \times 10^{-04}$

CpG ID: CpG-Locus Identifier. Chr: Chromosom. Position (hg19): Genomische Position des analysierten CpG-Locus nach UCSC Genome Browser. q-Wert: p-Wert korrigiert nach Benjamini-Hochberg. p-Wert: t-Test für unabhängige Stichproben. N/A: Nach UCSC Genome Browser konnte keine entsprechende Annotation mit dieser CpG-Position assoziiert werden.

10 Danksagung

Ich möchte mich gern bei allen Menschen und Freunden bedanken, die mich während dieser Arbeit direkt oder indirekt unterstützt haben.

Besonderer Dank gebührt meinem Betreuer, Prof. Dr. rer. nat. Ole Ammerpohl, für die Bereitstellung des interessanten Themas, die freundliche Betreuung und uneingeschränkte Unterstützung sowie für das Extrapaket an Zeit und Geduld gerade in der Endphase dieser Arbeit.

Bei Frau Prof. Dr. rer. nat. Manuela Dittmar möchte ich mich herzlich für die Übernahme der Betreuung und des Referates meiner Arbeit vor der Mathematisch-Naturwissenschaftlichen Fakultät der Christian-Albrechts-Universität zu Kiel bedanken.

Großer Dank gilt ebenfalls dem Direktor des Institutes für Humangenetik, Prof. Dr. med. Reiner Siebert, für die Aufnahme und Möglichkeit, diese Arbeit anzufertigen sowie für die wertvolle Unterstützung und das mir entgegengebrachte Vertrauen.

Des Weiteren danke ich allen Kooperationspartnern des SAME-Projektes für die konstruktive und angenehme Zusammenarbeit. Besonderer Dank gilt Frau Dr. rer. nat. Denise Olbrich für die Auswertung und Übertragung der Fragebögen sowie deren bereitwillige Weitergabe.

Weiterhin gilt mein Dank allen Mitdoktoranden und Mitarbeitern am Institut für Humangenetik für die gute Zusammenarbeit, die freundliche Atmosphäre sowie für die mir jederzeit entgegengebrachte Hilfsbereitschaft. Insbesondere möchte ich mich bei Frau Ute Jacobsen für die Durchführung der Pyrosequenzierungen bedanken.

Bei Herrn Prof Dr. rer. nat. Ulrich Lehmann und seiner Arbeitsgruppe von der Medizinischen Hochschule Hannover möchte ich mich für die angenehme und konstruktive Zusammenarbeit bedanken, aus der eine Publikation hervorging, die Bestandteil dieser Dissertation ist.

Schließlich danke ich ganz herzlich meiner Familie für ihre liebevolle Unterstützung und ihren Glauben an mich.

Besonderer Dank gilt meinem Verlobten Olaf für seine unendliche Geduld und Liebe.

11 Publikationsliste

Publizierte Originalarbeiten

Roessler J, Ammerpohl O, **Gutwein J**, Hasemeier B, Anwar SL, Kreipe H, Lehmann U. Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. BMC Res Notes. 2012;5:210.

Ammerpohl O, **Gutwein J**, Siebert R. Epigenetik: Die Brücke zwischen Umwelt und Genom. umwelt medizin gesellschaft 2013; 26(4): 263–269.

Friemel C, Ammerpohl O, **Gutwein J**, Schmutzler AG, Caliebe A, Kautza M, von Otte S, Siebert R, Bens S. Array-based DNA methylation profiling in male infertility reveals allele-specific DNA methylation in PIWIL1 and PIWIL2. Fertil Steril. 2014;101(4):1097–1103.

Roessler J, Ammerpohl O, **Gutwein J**, Steinemann D, Schlegelberger B, Weyer V, Sariyar M, Geffers R, Arnold N, Schmutzler R, Bartram CR, Heinrich T, Abbas M, Antonopoulos W, Schipper E, Hasemeier B, Kreipe H, Lehmann U. The CpG island methylator phenotype in breast cancer is associated with the lobular subtype. Epigenomics. 2014:1–13.

Eingereichte Originalarbeiten und Letter

Vater I, Montesinos-Rongen M, Schlesner M, Haake A, Sprute R, Mettenmeyer N, Purschke F, Nazzari I, Nagel I, **Gutwein J**, Richter J, Buchhalter I, Eils R, Deckert M, Siebert R. The mutational landscape of primary lymphoma of the central nervous system determined by whole exome sequencing. Acta Neuropathol. 2014.

Kolarova J, Ammerpohl O, **Gutwein J**, Welzel M, Baus I, Riepe FG, Eggermann T, Caliebe A, Holterhus PM, Siebert R, Bens S. In vivo investigations of the effect of short- and long-term recombinant growth hormone treatment on DNA-methylation in humans. PLoS One. 2014.

Kretzmer H, Bernhart SH, Wang W, Haake A, Weniger MA, Bergmann AK, Betts MJ, Carillo6 E, Doose G, **Gutwein J**, Richter J, Hovestadt V, Huang B, Jühling F, Kolarova J, Lu Q, Otto C, Wagener R, Arnolds J, Burkhardt B, Claviez A, Drexler HG, Eberth S, Eils R, Flicek P, Haas S, Hummel M, Karsch D, Kerstens HHD, Klapper W, Kreuz M, Lawerenz C, Lenze D, Loeffler M, López C, MacLeod RAF, Martens JHA, Martín-Subero JI, Möller P, Nagel I, Picelli S, Vater I, Rohde M, Rosenstiel P, Rosolowski M, Russel RB, Schilhabel M, Schlesner M, Stadler PF,

Szczepanowski M, Trümper L, Stunnenberg HG, Küppers R, Ammerpohl O, Lichter P, Siebert R, Hoffmann S, Radlwimmer B. Linking differential DNA methylation in gene regions with the mutational landscape of Burkitt lymphoma. *Nat Genet.* 2014.

Kolarova J, Tangen I, Bens S, Gillessen-Kaesbach G, **Gutwein J**, Kautza M, Rydzanic M, Stephani U, Siebert R, Ammerpohl O, Caliebe A. Array-based DNA-methylation analysis in individuals with developmental delay/intellectual disability and normal molecular karyotype. *Eur J Med Genet.* 2015.

Veröffentlichte Tagungsbeiträge

Gutwein J, Sieprath S, Nagel I, Wilhelms C, Arnold N, Weimer J, Siebert FA, Niehoff P, Grunewald R, Schreer I, Jonat W, Siebert R, Heidemann S. Ist die Strahlensensibilität von *BRCA1*-Mutationsträgerinnen erhöht? 29. Treffen der Norddeutschen Humangenetiker/Innen, 13. November 2010, Hamburg, Deutschland.

Gutwein J, Sieprath S, Nagel I, Wilhelms C, Arnold N, Weimer J, Siebert FA, Niehoff P, Grunewald R, Schreer I, Jonat W, Siebert R, Heidemann S. Is there a risk for *BRCA1* mutation carriers by the frequent use of mammography? 22. Jahrestagung der deutschen Gesellschaft für Humangenetik (GfH), 16.–18. März 2011, Regensburg, Deutschland.

Ammerpohl O, Bens S, Bergmann A, **Gutwein J**, Haake A, Hoff K, Kolarova J, Richter J, Siebert R. Der zweite Code. 108. Jahrestagung der Deutschen Gesellschaft für Kinder- und Jugendmedizin e. V. *MedReport* 22/36. Jahrgang 2012.

Friemel C, Bens S, Caliebe A, **Gutwein J**, Richter J, Schmutzler A, Siebert R, Ammerpohl O. A methylation-sensitive SNP in *PIWIL2* is associated with male infertility. 23. Jahrestagung der deutschen Gesellschaft für Humangenetik (GfH) gemeinsam mit der European Journal of Human Genetics (ESHG), 23.–26. Juni 2012, Nürnberg, Deutschland, *Medizinische Genetik*, Heft 1, Juni 2012.

Kolarova J, **Gutwein J**, von Kampen O, Schafmayer C, Kalthoff C, Hampe J, Siebert R, Ammerpohl O. A DNA methylation survey of 117 tumor samples and corresponding normal controls revealed more than 640 genes aberrantly methylated in colorectal cancer. 23. Jahrestagung der deutschen Gesellschaft für Humangenetik (GfH) gemeinsam mit der

Publikationsliste

European Journal of Human Genetics (ESHG), 23.–26. Juni 2012, Nürnberg, Deutschland, Medizinische Genetik, Heft 1, Juni 2012.

Tangen I, Sporns P, Bens S, **Gutwein J**, Jacobsen U, Rydzanicz M, Stephani U, Gillessen-Kaesbach G, Siebert R, Ammerpohl O, Caliebe A. 23. Jahrestagung der deutschen Gesellschaft für Humangenetik (GfH) gemeinsam mit der European Journal of Human Genetics (ESHG), 23.–26. Juni 2012, Nürnberg, Deutschland, Medizinische Genetik, Heft 1, 2012.

Siebert R, Richter J, Bergmann A, **Gutwein J**, Kolarova J, Haake A, Radlwimmer B, Kabelitz D, Schrappe M, Lichter P, Ammerpohl O. DNA-Methylation Profiling of Normal and Malignant B- and T-Cells. Blueprint-Epigenome: International Human Epigenome Consortium (IHEC)-Treffen, 6.–8. September 2012, Seoul, Südkorea.

Papakonstantinou N, **Gutwein J**, Ammerpohl O, Koravou E, Belessi C, Siebert R, Kneba M, Ritgen M, Stamatopoulos K, Pott C. CLL Subsets with Distinct Stereotyped B Cell Receptors Have Distinct Epigenetic Make-up, Even Beyond IGHV Gene Mutational Status: DNA Methylation Profiling of IGHV-Unmutated CLL Stereotyped Subsets #6 and #8. 54th ASH Annual Meeting and Exposition, 8.–11. Dezember 2012, Atlanta, USA.

Gutwein J, Richter J, Altmüller J, Frommolt P, Nürnberg P, Siebert R, Caliebe A. Investigation of maternal effect mutations in the NLRP gene family in women with spontaneous abortions. 24. Jahrestagung der Deutschen Gesellschaft für Humangenetik (GfH) gemeinsam mit der Österreichischen Gesellschaft für Humangenetik (ÖGH) und der Schweizerischen Gesellschaft für Medizinische Genetik (SGMG), 20.–22. März 2013, Dresden, Deutschland.

Wagener R, Ammerpohl O, Kolarova J, **Gutwein J**, Murga Penas EM, Richter J, Drexler HG, MacLeod RAF, Siebert R. Significant differences in the DNA-methylation profile between EBV-positive and EBV-negative Burkitt Lymphoma cell lines. 24. Jahrestagung der Deutschen Gesellschaft für Humangenetik (GfH) gemeinsam mit der Österreichischen Gesellschaft für Humangenetik (ÖGH) und der Schweizerischen Gesellschaft für Medizinische Genetik (SGMG), 20.–22. März 2013, Dresden, Deutschland.

Publikationsliste

Gutwein J, Dalgård C, Kyvik K, Olbrich D, Dittmar M, Siebert R, Ammerpohl A. Shift and night work may shape the DNA methylome in human. *European Journal of Human Genetics*. 2013;Volume 21 Supplement 2. European Human Genetics Conference (ESHG), 08.–11. Juni 2013, Paris, Frankreich.

Papakonstantinou N, Malousi A, Tsagiopoulou M, **Gutwein J**, Ammerpohl O, Kneba M, Anagnostopoulos A, Ritgen M, Maglaveras N, Siebert R, Pott C, Stamatopoulos K. DNA Methylation Changes In Patients With Chronic Lymphocytic Leukemia Relapsing After Treatment Are Not Stochastic But Rather Selectively Affect Critical Pathways For B-Cell Physiology. 55th ASH Annual Meeting and Exposition, 7.–10. Dezember 2013, New Orleans, USA.

Vater I, Montesinos Rongen M, Schlesner M, Alexandrov LB, **Gutwein J**, Haake A, Nagel I, Nazzari I, Purschke F, Richter J, Sprute R, Deckert M, Siebert R. The mutational landscape of primary lymphomas of the central nervous system determined by whole exome sequencing. 25. Jahrestagung der Deutschen Gesellschaft für Humangenetik (GfH) gemeinsam mit der Österreichischen Gesellschaft für Humangenetik (ÖGH) und der Schweizerischen Gesellschaft für Medizinische Genetik (SGMG), 19.–21. März 2014, Essen, Deutschland. *Medizinische Genetik*, Heft 1, 2014.

Kolarova J, Bens S, **Gutwein J**, Monk D, Eggermann T, Buiting K, Lehold J, Riepe FG, Thiele-Schmitz S, Gillessen-Kaesbach G, Prawitt D, Ammerpohl O, Horsthemke B, Caliebe A, Siebert R. Array based DNA Methylation analysis in a cohort of patient with multi locus methylation defects. 1st European Imprinting School 19.–21. Mai 2014, Mery-sur-Oise, Frankreich.

12 Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation eigenständig und nur mit den von mir angegebenen Quellen und Hilfsmitteln bzw. mit der wissenschaftlichen Beratung meiner akademischen Betreuer angefertigt habe. Die vorliegende Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden und wurde bisher in gleicher oder ähnlicher Form an keiner anderen Universität zur Begutachtung im Rahmen eines Prüfungsverfahrens vorgelegt. Weiterhin bestätige ich, dass ich zuvor noch keine Promotionsversuche unternommen habe.

Folgende Teile der vorliegenden Arbeit wurden in Form eines Artikels in einer internationalen Fachzeitschrift veröffentlicht bzw. auf einem Kongress präsentiert:

- Roessler J, Ammerpohl O, Gutwein J, Hasemeier B, Anwar SL, Kreipe H, Lehmann U. Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. BMC Res Notes. 2012;5:210.
- Gutwein J, Dalgård C, Kyvik K, Olbrich D, Dittmar M, Siebert R, Ammerpohl A. Shift and night work may shape the DNA methylome in human. European Journal of Human Genetics. 2013;Volume 21 Supplement 2. European Human Genetics Conference, 08.–11. Juni 2013, Paris, Frankreich.

Kiel, den 29.01.2015

Jana Gutwein