

Analyse genomweiter DNA-Methylierungsmuster humaner und muriner Zelltypen mittels MeDIP-Chip

Dissertation

zur Erlangung des Grades des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät III

Chemie, Pharmazie, Bio- und Werkstoffwissenschaften

der Universität des Saarlandes

vorgelegt von

Nina Gasparoni

Saarbrücken, im April 2012

Datum des Kolloquiums: 19.12.2012

Dekan: Prof. Dr. Stefan Diebels
Vorsitzender: Prof. Dr. Alexandra Kiemer
1. Berichterstatter: Prof. Dr. Jörn Walter
2. Berichterstatter: Prof. Dr. Uli Müller
Akad. Mitarbeiter: Dr. Daniel Müller

Danksagung

Allen voran möchte mich ganz herzlich bei meinem Doktorvater Herrn Prof. Dr. Jörn Walter dafür bedanken, dass er mir diese Arbeit ermöglicht, mich die letzten Jahre mit seinem Vertrauen in mich als auch mit hilfreichen Diskussionen und Anregungen unterstützt sowie meinen Perfektionismus wenn nötig gebremst hat.

Auch möchte ich Herrn Prof. Dr. Uli Müller dafür danken, sich als Zweitgutachter zur Verfügung zu stellen.

Meiner „Mitstreiterin“ Diana Santacruz, die in der gemeinsamen MeDIP-Zeit mit mir durch dick und dünn ging, gilt ein besonderer Dank für die tolle Zusammenarbeit und die nützlichen Diskussionen.

Weiterhin bedanken möchte ich mich bei Jasmin Gries und Christina LoPorto für ihre Hilfeinsätze im Labor sowie bei Pavlo Lutsik für seine Unterstützung in bioinformatischen Fragen. Herzlichen Dank auch an Dr. Konstantin Lepikhov, Dr. Sascha Tierling sowie an ehemalige und derzeitige Doktoranden (Corry Schwierzy, Julia Arand, Mark Wossidlo, Sarah Fuchs, Anne Dubberke,..) für ihre Ratschläge und Ermunterungen. Auch Nicole Jundel danke ich für ihre Unterstützung nicht nur in administrativer Hinsicht.

Ganz besonderer Dank gilt meinen Eltern, meiner Schwester sowie meinem Ehemann, die immer für mich da sind. Mit ihrer immer währenden Unterstützung, ihrem Verständnis und ihrem Vertrauen in meine Fähigkeiten haben sie mich auf diesem Weg begleitet und auch in schwierigen Phasen immer wieder motiviert.

Nina Gasparoni

Saarbrücken, im April 2012

„Ohne Spekulation gibt es keine neue Beobachtung.“
(Charles Darwin)

Inhaltsverzeichnis

Danksagung.....	II
Inhaltsverzeichnis	IV
1 Einleitung	1
1.1 Epigenetik.....	1
1.1.1 Epigenetische Modifikationen und Chromatin-Struktur	1
1.1.2 Epigenetische Modifikationen in pluripotenten und differenzierten Zellen.....	6
1.2 Stammzellen.....	9
1.2.1 Eigenschaften und Generierung pluripotenter Stammzellen.....	9
1.2.2 gPS: aus der Keimbahn abstammende pluripotente Stammzellen	13
1.2.3 EpiSC: Stammzellen des Epiblasten	14
1.3 Kryokonservierung.....	16
1.4 Epigenetik und Stress.....	18
1.4.1 Kryokonservierung, abiotischer Stress und Epigenetik	18
1.4.2 Oxidativer Stress und DNA-Methylierung	19
1.5 Methoden zur Analyse von DNA-Methylierung.....	21
1.6 Aufgabenstellung	22
1.6.1 Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung	22
1.6.2 Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryopreservation	23
1.6.3 Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen	23
2 Material und Methoden	25
2.1 Material.....	25
2.1.1 Chemikalien.....	25
2.1.2 Lösungen und Puffer	26
2.1.3 Kulturmedien	28
2.1.4 Reaktionskits	28
2.1.5 Größenstandards für Agarosegele.....	29
2.1.6 Enzyme	29
2.1.7 Antikörper.....	29
2.1.8 Oligonukleotide	30
2.1.9 Microarrays.....	31
2.1.10 Eukaryotische Zellen + DNA.....	32
2.1.11 Bakterienstämme	34
2.1.12 Instrumente und Maschinen.....	34
2.2 Methoden.....	35
A. Molekularbiologische Methoden.....	35
2.2.1 Isolation und Analyse von Nukleinsäuren	35
2.2.2 Analyse der DNA-Methylierung	43
B. Bioinformatische Methoden und Programme	51

2.2.3	Analyse der MeDIP-Chip-Daten	51
2.2.4	Analyse von mRNA-Expressions-Daten.....	58
2.2.5	Funktionelle Analyse mittels Gen-Ontologie	59
2.2.6	Analyse von Daten der Bisulfit-Sequenzierung	59
2.2.7	Liste verwendeter Datenbanken, Programme, Softwares und Web-Seiten.....	59
3	Ergebnisse	62
3.1	Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung.....	62
3.1.1	Etablierung experimenteller Parameter der Immunopräzipitations-basierten Methode ..	63
3.1.2	Etablierung bioinformatischer Bausteine zur Überprüfung und Analyse der Daten aus MeDIP-Chip	70
3.2	Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryopreservation	84
3.2.1	Kryokonservierung von HepG2 Zellen: experimentelles Design	85
3.2.2	Immunopräzipitation und Qualitätskontrolle vor/nach der Amplifikation.....	86
3.2.3	Analyse chromosomenweiter DNA-Methylierungsmuster nach Kryokonservierung mit MeDIP-Chip	88
3.2.4	Genomweite mRNA-Expression in Zusammenhang mit Kryopreservation.....	92
3.2.5	Quantitative Methylierungsanalysen ausgesuchter Zielregionen	96
3.3	Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen	99
3.3.1	Immunopräzipitation und Qualitätskontrolle	100
3.3.2	Analyse der DNA-Promotormethylierung mittels MeDIP-Chip - Gruppierung der Zelltypen zueinander	102
3.3.3	Analyse der keimbahnassoziierten Stammzellen: gPS gegenüber GSC und ESC	104
3.3.4	Analyse späterer Pluripotenz-Stadien: EpiSC gegenüber MEF und ESC.....	108
3.3.5	Validierung der MeDIP-Chip-Daten mittels Bisulfit-Sequenzierung	112
4	Diskussion	115
4.1	Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung.....	115
4.1.1	Wesentliche Aspekte hinsichtlich experimenteller Parameter	115
4.1.2	Unterschiedliches Präzipitationsverhalten humaner und muriner DNA	117
4.1.3	Qualitätskontrollen, Prä-Prozessierung und Vergleichbarkeit von Datensätzen.....	118
4.1.4	Verlust der Sensitivität durch genomrepräsentative Amplifikation.....	120
4.1.5	Identifikation von DMRs und Sensitivität der MeDIP-Chip-Pipeline	121
4.1.6	Kritische Betrachtung der Einbeziehung von Sequenzinformationen in die Daten-Evaluation	122
4.1.7	Fazit.....	123
4.2	Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryokonservierung	124
4.2.1	Geringfügige Abweichungen der MeDIP-Anreicherung nach Kryokonservierung reflektieren technische und biologische Schwankungen.....	124
4.2.2	Änderungen der mRNA-Expression nach Kryokonservierung zeigen keinen Zusammenhang mit DNA-Methylierung.....	127
4.2.3	Fazit.....	129
4.3	Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen	130

4.3.1	Die MeDIP-Chip-Daten erlauben eine sensitive Erfassung von Metylierungsunterschieden im Promotor-Bereich	131
4.3.2	Distinkte Methylierungsmuster der Aufwärts- und Kern-Promotoren.....	131
4.3.3	Genomweite Methylierungs- und Expressionsmuster der pluripotenten gPS heben sich deutlich von den keimbahnspezifischen Profilen der GSC.....	132
4.3.4	Die Methylierungsprofile der EpiSC spiegeln das Stadium der eingeschränkten Pluripotenz wider	138
Zusammenfassung		143
Summary		144
5	Referenzen	145
6	Anhang.....	156
6.1	Anhang A: Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung	156
6.1.1	Resultate zur Etablierung experimenteller Parameter.....	156
6.1.2	Resultate zur bioinformatischen Analyse von MeDIP-Chip-Daten	158
6.2	Anhang B: Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryopreservation	160
6.2.1	Protokoll zur Kryopreservation von HepG2 Zellen	160
6.2.2	Resultate zur Immunopräzipitation von HepG2-Zellen.....	160
6.2.3	Resultate zur Evaluierung differentiell angereicherter bzw. exprimierter Gene in Verbindung mit Kryopreservation	161
6.3	Anhang C: Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen.....	168
6.3.1	Resultate zur Evaluierung differentiell angereicherter Gene in murinen Stammzellen..	168
6.3.2	Analyse der DNA-Methylierung mittels Bisulfit-Sequenzierung	179
7	Abkürzungsverzeichnis	VII
8	Abbildungsverzeichnis	IX
9	Tabellenverzeichnis	XI
10	Lebenslauf	XIII
Eidesstattliche Erklärung		XV

1 Einleitung

1.1 Epigenetik

Als Teilgebiet der Genetik umfasst die Epigenetik („epi“[griechisch]: daneben, darüber) solche Mechanismen, welche jenseits der bekannten Prinzipien hinsichtlich der DNA-Sequenz einen vererbaren zellulären Phänotyp bestimmen (Bird 2002, Jaenisch und Bird 2003). Zelltypspezifische Signaturen, die durch DNA-Methylierung, Histon-Modifikationen, Nukleosomen-Remodelierung und nicht-kodierende RNAs bestimmt werden, üben dabei Einfluss auf die Regulation bzw. Expression von Genen aus (Wu et al. 2012). Störungen hinsichtlich der Festlegung oder Interaktion der genannten Komponenten können aufgrund einer veränderten Genexpression zu diversen Entwicklungsdefekten sowie epigenetischen Erkrankungen führen.

1.1.1 Epigenetische Modifikationen und Chromatin-Struktur

1.1.1.1 Basen-Modifikationen der DNA / DNA-Methylierung

Die Modifikation genomischer DNA an den Nukleinbasen Cytosin und Adenin stellt eine wichtige Form epigenetischer Modifikationen dar. Chemische Gruppen werden dabei kovalent an die Base des Nukleotids gebunden. In Säugern ist insbesondere die (symmetrische) Methylierung der C5-Position eines Cytosins (**5-Methylcytosin**, 5mC) im CpG-Kontext von Bedeutung. CpG-Dinukleotide sind nicht gleichmäßig über das Genom verteilt sondern größtenteils in Form CG-reicher Abschnitte, sogenannter **CpG-Inseln**, organisiert (Gardiner-Garden und Frommer 1987, Deaton und Bird 2011). Diese sind meist mit regulatorischen Regionen wie Promotoren assoziiert, welche im Gegensatz zu den übrigen über das Genom verstreuten CpGs vornehmlich methylierungsfrei sind (Bird 2002, Deaton und Bird 2011, Wu et al. 2012). Eine erhöhte Methylierung von Promotoren hat in der Regel eine transkriptionelle Repression des entsprechenden Gens zur Folge. Dabei wird die Bindung von Transkriptionsfaktoren (TF) behindert bzw. durch Rekrutierung MBD-assoziiierter Repressor-Komplexe (MBD = Methyl-bindende Domäne) eine repressive Chromatin-Konfiguration gefördert (Bird und Wolffe 1999, Ng et al. 1999, Bird 2002, Suzuki und Bird 2008) (Abbildung 1.1-1 (A)). Im Gegensatz dazu wurde intra-genische Methylierung

jüngst mit transkriptioneller Aktivität in Verbindung gebracht. Die Bedeutung der Genkörper sowie CpG-armer Ausläuferregionen der CpG-Inseln wurde erst kürzlich hinsichtlich gewebe-, entwicklungs- und krankheitsspezifischer Methylierungsmuster erkannt (Irizarry et al. 2009, Ball et al. 2009, Dindot et al. 2009, Lister et al. 2009, Laurent et al. 2010). Auch CpG-arme distale Regionen geringer Methylierung (LMR, „*low methylated regions*“) scheinen in Kombination mit zelltypspezifischen TFs regulatorische Funktionen (Enhancer) zu übernehmen (Stadler et al. 2011).

Die epigenetische Kontrollfunktion der DNA-Methylierung steht mit verschiedenen Prozessen in Zusammenhang wie: i) Regulation gewebespezifischer Expression während Entwicklung und Differenzierung, ii) genomisches Imprinting, iii) Stilllegung transposabler Elemente zur Erhaltung genomischer Stabilität, iv) X-Chromosomen-Inaktivierung im Zuge der Dosiskompensation und v) Organisation der Chromatin-Struktur (Bird 2002, Suzuki und Bird 2008).

Drei katalytisch aktive Enzyme sind für die Etablierung und Erhaltung der DNA-Methylierungsmuster verantwortlich. Die in der Keimbahn exprimierten **Methyltransferasen** Dnmt3a und Dnmt3b dienen einer *de novo* Methylierung zuvor unmethylierter Bereiche während der Embryonalentwicklung/ Differenzierung (Okano et al. 1999, Okano et al. 1998, Goll und Bestor 2005). Dnmt3b konnte mit einer Stilllegung repetitiver Elemente in Verbindung gebracht werden, während Dnmt3a offenbar für die Etablierung allelspezifischer Methylierungsmuster zuständig ist. Dabei scheint auch die Interaktion mit Dnmt3L von Bedeutung zu sein, für welches indessen keine katalytische Aktivität nachgewiesen werden konnte (Bourc'his und Bestor 2004, Kaneda et al. 2004). Eine Erhaltung vorhandener Methylierungsmuster während späterer Zellteilungen wird durch die Funktion der DNA-Methyltransferase Dnmt1 ermöglicht. Diese wird durch Uhrf1 zu Replikationsfoci rekrutiert und kann aufgrund ihrer Präferenz für hemi-methylierte DNA während der Replikation die Methylierung vom Elternstrang auf den Tochterstrang übertragen und so erhalten (Bestor 2000, Goll und Bestor 2005, Bostick et al. 2007).

Während die Methylierung im **nicht-CpG-Kontext** in Pflanzen schon lange bekannt und erforscht ist, wurde diese Modifikation in Säugern erst kürzlich erkannt. Asymmetrische CHG- und CHH-Methylierung (H = A, C oder T) kommt in niedrigen Leveln in embryonalen Stammzellen (Maus und Mensch) und frühen Embryonen vor, jedoch kaum in somatischen Zellen (Lister et al. 2009, Dodge et al. 2002, Ramsahoye et al. 2000, Haines et al. 2001, Laurent et al. 2010). Interessanterweise zeigen diese Modifikationen eine starke Variabilität in pluripotenten Zellen (verschiedene ESC-Linien und -Passagen) und scheinen anfälliger für stochastische Variation zu sein als CpG-Methylierung (Ziller et al. 2011). Die

nicht-CpG-Methylierung und ihre mögliche Funktion ist noch Gegenstand aktueller Forschung. Man vermutet jedoch eine Beteiligung an der Regulation von Enhancern, welche für die Wahl olfaktorischer Rezeptoren im Maus-Gehirn nötig sind (Lomvardas et al. 2006).

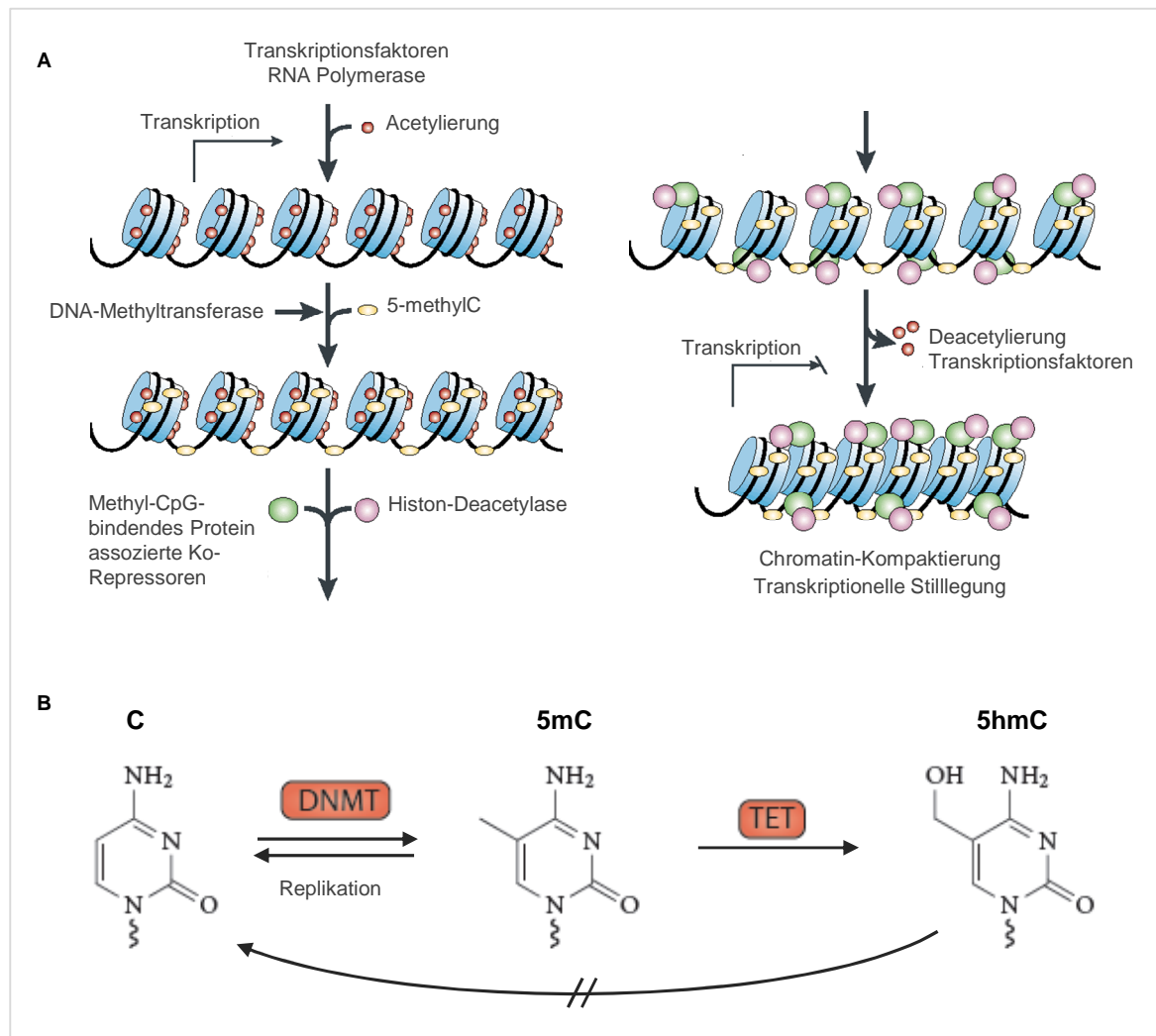


Abbildung 1.1-1: DNA-Methylierung in Säugern. (A) Schematische Darstellung der Stilllegung transkriptionell aktiver Bereiche durch Interaktion von DNA-Methylierung, Methyl-bindender Faktoren sowie Histon-Deacetylasen, wodurch eine Kompaktierung des Chromatins bewirkt wird (verändert nach Robertson und Wolffe 2000). (B) Schematische Darstellung der Basenmodifikationen 5mC und 5hmC. Methylierung des Cytosins mittels DNMTs zu 5mC bzw. passive Demethylierung desselben während der Replikation sind gegenüber der Konvertierung von 5mC mittels TET zu 5hmC und einer passiven/aktiven Demethylierung desselben (verschiedene Mechanismen und Zwischenstufen) dargestellt (verändert und vereinfacht nach Branco et al. 2012).

Obwohl DNA-Methylierung als eine stabile Modifikation zur Langzeit-Stilllegung von Genen erachtet wird, können auch bestimmte **Demethylierungsereignisse** beobachtet werden. Diese sind während der Entwicklung offenbar essentiell zur Etablierung pluripotenter Stadien in Zygoten (Preimplantations-Embryonen) bzw. elternspezifischer Imprints in

primordialen Keimzellen (Reik et al. 2001). Eine dynamische Regulation der DNA-Methylierung findet sich zudem während der Differenzierung und in somatischen Zellen als Antwort auf spezifische Signale (Meissner 2010, Wu und Zhang 2010). Die bei der Modulation existierender Methylierungsmuster involvierten aktiven Mechanismen sind noch nicht komplett bekannt und verstanden. Es wird jedoch ein Zusammenhang mit Hydroxylasen der TET-Familie („ten-eleven translocation“) und durch sie katalysierte iterative Oxidationsvorgänge sowie anschließende Reparaturvorgänge vermutet (Wu und Zhang 2010).

5-Hydroxymethylcytosin (5hmC), ein TET-vermitteltes Oxidationsprodukt von 5mC (Abbildung 1.1-1 (B)), wurde erst kürzlich als neue Basenmodifikation identifiziert und ist in undifferenzierten murinen embryonalen Stammzellen (ESC, „*embryonic stem cell*“) sowie Purkinje-Neuronen vorhanden (0,6% und 0,2% aller Nukleotide) (Kriaucionis und Heintz 2009, Tahiliani et al. 2009, Ito et al. 2010). Neben weiteren Derivaten wie 5-Formylcytosin (5fC) und 5-Carboxylcytosin (5caC) (Ito et al. 2011, He et al. 2011) könnte 5hmC lediglich das Intermediat einer passiven oder aktiven Demethylierung darstellen oder aber ein neues epigenetisches Level. Verschiedene aktuelle Studien, welche sich mit der genomweiten Verteilung von hmC und Tet1 in murinen sowie humanen ESC befassen (Ficz et al. 2011, Pastor et al. 2011, Stroud et al. 2011, Szulwach et al. 2011, Williams et al. 2011, Wu et al. 2011a, Wu et al. 2011b, Xu et al. 2011), deuten auf kontextabhängige duale Funktionen (Repression und Aktivierung von Genen) während der Entwicklung hin, welche offenbar hinsichtlich Pluripotenz und Differenzierung eine Rolle spielen (Wu und Zhang 2011) (siehe auch Abschnitt 1.1.2.1).

1.1.1.2 Histon-Modifikationen

Die DNA eukaryotischer Organismen erreicht durch ihre Assoziation mit Histon-Proteinen und Nicht-Histon-Proteinen einen erstaunlichen Verpackungsgrad. Die kleinste, sich wiederholende Einheit der als Chromatin bezeichneten Struktur stellt das Nukleosom dar. Dieses besteht aus 147 nt doppelsträngiger DNA, welche um ein Histon-Oktamer aus je zwei Molekülen der Kern-Histone H2A, H2B, H3 und H4 gewunden ist. Die Nukleosomen sind durch „linker-DNA“ miteinander verbunden, welche mit dem „linker-Histon“ H1 assoziiert sind. Kovalente Markierungen aminoterminaler und auch carboxyterminaler Histonschwänze stellen eine weitere Ebene epigenetischer Regulation dar. An mehreren Positionen können N-terminale Enden der Histone mittels Acetylierung, Methylierung (Mono-, Di-, Tri-), Phosphorylierung, Ubiquitinierung und Sumoylierung modifiziert werden (Jenuwein und Allis 2001, Kouzarides 2007, Bernstein et al. 2007) (Abbildung 1.1-2). Die

Vielfalt dieser posttranslationalen Modifikationen und ihre Kombinationen untereinander legen einen **Histone-Code** nahe, welcher das Informationspotential des genetischen Codes erweitert (Jenuwein und Allis 2001). Das Bindungsverhalten zwischen DNA und Histonen wird abhängig von Typ und Lokalisation der Modifikationen beeinflusst, wodurch der Verpackungsgrad des Chromatins und damit auch die Genaktivität bestimmt werden. Eine Acetylierung von Lysinresten, welche durch Schwächung der DNA-Histon-Interaktion eine Auflockerung der Chromatinstruktur bewirkt, gilt als Kennzeichen euchromatischer, transkriptionell aktiver Regionen des Genoms (Bernstein et al. 2007). Eine solch klare Zuordnung ist im Falle der Methylierung nicht gegeben - der Einfluss ist abhängig von Position und Grad der Methylierung. Eine Methylierung des Lysins an der vierten Position des Histons 3 (H3K4me) in Promotor- bzw. Enhancer-Bereichen fördert eine transkriptionell aktive Konformation, während H3K9me3 und H3K27me3, generell stillgelegte/heterochromatische Bereiche markieren (Bernstein et al. 2005, Barski et al. 2007, Heintzman et al. 2007, Bernstein et al. 2007). H3K36me3 kann in Abhängigkeit der Entfernung zur Promotor-Region mit einer Gen-Aktivierung oder -Stilllegung assoziiert sein. In Genkörpern aktiv transkribierter Gene ist es v.a. in Exons angereichert und kolokalisiert mit intra-genischer DNA-Methylierung (Ball et al. 2009, Hodges et al. 2009).

Viele der **Histon-modifizierenden Enzyme** sind bereits bekannt bzw. näher charakterisiert, wie Histon-Acetyltransferasen, Deacetylasen, Methyltransferasen sowie Histon-Demethylasen. Zu den am besten untersuchten Faktoren gehören Proteinkomplexe der Polycomb- und Trithorax-Gruppen (PcG, trxG). Diese sind an der Ausbildung bivalenter Domänen beteiligt und spielen eine wesentliche Rolle bei der Regulationentwicklungsspezifischer Gene (siehe Abschnitt 1.1.2.1) (Ringrose und Paro 2004, Schuettengruber et al. 2007, Bernstein et al. 2007).

Die Chromatinstruktur wird des Weiteren durch **DNA-interagierende Proteine** manipuliert, wodurch ebenfalls Einfluss auf die Genaktivität genommen wird. Nukleosomen-Remodelierungs-Komplexe können die in Nukleosomen gepackte DNA durch Verschiebung der Histone exponieren (und so z.B. der Transkriptions- oder Methylierungsmaschinerie Zugang gewähren) oder Untereinheiten der Histon-Oktamere ersetzen bzw. entfernen (Cairns 2007).

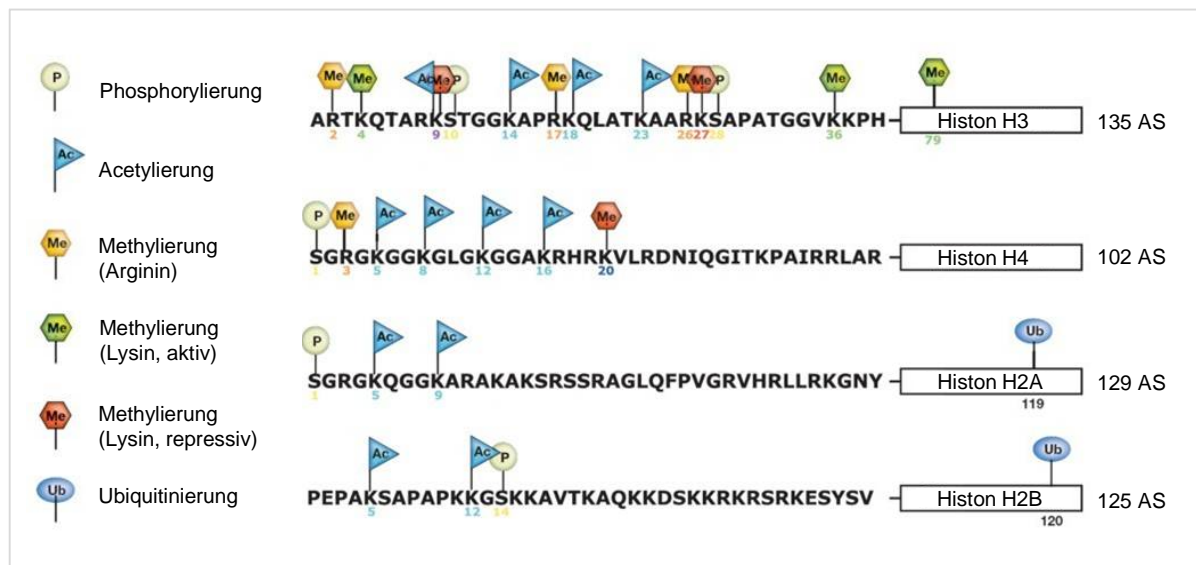


Abbildung 1.1-2: Schematische Darstellung des Histon-Codes. Post-Translationale Modifikationen wie Methylierung, Acetylierung und Phosphorylierung sind für verschiedene Positionen der Histone H3, H4, H2A und H2B mit potentieller Funktion (Aktivierung, Reprimierung) gekennzeichnet (verändert nach http://www.cshteaching.org/main/epigenetics/CH05/06_EPIGEN_CH05.jpg).

1.1.1.3 Nicht-kodierende RNAs

Neben ihrem eher direkten Einfluss auf die Genexpression auf transkriptioneller und post-transkriptioneller Ebene, können nicht-kodierende RNAs (ncRNAs) auch auf epigenetischer Ebene agieren. Verschiedene Typen an ncRNAs (z.B. lange intergenische ncRNAs, miRNAs und Promotor-assoziierte RNAs) wurden mit epigenetischen Mechanismen transkriptioneller Stilllegung in Verbindung gebracht, wobei sie Chromatin-modifizierende Komplexe zu rekrutieren bzw. mit diesen interagieren zu scheinen (z.B. X-Inaktivierung, Regulation imprinteter Gene) (Kaikkonen et al. 2011).

1.1.2 Epigenetische Modifikationen in pluripotenten und differenzierten Zellen

Während der Säuger-Entwicklung und der Ausbildung distinkter Zelltypen werden weitreichende Änderungen der Transkriptionsprogramme beobachtet, welche ein flexibles und dynamisches Epigenom erfordern. In Abhängigkeit der Anforderungen an die Zelle stehen dabei Markierungen für ein epigenetisches Kurz- und Langzeit-Gedächtnis zur Verfügung.

1.1.2.1 Epigenetische Modifikationen in pluripotenten Zellen

Pluripotente Zellen (siehe auch Abschnitt 1.2) repräsentieren ein Grundstadium der Entwicklung, das durch komplexe autoregulative Netzwerke von Transkriptionsfaktoren (TF) - u.a. Oct4, Sox2, Nanog - definiert und erhalten wird (Jaenisch und Young 2008, Silva und Smith 2008, Meissner 2010).

Zusätzlich wird dieses Stadium, bzw. das Differenzierungspotential, durch spezifische epigenetische Signaturen reflektiert. Embryonale Stammzellen (ESC) sind durch offenes hochdynamisches Chromatin mit aktiven Histon-Modifikationen (H3K9a, H3K4me3, H3K36me3) gekennzeichnet (Mikkelsen et al. 2007, Guenther et al. 2007, Meissner 2010). Diese **offene Konformation** in nicht-kompaktem Euchromatin erlaubt einen einfachen Zugang für Transkriptionsfaktoren bzw. die Transkriptionsmaschinerie und ist vermutlich verantwortlich für die genomweite basale, stochastische Expression (Fisher und Fisher 2011). CpG-reiche Promotoren sind in der Regel unmethyliert und mit Haushalts- oder entwicklungspezifischen Genen assoziiert. Letztere sind meist durch **bivalente Chromatinstrukturen** stillgelegt (Bernstein et al. 2006, Mikkelsen et al. 2007, Meissner et al. 2008, Meissner 2010). PcGs katalysieren und binden H3K27me3 und/oder H2AK119Ub1 (Polycomb-repressiver Komplex (PRC) 2 bzw. 1), wodurch die RNA-Polymerase II aufgrund der Chromatin-Kompaktierung behindert wird. Gleichzeitig sind diese Regionen mit aktivierenden Histon-Modifikationen wie H3K9ac und H3K4me3 ko-markiert (durch trxG MII mit Histon-Methyltransferasen). Diese bivalenten Chromatin-Domänen dienen vermutlich einer Vorbereitung entwicklungspezifischer Promotoren zur schnellen Aktivierung/Inaktivierung während der Differenzierung (Bernstein et al. 2006, Mikkelsen et al. 2007, Fisher und Fisher 2011, Meissner 2010). Demgegenüber sind CpG-arme Promotoren, welche meist mit keimzell- bzw. gewebespezifischen Genen assoziiert sind, tendenziell methyliert (Mohn et al. 2008, Meissner et al. 2008, Weber et al. 2007, Fouse et al. 2008, Meissner 2010). Diese Gruppe von Genen wird z.T. unabhängig der PcGs oder Oct4/Nanog-Zielgene durch DNA-Methylierung reguliert (Fouse et al. 2008).

Des Weiteren tragen auch **Tet-Proteine** und **5hmC** zur Signatur und Regulation pluripotenter Zellen bei und wurden bereits in verschiedene Studien bzgl. ihrer genomweiten Verteilung analysiert (humane und murine ESC) (Ficz et al. 2011, Pastor et al. 2011, Williams et al. 2011, Wu et al. 2011a, Wu et al. 2011b, Xu et al. 2011). Dabei werden duale Funktionen hinsichtlich der transkriptionellen Regulation nahegelegt (Wu und Zhang 2011). Zum Einen sind Tet1 und hmC deutlich in Promotoren bivalenter Domänen angereichert und scheinen an der PcG-geförderten Repression entwicklungspezifischer Gene beteiligt zu sein (Pastor et al. 2011, Williams et al. 2011, Wu et al. 2011a, Wu et al.

2011b, Xu et al. 2011). Zum Anderen finden sich Tet1 und 5hmC in Promotor-Regionen bzw. Genkörpern aktiv transkribierter Gene, u.a. mit Pluripotenz in Zusammenhang stehende TFs (Nanog, Tcf1, Esrrb) (Ficz et al. 2011, Pastor et al. 2011, Williams et al. 2011, Wu et al. 2011b). Dabei scheinen sie den hypomethylierten Status der CGI zu fördern bzw. zu erhalten. Ein Tet1 knockdown führt zu einer Behinderung der Selbsterneuerung und Erhaltung von ESC (Ito et al. 2010, Wu und Zhang 2010). Tet-Proteine scheinen also ein Teil des Pluripotenznetzwerks zu sein und sind selbst Zielgene wichtiger Pluripotenz-Faktoren (Koh et al. 2011). Das in aktiven Enhancern vorkommende 5hmC könnte zudem eine Rolle bei der Regulation der Enhancer-Funktion innehaben (Williams et al. 2011, Wu et al. 2011a, Ficz et al. 2011, Pastor et al. 2011, Stroud et al. 2011, Szulwach et al. 2011, Wu und Zhang 2011).

1.1.2.2 Epigenetische Modifikationen während der Entwicklung

Während der Differenzierung kommt es zu globalen Änderungen epigenetischer Modifikationen und damit der Chromatin-Organisation. Dabei ist insgesamt eine Kompaktierung der Chromatinstruktur zu verzeichnen, was zu einer Stilllegung nicht benötigter genomischer Abschnitte führt (Meissner 2010). In Euchromatin vorkommende Bereiche mit H3K9me₂, sogenannte LOCKs („*large chromatin domains of H3K9me₂*“) nehmen während der ESC-Differenzierung in Größe und Anzahl zu (4-46% Genom-Abdeckung) (Wen et al. 2009, Fisher und Fisher 2011). Tet1- und hmC-level nehmen ab (Tahiliani et al. 2009). Bivalente Domänen lösen sich, vermutlich unter Einbeziehung von Histon-Demethylasen und DNA-Methyltransferasen, in univalente Regionen auf. Dies kann zu einer Aktivierung oder stabilen Stilllegung der entsprechenden Gene führen. Auch konnten bivalente Regionen in Zellen mit eingeschränktem Entwicklungspotential festgestellt werden. Dies lässt vermuten, dass einige Regionen - bis zu einer späteren Auflösung - unverändert bleiben bzw. neue bivalente Regionen gebildet werden (Mohn et al. 2008, Mikkelsen et al. 2007, Fisher und Fisher 2011, Meissner 2010) (Abbildung 1.1-3). Auf Ebene der DNA-Methylierung sind ebenfalls weitreichende Änderungen in Hinblick auf pluripotenz- bzw. entwicklungspezifische Gene zu verzeichnen. Promotoren werden insbesondere im Zuge der Stilllegung bivalenter Domänen methyliert. Nicht PRC2-regulierte Promotoren weisen dagegen seltener Methylierungsänderungen auf (Laurent et al. 2010, Lister et al. 2009, Mohn et al. 2008, Meissner 2010, Fisher und Fisher 2011). Distale regulatorische Regionen (1-100 kb vom Promotor entfernt) sind mit den meisten dynamischen Änderungen während der Differenzierung assoziiert (Meissner et al. 2008).

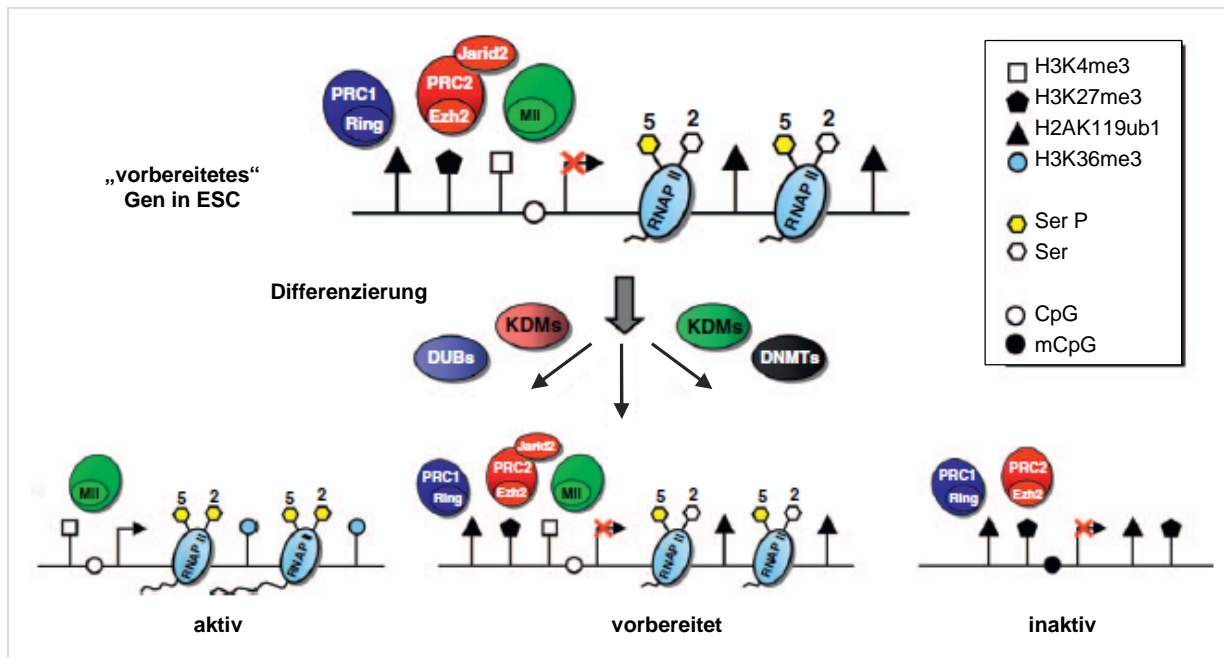


Abbildung 1.1-3: Auflösung bivalenter Domänen während der Differenzierung von ESC. Schematisch dargestellt sind (entwicklungsspezifische) Gene, welche über bivalente Chromatin-Domänen reguliert sind. In ESC sind diese Gene mit reprimierenden und aktivierenden Histonmodifikationen ko-markiert und transkriptionell inaktiv. Während der Differenzierung werden diese Regionen z.T. in aktive Regionen aufgelöst bzw. stabil stillgelegt. KDM: Lysin-Demethylase, DUB: Histon-Deubiquitinase, RNAP II: RNA-Polymerase II (verändert nach Fisher und Fisher 2011).

Ein Großteil der Methylierungsänderungen treten offenbar direkt während des Übergangs von Pluripotenz zu eher festgelegten Vorläuferzellen auf und weniger während der sukzessiven Differenzierung zu terminal differenzierten Zellen (Studien zu Differenzierung von ESC über neurale Vorläuferzellen zu Neuronen) (Mohn et al. 2008, Meissner et al. 2008, Fisher und Fisher 2011). DNA-Methylierung scheint dabei einen sekundären Stilllegungsmechanismus darzustellen, der als Langzeit-Gedächtnis zur Reprimierung nicht mehr benötigter Gene (wie Oct4) dient. Obwohl Methylierung auch an der Regulationentwicklungsspezifischer Gene beteiligt ist, werden diese während der Differenzierung eher durch flexible Histonmodifikationen (bivalente Domänen) reguliert (Meissner 2010).

1.2 Stammzellen

1.2.1 Eigenschaften und Generierung pluripotenter Stammzellen

Bei Stammzellen handelt es sich um Zellen, welche sich in Abhängigkeit ihres Potentials in festgelegte Gewebetypen oder sogar jegliches Gewebe ausdifferenzieren können. Gemäß ihrer Fähigkeit der Selbsterneuerung sowie ihrer spezifischen Kapazität bzgl. der

Differenzierung zu spezialisierten Zelltypen, können zwei Gruppen definiert werden. Embryonale Stammzellen der inneren Zellmasse eines frühen Embryos (Blastozyste) sind in der Lage, alle drei Keimblätter des sich entwickelnden Organismus zu bilden, weshalb sie als pluripotent bezeichnet werden. Post-embryonale oder adulte Stammzellen hingegen bilden während der gesamten Lebensdauer des Organismus neue spezialisierte Zellen, z.B. zur Reparatur geschädigter Gewebe. Aufgrund ihres eingeschränkten, determinierten Differenzierungspotentials handelt es sich hierbei um multi- oder unipotente Zellen (Hochedlinger und Plath 2009).

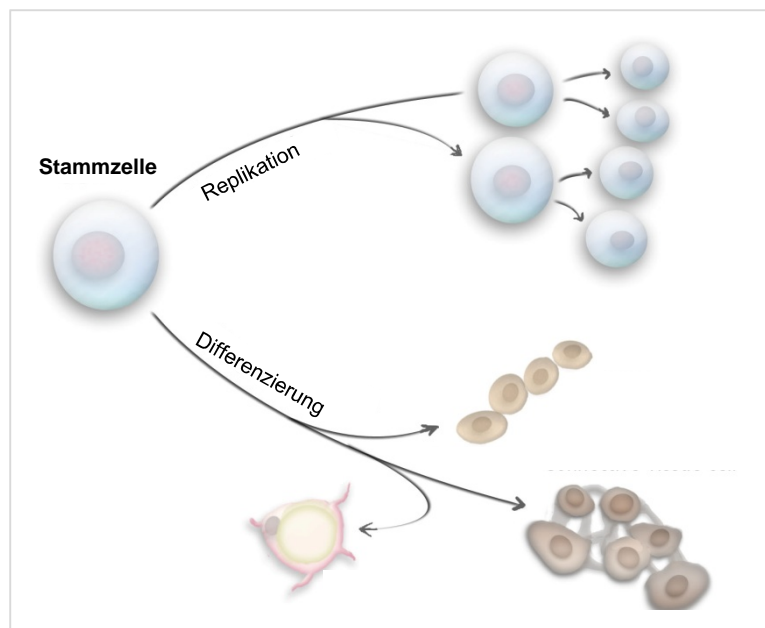


Abbildung 1.2-1: (Pluripotente) Stammzellen, mit ihren Hauptcharakteristika der Selbsterneuerung sowie des Potentials zur Differenzierung in (jegliche) Zelltypen (<http://www.purstem.eu/cmsms/uploads/images/fionn/stem%20cellsdraft4.jpg>).

In den vergangenen Jahren konnten einige Stammzelllinien generiert werden, die der Erforschung des Differenzierungspotentials verschiedener Entwicklungsstadien dienen. Dies erweitert nicht nur das Grundverständnis ontogenetischer Vorgänge (Festlegung des Zellschicksals, Differenzierung) sondern bietet auch Ansatzmöglichkeiten im Bereich der medizinischen Forschung (Stammzelltherapie).

In den 1980er-Jahren wurden erstmals Zellen der inneren Zellmasse von Prä-Implantations-Embryonen als permanente Zelllinie etabliert. Die aus Blastozysten etwa 3,5 Tage nach der Fertilisation explantierten Zellen (Maus) werden dabei durch bestimmte Kultivierungsbedingungen in ihrer Differenzierung blockiert (z.B. mittels Lif („*leukemic inhibitory factor*“)). Die Pluripotenz und Fähigkeit zur Selbsterneuerung dieser

sogenannten **embryonalen Stammzellen** (ESC, „*embryonic stem cell*“) wird damit über lange Zeit *in vitro* aufrechterhalten (Evans und Kaufman 1981, Wobus et al. 1984, Silva und Smith 2008, Nichols und Smith 2009, Albert und Peters 2009) (Abbildung 1.2-2).

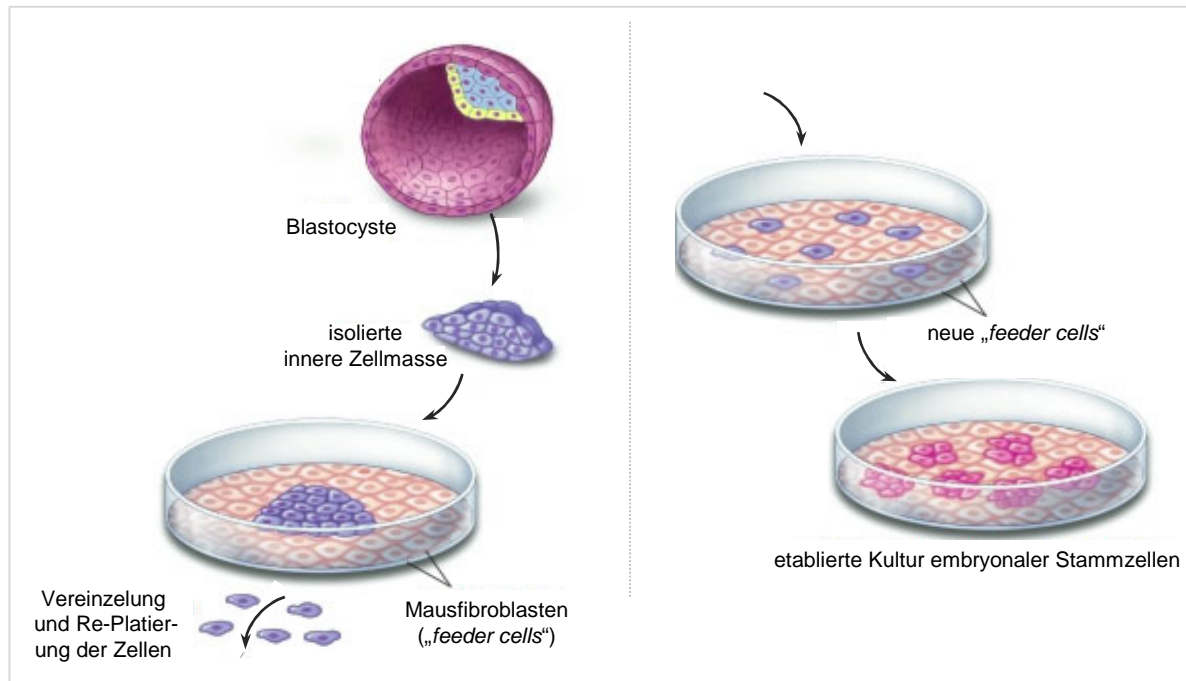


Abbildung 1.2-2: Etablierung einer Zellkulturlinie aus embryonalen Stammzellen. Schematische Darstellung der Isolierung und Kultivierung von Zellen der inneren Zellmasse einer Blastozyste. Die Differenzierung wird in diesem Beispiel durch Faktoren gehemmt, welche von ko-kultivierten embryonalen Mausfibroblasten produziert werden (<http://stemcells.nih.gov/StaticResources/info/scireport/images/figurec1.jpg>).

Humane ESC (hESC) sind - nicht zuletzt aufgrund der ethischen Einschränkungen hinsichtlich der Manipulation menschlicher Embryonen - weniger gut charakterisiert. Interessanterweise zeigen sie wesentliche Unterschiede zu murinen ESC und gleichen eher Stammzellen der Post-Implantations-Phase (EpiSC). Warum hESC, obwohl aus der inneren Zellmasse von Blastozysten isoliert, eher mit dem „vorbereiteten“ als dem naiven Pluripotenzstadium übereinstimmen, ist noch nicht verstanden (siehe auch Abschnitt 1.2.3). Jedoch könnten suboptimale Kulturbedingungen für diese Diskrepanz verantwortlich sein (Yu und Thomson 2008, Hanna et al. 2010).

Des Weiteren werden Stammzellen unterschiedlichen Differenzierungspotentials durch **Reprogrammierung** erhalten. In dieser Hinsicht konnte 2006 mit der Generierung pluripotenter Zellen aus adulten Maus-Fibroblasten ein bedeutender Durchbruch erzielt werden (Takahashi und Yamanaka 2006). Die sogenannten iPS-Zellen (induzierte pluripotente Stammzellen) wurden dabei durch retrovirale Überexpression definierter

Transkriptionsfaktoren (Oct4, Sox2, Klf4, cMyc - „Yamanaka-Cocktail“) erzeugt. Die aktuelle Forschung beschäftigt sich mit einer Optimierung der Reprogrammierungseffizienz (0,01-0,1%) und -Technik sowie einer Beurteilung der erhaltenen Zellen im Vergleich zu ESC (Hanna et al. 2010).

Die Bedeutung einer gründlichen **Charakterisierung von Stammzelllinien**, welche als *in vitro* Modell zum Verständnis der Embryonalentwicklung dienen sollen, wurde durch aktuelle Studien hervorgehoben. Neben den bereits bekannten Variabilitäten zwischen verschiedenen Linien und Passagen (insbesondere im Falle von hESC) (Pannetier und Feil 2007) weisen ESC im Vergleich zu frühen Embryonen aberrante epigenetische Modifikationen auf, welche möglicherweise auf Kultur-Bedingungen oder den genetischen Hintergrund zurückgeführt werden können. Dabei waren in ESC DNA-Methylierungsmuster und Marker konstitutiven Heterochromatins (H4K20me3 und HP1a) zu verzeichnen, welche *in vivo* für spätere Entwicklungsphasen spezifisch sind (Borgel et al. 2010, Wongtawan et al. 2011).

Zudem haben verschiedene Studien gezeigt, dass ESC interessanterweise keine homogene Population darzustellen scheinen. Anhand variabler Expressionsmuster einzelner Gene (Nanog, Stella, Pecam1, Rex1) konnten in murinen ESC Subpopulationen identifiziert werden, welche offenbar verschiedene Entwicklungsphasen repräsentieren. Dabei konnten dynamische Übergänge zwischen ICM- und Epiblast-ähnlichen Phänotypen und damit verschiedenen Pluripotenz-Stadien beobachtet werden (siehe auch Abschnitt 1.2.3) (Chambers et al. 2007, Hayashi et al. 2008, Toyooka et al. 2008, Nichols und Smith 2009, Albert und Peters 2009). In dieser Hinsicht scheint auch ein Zusammenhang mit Histon-Modifikationen nicht aber der DNA-Methylierung zu bestehen (Hayashi et al. 2008).

Neben einer Beurteilung der Morphologie, Genexpression und epigenetischen Profile dienen **funktionelle Tests** einer Einschätzung der Pluripotenz und des Differenzierungspotentials verfügbarer Stammzelllinien (und geben z.B. Hinweise auf unvollständige Reprogrammierung). Dazu zählen insbesondere: i) Fähigkeit zur *in vitro* Differenzierung in Zelllinien aller drei Keimblätter durch Bildung von Zellaggregaten („embryoid body“ (EB) ~ embryonale Körperchen) (Itskovitz-Eldor et al. 2000, Smith et al. 2009); ii) Generierung chimärer, Keimbahn-kompetenter Organsimen nach Einbringen von Zellen in Blastozysten (Gold-Standard Test) (Okita et al. 2007, Smith et al. 2009); iii) Bei Einbringen der Zellen in immunodefiziente Mäuse (z.B. mittels subdermaler oder intramuskulärer Injektion) Bildung Teratoma-artiger Strukturen mit Zellen der ektodermalen, endodermalen und mesodermalen Linien (Smith et al. 2009) und iv) Erzeugung einer 4n-Blastozyste durch

Fusion somatischer Zellen und Einbringen pluripotenter Zellen, welche alleinig zur Bildung des resultierenden Organismus beitragen (tetraploide Komplementation) (Nagy et al. 1993, Smith et al. 2009).

Bisher bestehen keine einheitlichen Standards zur Kontrolle der Pluripotenz. Oftmals wird aus Zeitgründen nur eine Auswahl der möglichen Tests vorgenommen oder nur eine begrenzte Anzahl an Keimblättern bzw. daraus entstehender Zelltypen werden überprüft (Smith et al. 2009).

Zwei im Zuge dieser Arbeit näher untersuchte Stammzelltypen sollen im Folgenden kurz vorgestellt werden.

1.2.2 gPS: aus der Keimbahn abstammende pluripotente Stammzellen

Keimbahnstammzellen (GSC: „*germline stem cell*“) sind unipotente Vorläuferzellen der Spermatogenese. Aufgrund ihres eingeschränkten Entwicklungspotentials haben sie unter natürlichen Bedingungen nur die Fähigkeit zur Selbsterneuerung sowie zur Bildung von Spermien. Dazu ist offensichtlich die, wenn auch geringe, Expression von Oct4 erforderlich (Dann et al. 2008). Die unter den Hodenzellen in geringem Maße vorkommenden GSC (0,02-0,03%) (Tegelenbosch und de Rooij 1993) können einzeln isoliert und als Zelllinie vermehrt werden (Kanatsu-Shinohara et al. 2003, Kubota et al. 2004).

Mehrere Studien haben sich in den letzten Jahren mit der Generierung ESC-ähnlicher Zellen aus Hodenzellen adulter oder neugeborener Mäusen beschäftigt (Guan et al. 2006, Izadyar et al. 2008, Kanatsu-Shinohara et al. 2003, Kanatsu-Shinohara et al. 2008, Seandel et al. 2007). Die tatsächliche Abstammung dieser Zellen wurde allerdings angezweifelt und auch die Eigenschaften pluripotenter Zellen wurden nicht immer erfüllt (z.B. Teratoma-Bildung).

In der Studie von Ko und Mitarbeitern (Ko et al. 2009) konnte erstmals die Generierung pluripotenter Zellen aus unipotenten GSC der adulten männlichen Keimbahn der Maus gezeigt werden. Kolonien der sogenannten gPS („*germline-derived pluripotent stemcell*“) bildeten sich bei Kultivierung von GSC spontan innerhalb von 4 Wochen mit einer Effizienz von 0,01%. Ausschlaggebend für diese Konvertierung, welche im Vergleich zu somatischen Zellen ohne Zugabe von Reprogrammierungsfaktoren erfolgte, waren offenbar spezielle Kultivierungsbedingungen. Dabei war insbesondere eine geringe Anzahl/Dichte plattierter Ursprungszellen von Bedeutung.

Mit Hilfe klonaler GSC-Linien konnte einwandfrei die Abstammung der gPS von GSC bewiesen werden. Dies wurde zudem durch die spezifischen Methylierungsmuster der imprinteten Gene H19 und Igf2r verdeutlicht. Die auf eine epigenetische Reprogrammierung in der Keimbahn zurückzuführenden androgenetischen Methylierungsmuster der GSC waren auch in gPS erkennbar. Eine nähere Charakterisierung der gPS konnte die große Ähnlichkeit dieser Zellen zu ESC hinsichtlich Morphologie, Expression und Differenzierungspotential aufweisen. Anhand verschiedener *in vivo* und *in vitro* Tests (Abschnitt 1.2) wurde die Pluripotenz der gPS, im Gegensatz zu GSC, dargestellt (*in vitro* Differenzierung in Herz-, Nerven- und Endothelzellen, Teratoma-Test, Bildung von Chimären). Einzig die tetraploide Komplementation ergab negative Resultate, was auf die androgenetischen Methylierungsmuster imprinteter Gene zurückgeführt wird (Ko et al. 2009).

1.2.3 EpiSC: Stammzellen des Epiblasten

Nach Einnisten des frühen Säugerembryos in den Uterus segregiert die innere Zellmasse in den Hypoblasten und den Epiblasten. Während der Hypoblast später u.a. den Dottersack, das Ernährungsorgan des Embryos, bildet, entwickelt der Epiblast sich zum eigentlichen Embryo. Im Zuge der Gastrulation differenziert er zu den drei Keimblättern (Endoderm, Ektoderm, Mesoderm), aus welchen während der folgenden Entwicklung Organe und Gewebe entstehen.

Durch Explantation von Epiblast-Zellen 5,5-7,5 Tage nach Fertilisation können pluripotente Zellen des Post-Implantations-Stadiums ebenfalls als stabile Zelllinien etabliert werden (EpiSC: „*Epiblast stemcell*“) (Brons et al. 2007, Tesar et al. 2007). Obwohl die EpiSC einige Merkmale pluripotenter Zellen aufweisen (z.B. Expression der TFs Oct4, Sox2), konnten molekulare und funktionelle Analysen auch wesentliche Unterschiede zu ESC aufdecken. Dazu zählen distinkte, die fortgeschrittene Entwicklungsphase des Epiblasten repräsentierende Expressionsmuster (mRNA und miRNA), epigenetische Profile (Chromatin-Modifikationen und DNA-Methylierung in Einzelgenanalysen, z.B. Stella) und Signalwege zur Steuerung der Selbst-Erhaltung bzw. Differenzierung (Chou et al. 2008, Tesar et al. 2007, Han et al. 2010, Hayashi et al. 2008, Nichols und Smith 2009). EpiSC sprechen in dieser Hinsicht bei Kultivierung nicht auf ESC-spezifische Faktoren wie Lif und Bmp4 an, sondern benötigen FGF2 und Aktivin, welche in ESC zur Differenzierung führen (Tesar et al. 2007, Brons et al. 2007). Des Weiteren scheint das Potential der EpiSC im Vergleich zu ESC eingeschränkt zu sein, da sie - trotz der Fähigkeit zur *in vitro* Differenzierung und

Teratoma-Bildung - offenbar nicht zu Chimären beitragen können (Guo et al. 2009, Tesar et al. 2007) (Abbildung 1.2-3). Dieses durch EpiSC repräsentierte Stadium, in welchem die Zellen offenbar für die weitere Spezifikation in Antwort auf Signale der extraembryonalen Gewebe gerüstet sind, wird mit Hilfe des Begriffs „vorbereitete“ („primed“) Pluripotenz von den ESC (naive Pluripotenz) abgegrenzt (Nichols und Smith 2009).

Neuere Studien konnten zeigen, dass EpiSC, ebenso wie ESC, eine heterogene Population darstellen. Dabei konnten anhand von Expressions- und DNA-Methylierungsmustern einzelner Gene (z.B. Oct4, Dppa5) zwei Subpopulationen unterschieden werden, welche frühe und späte Entwicklungsstadien des Epiblasten (*in vivo*) zu repräsentieren scheinen. Den frühen Epiblasten widerspiegelnde Zellen, eine lediglich geringe Fraktion der EpiSC, konnten dabei zu Chimären beitragen (Han et al. 2010).

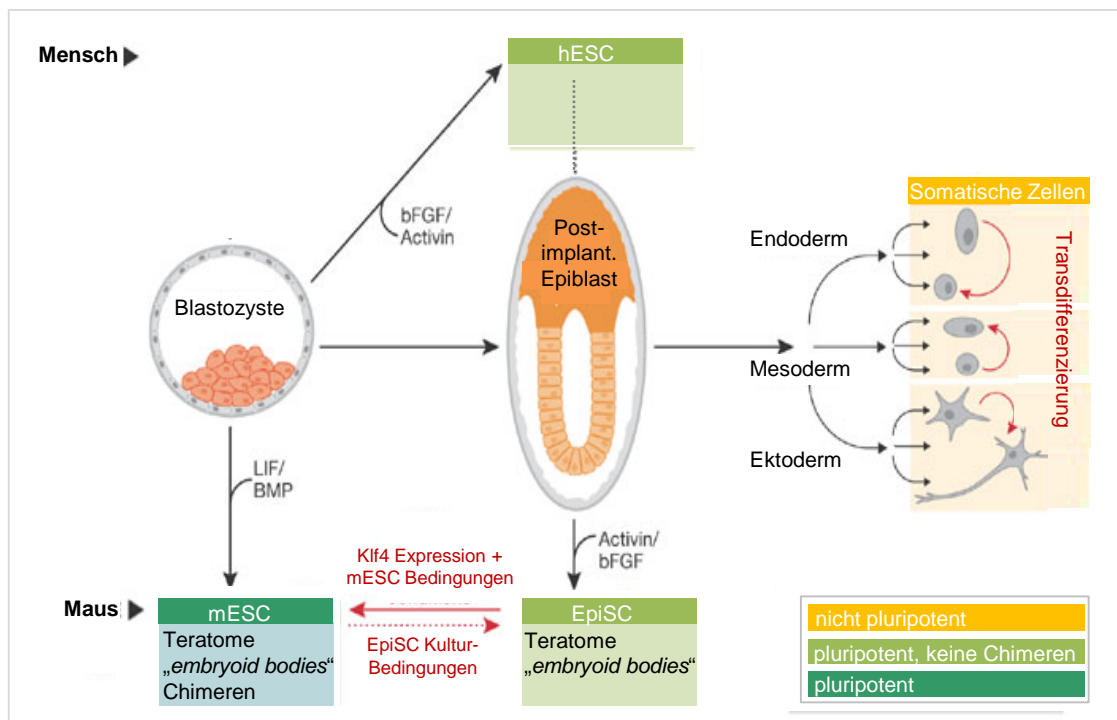


Abbildung 1.2-3: Stammzellen des Epiblasten im Vergleich zu ESC. Schematisch dargestellt sind verschiedene Stadien der Embryonalentwicklung und die Gewinnung entsprechender Stammzelllinien in Maus und Mensch mit unterschiedlichem Entwicklungspotential (siehe auch Farbkodierung der Legende). Kulturzellen des Post-Implantations-Epiblasten (EpiSC) können im Gegensatz zu Kulturzellen der frühen Blastozyste (mESC) nicht zur Chimären-Bildung beitragen und repräsentieren ein Stadium der eingeschränkten/vorbereiteten Pluripotenz. In ihren Kultivierungsbedingungen gleichen EpiSC humanen ESC (hESC) (verändert und vereinfacht nach Colman und Dreesen 2009).

EpiSC sind nicht nur als *in vitro* Modell zur Erforschung der Pluripotenz und Spezifikation des Zellschicksals während der Säugereentwicklung von Interesse. Tatsächlich stand

zunächst die hohe Ähnlichkeit mit humanen ESC bzgl. Morphologie, Expression und Signalwegen zur Regulation der Selbsterneuerung bzw. Differenzierung (und damit den Kulturbedingungen) im Vordergrund - Eigenschaften, in welchen humane ESC sich wesentlich von murinen ESC unterscheiden (Tesar et al. 2007, Brons et al. 2007, Nichols und Smith 2009, Yu und Thomson 2008). Kürzlich wurde jedoch gezeigt, dass FGF2 zwar in hESC und EpiSC zur Aufrechterhaltung der Pluripotenz und Selbst-Erneuerung in Kultur beitragen, die biochemischen Signalkaskaden sich aber unterscheiden (Kontrolle der Nanog-Expression vs. Unterdrückung von Reifungsprozessen)(Greber et al. 2010).

EpiSC können *in vitro* auch durch entsprechende Kultivierung von ESC (Guo et al. 2009) oder Reprogrammierung somatischer Zellen (iEpiSC) (Han et al. 2011) erhalten werden, wodurch sich weitere Erkenntnisse zur Charakterisierung naiver und vorbereiteter Stadien der Pluripotenz bieten.

1.3 Kryokonservierung

Kryopreservation, die Aufbewahrung biologischen Materials bei Temperaturen von -80 bis -196°C, gilt heutzutage als Standardverfahren zur Konservierung lebender Zellen und Gewebe und findet weitreichende Anwendungen im medizinischen Bereich. Dabei ist nicht nur die Erhaltung fortpflanzungsfähiger Zellen von Bedeutung sondern ebenso eine Bewahrung verschiedener Organgewebe.

Die Lagerung des Materials in flüssigem Stickstoff bzw. dessen kalter Gasphase ermöglicht eine beinahe unbegrenzte Erhaltung der Vitalität des lebenden Materials. Dies ist in der Tatsache begründet, dass das biologische System in den Aggregatzustand eines Festkörpers übergeht und Stoffwechselfvorgänge sowie biochemische Abläufe zum Stillstand kommen.

Insbesondere während des Einfrierens und Auftauens kommt es durch **physikalische und biochemische Kräfte** zu einer Schädigung der Zellen. Als Hauptursache dafür gilt gemäß der Zweifaktoren-Hypothese (Mazur 1963, Mazur et al. 1972, Gao und Critser 2000) der Phasen-Wechsel des Wassers. Während des Abkühlens (-5 bis -15°C) entsteht ein intrazellulärer osmotischer Druck, der zu einem Abstrom von Zellwasser in den extrazellulären Raum führt. Bei einer sehr langsamen Abkühlung führt das Gefrieren extrazellulärer Flüssigkeit zu einem weiteren Anstieg des osmotischen Drucks und einer Dehydrierung der Zelle. Neben einer mechanischen Schädigung durch die Eiskristalle selbst, ist auch die Konzentration gelöster Substanzen infolge der Dehydrierung und Eisbildung („*solution-effect*“) nachteilig. Eine zu schnelle Abkühlung hingegen führt, da

das Wasser die Zelle nur langsam verlässt, zur Bildung letaler intrazellulärer Eiskristalle (Mazur 1963, Mazur et al. 1972) (Abbildung 1.3-1 (A)).

Die **optimale Kühlrate** stellt somit einen kritischen Faktor zur Minimierung von Zellschädigungen dar. Noch immer wird Forschung zur Ermittlung der idealen Einfrier-Bedingungen betrieben, welche sich von Zelltyp zu Zelltyp unterscheiden. Dies ist hauptsächlich in den spezifischen Membran-Permeabilitäten begründet (Mazur 1984, Gao und Critser 2000) (Abbildung 1.3-1 (B)).

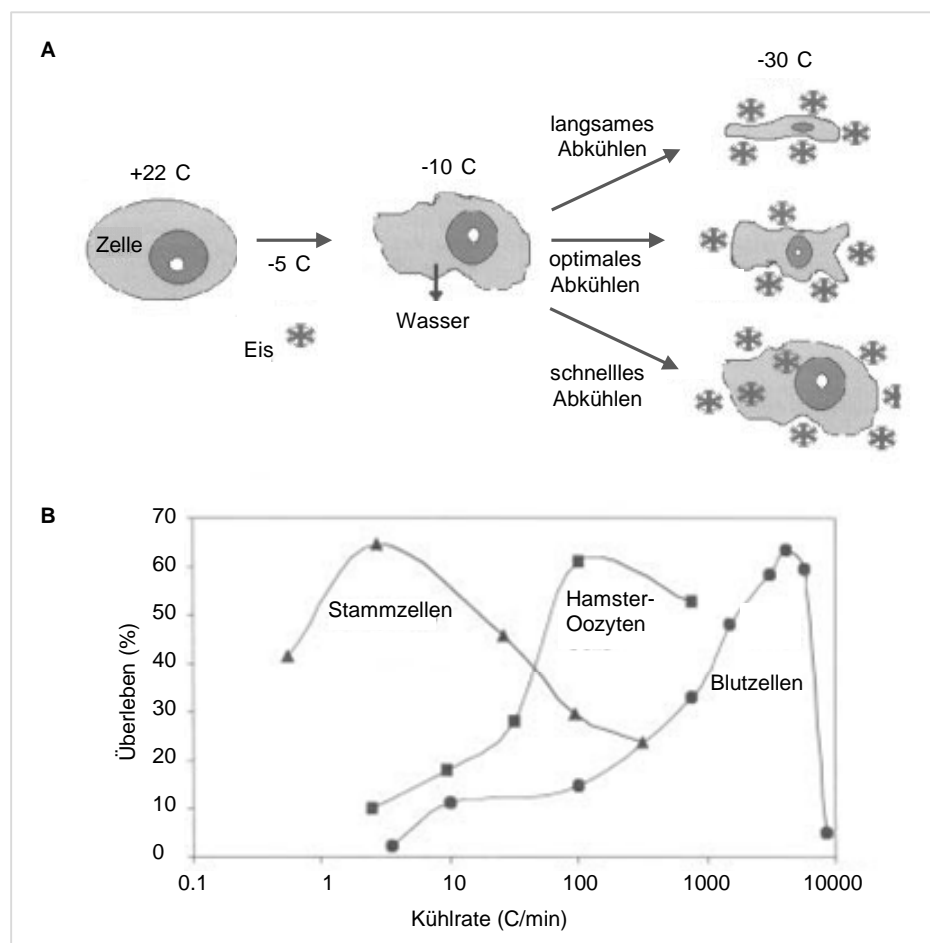


Abbildung 1.3-1: Zellschädigung in Abhängigkeit der Kühlrate bei Kryokonservierung. (A) Schematische Darstellung der Zellschädigungen aufgrund zu langsamer und zu schneller Kühlraten durch Dehydrierung und Konzentrationseffekte bzw. intrazelluläre Eiskristallbildung. (B) Anteil überlebender Zellen in Abhängigkeit der Kühlrate drei verschiedener Zelltypen (Stammzellen des Knochenmarks, Hamster Oocyten und rote Blutzellen) (Abbildungen verändert nach Gao und Critser 2000).

Eine weitere Minimierung von Zellschäden wird durch den Einsatz von Gefrierschutzmitteln (**Kryoprotektiva**) erreicht. Penetrierende Substanzen wie z.B. Glycerol und Dimethylsulfoxid (DMSO) verlangsamen die Dehydrierung der Zelle und verringern somit

schädigende Effekte aufgrund sich verändernder Solut-Konzentrationen. Zudem führen sie zur Herabsetzung der Temperatur, bei welcher intrazelluläre Eiskristalle gebildet werden. Allerdings kommt es durch die Kryoprotektiva selbst, insbesondere während des Auftauens, zu osmotischen und toxischen Effekten (Mazur und Schneider 1986, Wusteman et al. 2002, Meryman 2007, Pegg 2002). Aus diesem Grund ist eine vorsichtige Equilibrierung des Materials vor dem Einfrieren und eine gründliche Ausverdünnung während des Auftauens von Bedeutung (Meryman 2007).

Als **weitere schädliche Einflüsse** konnten u.a. eine elektrische Ladungstrennung der Zellmembranen (Steponkus et al. 1985) sowie oxidativer Stress durch vermehrt gebildete reaktive Sauerstoffradikale (ROS: „*reactive oxygen species*“) (Niu et al. 2010, Rauen et al. 1999, Ahn et al. 2002) identifiziert werden. In diesem Zusammenhang konnte in Spermazellen auch eine Schädigung/Degradation der DNA beobachtet werden (Thomson et al. 2009, Zribi et al. 2010). Ein Effekt der während der Kryokonservierung/Hypothermie wirkenden Stressfaktoren auf die Genexpression verschiedener Zelltypen wurde bereits in mehreren Studien gezeigt. Bei den zumeist transient auftretenden Expressions-Änderungen handelt es sich hauptsächlich um eine Aktivierung bekannter Stress-Regulatoren wie z.B. Hitzeschockproteine (HSP) (Liu et al. 2000, Sonna et al. 2006).

1.4 Epigenetik und Stress

1.4.1 Kryokonservierung, abiotischer Stress und Epigenetik

Abschnitt 1.3 verdeutlicht, dass bei einem Prozess wie der Kryokonservierung eine erhebliche Belastung des biologischen Materials aufgrund thermaler, osmotischer und toxischer Effekte auftritt. Einige Erkenntnisse hinsichtlich einer Beeinflussung epigenetischer Modifikationen in Zusammenhang mit solchem abiotischen Stress konnten bereits durch verschiedene Studien an Pflanzen und Säugern erbracht werden.

Zum Einen wurde eine veränderte Expression von **Schlüsselfaktoren für die Aufrechterhaltung/Setzung epigenetischer Muster** (in Pflanzen und Säugern) beobachtet, wodurch indirekt auch die epigenetischen Profile selbst verändert wurden. Chen und Mitarbeiter konnten eine Hypoxie-induzierte Zunahme von H3K9me2 in Verbindung mit gesteigerter Aktivität der Histon-Methyltransferase G9a in verschiedenen Säugerzelllinien feststellen (Chen et al. 2006). Eine bei längerer Inkubation mit toxischem DMSO zu verzeichnende Demethylierung einzelner Gene wurde zudem auf eine veränderte Expression von Dnmt3a zurückgeführt (Iwatani et al. 2006).

Des Weiteren liefert die Literatur einige Hinweise auf direkte Stress-induzierte Effekte auf die chemischen Modifikationen selbst, ohne Kenntnis der genauen Mechanismen. Insbesondere in Pflanzen-Studien konnten Änderungen der DNA-Methylierung mit verschiedenen Umwelteinflüssen (Temperatur, Osmose, Aluminium, Schwermetalle etc.) in Verbindung gebracht werden (Chinnusamy und Zhu 2009). Dabei wurde Hypomethylierung im Zusammenhang mit u.a. Salz-Stress (Weizen, Mais) (Zhong et al. 2009, Tan 2010), Hypothermie (Mais) (Steward et al. 2002, Steward et al. 2000) und Kryopreservation (Papaya, Hopfen) (Kaity et al. 2008, Peredo et al. 2008) festgestellt.

Entsprechende Erkenntnisse zu Stress-induzierten epigenetischen Veränderungen in Säugern stammen aus Studien zur Kryopreservation/Vitrifikation von Oozyten und frühen Embryonen. Dabei wurden z.B. Änderungen spezifischer **Histonmodifikationen** (Suo et al. 2010, Yan et al. 2010), nämlich ein Anstieg der H3K9me, H4K5ac und AcH4K12, verzeichnet. Einige Studien, welche sich mit der **DNA-Methylierung** einzelner Gene in Zusammenhang mit Kryopreservation beschäftigten, lieferten gegensätzliche Resultate. Während zwei Publikationen keine veränderte Methylierung ausgewählter Imprinting-Center in vitrifizierten Maus-Oozyten feststellen konnten (H19-DMR, KvDMR) (Trapphoff et al. 2010, Al-Khtib et al. 2011), wurde an anderer Stelle über eine entsprechende Demethylierung (H19-DMR) in vitrifizierten Maus-Embryonen berichtet (Wang et al. 2010). Zudem konnte eine, wenn auch geringe, Demethylierung (4-25%) der Pluripotenzgene Oct4 und Sox2 detektiert werden (vitrifizierte Oozyten) (Milroy et al. 2011).

1.4.2 Oxidativer Stress und DNA-Methylierung

Unter oxidativem Stress versteht man den Zustand einer Zelle, bei welchem physiologische Mengen reaktiver Sauerstoffverbindungen (ROS) überschritten werden.

ROS wie Superoxid-Anionenradikale (O_2^-), Wasserstoffperoxide (H_2O_2) und Hydroxylradikale (OH^\cdot) werden als Nebenprodukt im Zuge natürlicher Stoffwechselforgänge gebildet. Zelluläre Schutzmechanismen halten dabei durch Entgiftungsreaktionen ein „natürliches“ Gleichgewicht aufrecht. Enzymatische und nicht-enzymatische Radikalfänger und Antioxidantien, wie Superoxid-Dismutase (SOD), Glutathion-Peroxidase, Katalase und Glutathion (GSH), führen dabei zum Abbruch radikalischer Reaktionsketten. Eine Störung dieser Mechanismen bzw. eine durch Umwelteinflüsse verursachte Überproduktion der freien Radikale führt folglich zur Schädigung zahlreicher Zellstrukturen (Finkel und Holbrook 2000) (Abbildung 1.4-1 (A)).

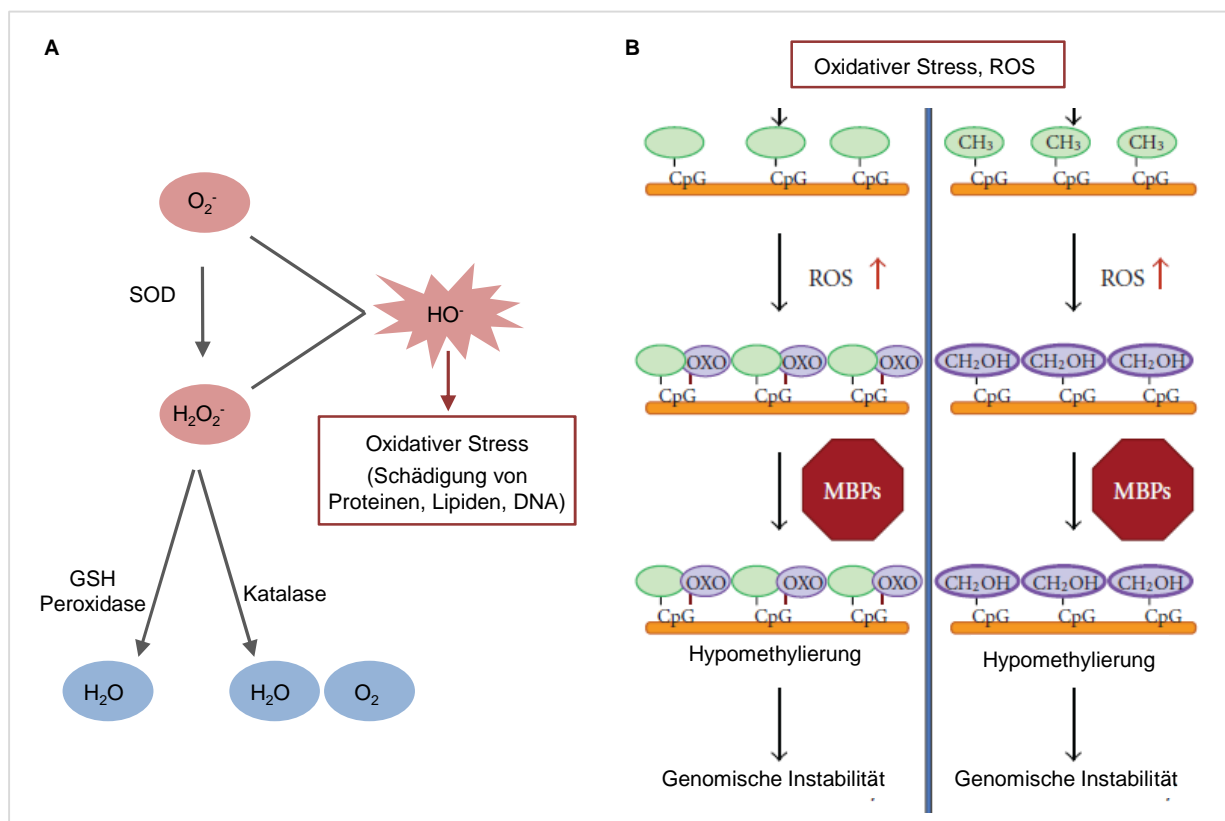


Abbildung 1.4-1: Oxidativer Stress, zelluläre Schutzmechanismen und DNA-Methylierung. (A) Verschiedene Sauerstoffradikale ($O_2^{\cdot-}$, H_2O_2 und HO^{\cdot}) werden durch enzymatische und nicht-enzymatische Radikalfänger und Antioxidantien, wie Superoxid-Dismutase (SOD), Glutathion-Peroxidase (GSH-Peroxidase), Katalase und Glutathion (GSH) in Wasser (H_2O) und Sauerstoff (O_2) umgewandelt. Bei einer Störung des Gleichgewichts können überschüssige Radikale zu oxidativen Schädigungen führen. (B) Modell des Einflusses erhöhter Mengen reaktiver Sauerstoffradikale (ROS) auf die DNA-Methylierung, wobei Hydroxylierung von Guanin oder Cytosin im CpG-Kontext die Interaktion mit Methyl-bindenden Faktoren (MBPs) bzw. Methyltransferasen behindern und so zu einer passiven Demethylierung sowie einem Verlust der genomischen Integrität führen (verändert nach <http://www.scielo.br/img/revistas/abo/v71n6s0/15f1.gif> (A) und Donkena et al. 2010 (B)).

Verschiedene DNA-Läsionen, wie Strangbrüche und Basen-Modifikationen, werden insbesondere durch Hydroxyl-Radikale (OH^{\cdot}) hervorgerufen und können u.a. auch auf epigenetische Ebene agieren. Eine Reihe von Krebsstudien befasst sich mit dem Zusammenhang zwischen oxidativem Stress und DNA-Methylierung. Dabei wird v.a. die **ROS-vermittelte Oxidation von Guanin und methyl-Cytosin** für eine globale Hypomethylierung verantwortlich gemacht, wobei die genauen Mechanismen noch nicht verstanden sind (Donkena et al. 2010). Einige Daten deuten allerdings darauf hin, dass Oxidationsprodukte wie 8-Hydroxyguanin, 8-Oxoguanin und 5-Hydroxymethylcytosin die Interaktion mit Methyl-bindenden Faktoren wie MeCP2 bzw. mit der Erhaltungsmethyltransferase DNMT1 beeinträchtigen (Valinluck et al. 2004, Valinluck und Sowers 2007, Turk et al. 1995, Weitzman et al. 1994). Dies könnte den Theorien zufolge

nicht nur zu einer passiven Demethylierung sondern aufgrund veränderter Chromatin-Kondensationen auch zum Verlust der genomischen Stabilität/Integrität führen (Donkena et al. 2010) (Abbildung 1.4-1 (B)).

1.5 Methoden zur Analyse von DNA-Methylierung

Technische Entwicklungen der letzten Jahre haben zu einer Auswahl verschiedener Methoden geführt, um genomweite DNA-Methylierungsmuster zu generieren und damit unterschiedliche Zelltypen und Entwicklungsstadien genauer charakterisieren zu können. Die Ansätze werden dabei mit Microarray-Hybridisierung oder Sequenzierung kombiniert und erreichen unterschiedliche Level der Abdeckung, Sensitivität und Auflösung.

Als Gold-Standard gilt dabei die Sequenzierung Bisulfit-behandelter DNA, welche quantitative Methylierungswerte in Basenpaar-Auflösung liefert. Die Methode basiert auf einer chemischen Konvertierung unmethylierter Cytosine, welche in den Folgeschritten eine direkte Unterscheidung methylierter und unmethylierter Positionen ermöglicht (Clark et al. 1994, Clark et al. 2006, Frommer et al. 1992). Der ehemals zeitaufwändige und Locus-spezifische Ansatz (Eckhardt et al. 2006, Zhang et al. 2009) wurde durch die Möglichkeit der Hochdurchsatz-Sequenzierung revolutioniert. Neben zielgerichteten Strategien, z.B. durch Einsatz von Restriktionsenzymen (RRBS, „*reduced representation bisulfite sequencing*“), wurden in den vergangenen Jahren die Methyloome verschiedener Spezies (Säuger, Pflanzen) mittels shotgun Sequenzierung bisulfitierter genomischer DNA erstellt (Meissner et al. 2005, Meissner et al. 2008, Stadler et al. 2011, Lister et al. 2009, Lister et al. 2011, Hodges et al. 2011, Lister et al. 2008, Cokus et al. 2008). Allerdings erfordert dieser Ansatz, nicht zuletzt aufgrund der durch Konvertierung degenerierten Sequenz, eine enorme Sequenzierleistung und stellt damit eine kostenintensive Methode dar. Des Weiteren ist keine Unterscheidung zwischen 5mC und hmC möglich (Huang et al. 2010, Jin et al. 2010).

Alternative Ansätze, kombiniert mit Microarrays oder Sequenzierung, basieren auf einer Selektion mittels Restriktionsenzymen oder Methyl-bindender Antikörper bzw. Proteine. Zu den bekanntesten Verfahren zählen hierbei: i) Methylierungssensitive bzw. -spezifische Restriktion, z.B. DMH- (Huang 1999), HELP- (Khuan 2006), CHARM-Methode (Irizarry 2008)); ii) Anreicherung methylierter DNA mittels Methyl-bindender Proteine, z.B. MeCP2, MBD2 (Cross et al. 1994, Gebhard et al. 2006b, Gebhard et al. 2006a, Jorgensen et al. 2006) und iii) Immunopräzipitation methylierter DNA (MeDIP bzw. mDIP) (Weber et al. 2005, Keshet et al. 2006). Obwohl diese Methoden aufgrund der Abhängigkeit von

Restriktionsschnittstellen bzw. CpG-Dichte einen reduzierten Informationsgehalt und eher qualitative als quantitative Daten bieten, stellen sie dennoch eine kostengünstige Alternative für eine (erste) Charakterisierung der DNA-Methylierung dar (Harris et al. 2010, Bock et al. 2010). Höherauflösende Methoden (wie Bisulfit-Sequenzierung) werden dann oftmals zur Validierung bzw. tieferen Analyse spezifischer Regionen von Interesse eingesetzt.

1.6 Aufgabenstellung

Die Immunopräzipitation methylierter DNA (MeDIP) stellt neben ähnlichen Ansätzen eine inzwischen gängige und kostengünstige Möglichkeit zur Analyse genomweiter Methylierungsprofile dar. Ziel der vorliegenden Arbeit war die laborinterne Etablierung und Überprüfung der Methode MeDIP-Chip inklusive bioinformatischer Analysen sowie deren Anwendung im Zuge zweier unterschiedlicher biologischer Fragestellungen, welche in getrennten Abschnitten abgehandelt werden.

1.6.1 Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung

Das erste Ziel der vorliegenden Arbeit bestand in der laborinternen Etablierung der Immunopräzipitations-basierten Methode MeDIP-Chip (Weber et al. 2005), für welche zu jenem Zeitpunkt kaum Erfahrungswerte vorlagen. Neben der Ermittlung der optimalen Präzipitationsrate sowie -Qualität sollte eine geeignete bioinformatische Daten-Analyse, insbesondere zur Erfassung chromosomenweiter Methylierungsmuster, erarbeitet werden. Mit Hilfe quantitativer Daten aus Bisulfit-Sequenzierung von >260 Promotoren verschiedener Gewebe (Zhang et al. 2009) sollten anschließend i) Datenqualität, Reproduzierbarkeit und Güte des analytischen Ansatzes bewertet sowie ii) die Anwendbarkeit der genomrepräsentativen Amplifikation, einer zu Verzerrungen neigenden Methode (Jia et al. 2010, Robinson et al. 2010, Johnson et al. 2008), kritisch überprüft werden.

1.6.2 Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryopreservation

Kryopreservation gilt zwar heutzutage als Standardverfahren im medizinischen und biotechnologischen Bereich zur Erhaltung lebender Zellen und Gewebe, jedoch ist der dabei verursachte zelluläre Stress noch nicht gänzlich verstanden. Abgesehen von physikalischen Schädigungen und Beeinträchtigungen auf biochemischer Ebene (siehe auch Abschnitt 1.3), ist wenig über potentielle sublethale Effekte bekannt, die zu Langzeitveränderungen der Zellen führen könnten. Epigenetische Signaturen stellen hierbei einen möglichen Angriffspunkt dar. Änderungen epigenetischer Modifikationen konnten bereits mit unterschiedlichen Umwelteinflüssen, einer Reihe von Erkrankungen (z.B. Krebs) sowie abiotischem und oxidativem Stress in Verbindung gebracht werden (Feil 2006, Esteller und Herman 2002, Robertson 2005, Yan et al. 2010, Suo et al. 2010, Wang et al. 2010, Milroy et al. 2011, Zhu et al. 2008, Hitchler und Domann 2009, Donkena et al. 2010) (siehe auch Abschnitt 1.4). Die im Zuge der Kryopreservation vermehrt gebildeten Sauerstoffradikale (Niu et al. 2010, Rauen et al. 1999) bzw. der Einsatz kryoprotektiver Substanzen könnten also direkt oder indirekt vererbare epigenetische Veränderungen verursachen, wobei eine Verbindung zu der bei Kälte- und Hitzeschock beobachteten deregulierten Genexpression denkbar wäre (Liu et al. 2000, Sonna et al. 2010, Sonna et al. 2006). Die Analyse einer potentiellen Stress-induzierten Epimutagenese und damit vererbaren phänotypischen Veränderung infolge von Kryopreservation ist damit von großer Bedeutung im Hinblick der Integrität des biologischen Materials. In der vorliegenden Arbeit sollten daher i) chromosomenweite DNA-Methylierungsmuster kryokonservierten Materials mittels MeDIP-Chip untersucht und ii) durch Vergleich mit mRNA-Expressionsdaten ein möglicher Einfluss auf den Phänotyp abgeschätzt werden. iii) Zusätzliche Analysen mittels Bisulit-Sequenzierung sollten zudem einer Validierung sowie Vervollständigung der erhaltenen Resultate dienen.

1.6.3 Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen

Neben ihrem potentiellen therapeutischen Nutzen bieten Stammzellen die Möglichkeit, das Grundverständnis entwicklungspezifischer Vorgänge zu erweitern. Als hilfreiche Modelle dienen dabei verschiedene in den vergangenen Jahren generierte Stammzelllinien, welche unterschiedliche Stadien der Entwicklung repräsentieren.

In diesem Zusammenhang sollten in der vorliegenden Arbeit murine Stammzelltypen unterschiedlichen ontogenetischen Alters bzw. Differenzierungspotentials hinsichtlich ihrer genomweiten DNA-Methylierungsprofile mittels MeDIP-Chip analysiert werden. Dabei standen zwei bisher epigenetisch kaum charakterisierte Stammzelllinien im Mittelpunkt. i) Zum Einen sollten die aus Keimbahnstammzellen (GSC) generierten pluripotenten gPS untersucht werden (Ko et al. 2009). Mit Hilfe der gewonnenen Methylierungsdaten sollte das Ausmaß der Reprogrammierung dieser ESC-ähnlichen Zellen auf epigenetischer Ebene abgeschätzt werden. ii) Des Weiteren sollte das Stadium der vorbereiteten Pluripotenz anhand von EpiSC (Han et al. 2010, Greber et al. 2010, Nichols und Smith 2009) im Vergleich zur naiven Pluripotenz (ESC) und terminal differenzierten Zellen (MEF) mittels der DNA-Methylierung näher beleuchtet werden.

2 Material und Methoden

2.1 Material

2.1.1 Chemikalien

Agar	Roth, Karlsruhe, Deutschland
Agarose (LE)	Biozym, Oldendorf, Deutschland
Ammoniumsulfat	Merck, Darmstadt, Deutschland
Ampicillin	Sigma-Aldrich, Taufkirchen, Deutschland
Aqua bidest. (Milli-Q-H ₂ O)	hauseigene Anlage, Millipore, Schwalbach, Deutschland
Borsäure	Roth, Karlsruhe, Deutschland
Bromphenolblau	Sigma-Aldrich, Taufkirchen, Deutschland
BSA (Rinderserum Albumin)	Roche, Mannheim, Deutschland
Chloroform	Roth, Karlsruhe, Deutschland
Desoxyribonukleosidtriphosphat (dNTPs)	Solis BioDyne, Tartu, Estland
Dinatriumhydrogenphosphat (Anhydrat)	Merck, Darmstadt, Deutschland
Dinatriumhydrogenphosphat (Dodecahydrat)	Merck, Darmstadt, Deutschland
1-4-Dioxan	Merck, Darmstadt, Deutschland
EDTA (Ethylen-diamin-tetra-essigsäure)	Sigma-Aldrich, Taufkirchen, Deutschland
Ethanol	Sigma-Aldrich, Taufkirchen, Deutschland
Ethidiumbromid	Roth, Karlsruhe, Deutschland
Glukose	Roth, Karlsruhe, Deutschland
Glykogen	Roche, Mannheim, Deutschland
Hefeextrakt	Roth, Karlsruhe, Deutschland
IPTG (Isopropyl-beta-D-thio-galactopyranosid)	Amersham, Little Chalfont, Großbritannien
Isoamylalkohol	Roth, Karlsruhe, Deutschland
Kaliumchlorid	Merck, Darmstadt, Deutschland
Kaliumhydrogenphosphat	Merck, Darmstadt, Deutschland
Magnesiumchlorid	Merck, Darmstadt, Deutschland
Magnesiumsulfat	Merck, Darmstadt, Deutschland
Natriumacetat	Roth, Karlsruhe, Deutschland
Natriumazid	Serva, Heidelberg, Deutschland
Natriumchlorid	Roth, Karlsruhe, Deutschland
Natriumdihydrogenphosphat (Monohydrat)	Merck, Darmstadt, Deutschland
Natriumdisulfit	Merck, Darmstadt, Deutschland
Natriumhydroxid (Pellets)	Grüssing, Filsum, Deutschland

Phenol	Roth, Karlsruhe, Deutschland
Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v), UltraPure™	Invitrogen (Life Technologies), Darmstadt, Deutschland
Radikalfänger (6-hydroxy-2,5,7,8-tetra- methylchroman-2-carbonsäure, 97%)	Sigma-Aldrich, Taufkirchen, Deutschland
Reference Dye (ROX) (aus Brilliant SYBR Green QPCR Core Reagent Kit)	Stratagene, La Jolla, USA
Saccharose	Roth, Karlsruhe, Deutschland
SDS (Sodiumdodecylsulfat)	Roth, Karlsruhe, Deutschland
SYBR Green I	BioWhittaker Molecular Application, Rockland, USA
Tris-Base	Roth, Karlsruhe, Deutschlan
Tris-HCl	Roth, Karlsruhe, Deutschland
Triton X100	Sigma-Aldrich, Taufkirchen, Deutschland
Trypton	Roth, Karlsruhe, Deutschland
Tween 20	Pharmacia Biotech, Uppsala, Schweden
Wasser (DNase-, RNase-frei)	Sigma-Aldrich, Taufkirchen, Deutschland
X-Gal	Eppendorf, Köln, Deutschland
Xylen-Cyanol	Chroma, Stuttgart, Deutschland

2.1.2 Lösungen und Puffer

Puffer zur Präparation genomischer DNA

Lösung A	25 mM EDTA 75 mM NaCl
Lösung B	10 mM EDTA 10 mM Tris-HCl (pH 8,0) 1% SDS

Puffer zur Immunopräzipitation methylierter DNA

Blockier-Puffer (10x)	1% BSA 0,1% Triton X100 0,05% NaN ₃ in 1x PBS Sterilfiltriert, Lagerung bei 4 °C.
IP-Puffer (10x)	100 mM Na-Phosphat-Puffer (pH 7,0) 1,4 M NaCl 0,5 % Triton X-100 Sterilfiltriert.
Na-Phosphat-Puffer (1M, pH 7,0) - Stammlösung 1 (2 M):	276 g NaH ₂ PO ₄ · H ₂ O ad 1 l Aqua bidest.

- Stammlösung 2 (2 M): 284 g Na_2HPO_4
ad 1 l Aqua bidest.
- Na-Phosphat-Puffer: 39 ml Stammlösung 1
61 ml Stammlösung 2
ad 200 ml Aqua bidest.

ProteinaseK-Puffer 50 mM Tris-Base (pH 8,0)
10 mM EDTA
0,5% SDS
Sterilfiltriert.

Lösungen zur Bisulfit-Behandlung von DNA

Natriumbisulfit-Lösung 1,9 g Natriumdisulfit (lichtgeschützt)
0,75 ml 2 M NaOH
2,5 ml Aqua bidest.
Bei 50 °C im Wasserbad lösen.

Dioxan-Radikalfänger-Lösung 0,098 g Radikalfänger (lichtgeschützt)
2,5 ml 1,4-Dioxan
Bei 50 °C im Wasserbad lösen.

Sonstige Puffer und Lösungen

Agarosegel-Färbelösung 0,5 µg/ml Ethidiumbromid in einfach deionisiertem Wasser

PCR-Puffer II (10x) 750 mM Tris-HCl (pH 8,8)
200 mM $(\text{NH}_4)_2\text{SO}_4$
20 mM MgSO_4
1% Tween 20

PBS-Puffer (Ca^{2+} -frei; Mg^{2+} -frei)

- Stammlösung 1 (10x) 29 g $\text{Na}_2\text{HPO}_4 \cdot 12 \text{H}_2\text{O}$
2 g KH_2PO_4
ad 1 l Aqua bidest.
- Stammlösung 2 (10x) 80 g NaCl
13,3 g KCl
ad 1 l Aqua bidest.
- Gebrauchslösung (1x) je 100 ml Stammlösung 1 und 2
ad 1 l Aqua bidest.

Probenpuffer 0,2% (w/v) Bromphenolblau
(für Agarosegele) 0,2% (w/v) Xylen-Cyanol
40% Saccharose

TBE-Elektrophorese-Puffer 40 mM Tris-Base
50 mM Borsäure

	1 mM EDTA
TE-Puffer (10x)	100 mM Tris-HCl (pH 8,0) 100 mM EDTA (pH 8,0) Lösung autoklavieren.

2.1.3 Kulturmedien

LB-Agar (1 l; pH 7,5)	10 g Trypton 5 g Hefeextrakt 10 g NaCl 1,5% Agar
-----------------------	---

Die Lösung wird autoklaviert. Zur Selektion wird diese später mit Antibiotikum (Ampicillin, 100µg/ml) sowie X-Gal (2%) und IPTG (0,1 M) versetzt.

SOC-Medium (1 l; pH 7,0)	20 g Trypton 5 g Hefeextrakt 0,5 g NaCl 10 ml KCl (250 mM)
--------------------------	---

Die Lösung wird autoklaviert, anschließend werden 5 ml MgCl₂ (2 M) und 20 ml Glukose (1 M) zugegeben.

2.1.4 Reaktionskits

Amplifikation von DNA

AllPrep DNA/RNA/Protein Mini Kit	Qiagen, Hilden, Deutschland
GenomePlex® Complete Whole Genome Amplification (WGA2) Kit	Sigma-Aldrich, Taufkirchen, Deutschland
Premix Ex Taq™ (Perfect Real Time)	TaKaRa Bio Inc., Shiga, Japan

Aufreinigung von DNA/RNA

Gel/PCR DNA Fragments Extraction Kit	Avegene, Taiwan
GenElute™ Mammalian Genomic DNA Miniprep Kit	Sigma-Aldrich, Taufkirchen, Deutschland
GenElute™ PCR Clean-Up Kit	Sigma-Aldrich, Taufkirchen, Deutschland
Microcon Centrifugal filters, Ultracell® YM-30	Millipore, Schwalbach, Deutschland
MinElute PCR Purification Kit	Qiagen, Hilden, Deutschland

Konzentrationsmessung von DNA

Quant-iT™ dsDNA BR Assay Kit	Invitrogen (Life Technologies), Darmstadt, Deutschland
Quant-iT™ dsDNA HS Assay Kit	Invitrogen (Life Technologies), Darmstadt, Deutschland
Quant-iT™ ssDNA Assay Kit	Invitrogen (Life Technologies), Darmstadt, Deutschland

<i>Quant-iT</i> TM RNA Assay Kit	Invitrogen (Life Technologies), Darmstadt, Deutschland
<u>Ligation</u>	
pGEM T System I Ligationskit	Promega, Madison, USA
<u>Nebulisierung</u>	
GS DNA Library Preparation Kit	Roche, Mannheim, Deutschland

2.1.5 Größenstandards für Agarosegele

Die bei der elektrophoretischen Auftrennung von DNA verwendeten Größenstandards sind im Folgenden mit den jeweils enthaltenen Fragmentgrößen in Basenpaaren angegeben:

GeneRulerTM 1kb-DNA-ladder (Fermentas, St. Leon Roth, Deutschland):

250 / 500 / 750 / 1000 / 1500 / 2000 2500 / 3000 / 3500 / 4000 / 5000 / 6000 / 8000 / 10000

100 bp DNA ladder (Solis BioDyne, Tartu, Estland):

100 / 200 / 300 / 400 / 500 / 600 / 700 / 800 / 900 / 1000 / 1500 / 2000 / 3000

2.1.6 Enzyme

HOT FIREPo [®] DNA Polymerase	Solis-BioDyne, Tartu, Estland
HotStarTaq DNA Polymerase	Qiagen, Hilden, Deutschland
HotStart-IT Bindingprotein	USB (Affymetrix), Cleveland, USA
HotStart-IT DNA Polymerase	USB (Affymetrix), Cleveland, USA
Platinum [®] Taq DNA Polymerase	Invitrogen (Life Technologies), Darmstadt, Deutschland
Proteinase K (PCR-Grade)	Roche, Mannheim, Deutschland
RNAse A	Invitrogen (Life Technologies), Darmstadt, Deutschland
T4-DNA-Ligase	Promega, Madison, USA
Taq-Polymerase	laboreigene Herstellung

2.1.7 Antikörper

Monoklonale anti-5mC-Antikörper

MAB-5MECYT-100	Diagenode, Liège, Belgien
15H61, gereinigt	Labor Schübeler, FMI, Basel, Schweiz

Sekundäre Antikörper

Sheep anti-Mouse IgG	Dynal Biotech (Invitrogen), Oslo, Norwegen
----------------------	--

2.1.8 Oligonukleotide

Alle im Folgenden aufgelisteten Oligonukleotide (Primer) wurden von den Firmen Metabion (Martinsried, Deutschland) oder Biomers (Ulm, Deutschland) bezogen. Zur besseren Übersicht sind die Primer nach Verwendungszweck geordnet (Tabelle 2.1-1, Tabelle 2.1-3, Tabelle 2.1-4). Für die GS FLX-Sequenzierung wurden Fusionsoligonukleotide in der PCR-Amplifikation eingesetzt. Hierfür wurde bei der Primersynthese eine Adaptorsequenz, bestehend aus Adaptor A/B sowie Erkennungssequenzen (Key, MID), an das 5'-Ende der spezifischen Oligonukleotid-Sequenzen angehängt (Tabelle 2.1-2).

Tabelle 2.1-1: Oligonukleotide zur Amplifikation Bisulfit-behandelter DNA. FP: Fusionsprimer, HS: *Homo sapiens*, Mm: *Mus musculus*

Amplikon	Primer	5' Primer (5' - 3')	3' Primer (5' - 3')	FP	Spezies
DP_DMSO	DP_Kav_DMSO_chr22_1398_F/R	GGTTAGGTTTGGTGGAAATAGTT	AACAAATCCAAATAACAACCATACA	-	Hs
DP_f48h	DP_Kav_48H_chr20_390_F/R	TAGAAAGTATTGTTAAAAGTTAATAA	TTCTCTCTTCTAACTACTATAAAT	-	Hs
DP_f120h	DP_Kav_120H_chr20_456_F/R	TAGTGTGGAGGAGAATAATTTTGA	TTCCAATCTCACCTAAAATATAAT	-	Hs
CEBPA_1	NP_Hs_CEBPA_F/R	GTTTTGTTAGGTTAAGGTTATTGT	AACCCTAAAACCCCTAAATAC	+	Hs
CEBPA_2	NP_Hs_CEBPA_F2/R2	GGAATTATAGGGGTAGTTGGAGA	CCCTCCTCTACCTACCCCTAAA	+	Hs
FOS_1	NP_Hs_FOS_F1/R1	TATAGGATGTTTATATTAGGATATTTG	CTCAATCTTAACTTCTCAAWTACTC	+	Hs
FOS_2	NP_Hs_FOS_F2/R2	GGTTTATGGTTATAGAGTTGGAGTT	CCCTAAATCTACAAAACCCCTCAA	+	Hs
HERVK	HERVK_F/R	TATTTTTTAATTTTAAGTATTTAGGGAT	ATACCTTCTCTTATCTCAACTACA	+	Hs
IGFBP1	NP_Hs_IGFBP1_F/R	GTTTTGGGTGATTAGTAAAATAA	CCCAAACCTAAACAATAATAAT	+	Hs
LINE1	LINE1_F/R	TGATTTTTGTATTTTATTTGAGGTAT	CCTCTAAACCAAATATAAAATATAATCTC	+	Hs
SPRY4	NP_Hs_SPRY4_F/R	GGTTTTATTATTTATTTGGTTAGTTT	TAAATATCCTTTCTCTATCCCAAT	+	Hs
Cacng5	NP_Mm_Cacng5_F/R	TTTTAGTTGATATTGAAAGAGGTAGG	AAATAAATACTATCCCCAAAATCAC	+	Mm
Dmc1h	NP_Mm_Dmc1h_F/R	GATTTGGTTTGGTGTGTTTATTT	AACCCTACCTACCTACTAACATCAC	+	Mm
Gabrg2	NP_Mm_Gabr2_F/R	AGGTTGTGGAAGTAGGGAGATTAT	AAATACCTTCAATATAACCTCAAATACC	+	Mm
Grin2b	NP_Mm_Grin2b_F/R	TTTAGTATTTAGATATTTAAGGATG	ATAATTATACTTTAAATCAATCTC	+	Mm
Pik3r1	NP_Mm_Pik3r1_F/R	TAGGGTTGGGAGAGAGTAGGTAAT	ATCCAAATCTTCCATAATCCAAAC	+	Mm
Ppp2r5c	NP_Mm_Ppp2r5c_F/R	TTGGGTAGGTTAGGGGTGTTATT	ACACATTCTCTACCTACAATTCATCCT	+	Mm
Snap25	NP_Mm_Snap25_F/R	TTAAGAGGTTATTAAGTTTTAAGATTG	AACAAAATATCTACTTCTACTAATTCAA	+	Mm
Usp29	NP_Mm_Usp29_F/R	TAGATGAGATAGGGAGTTTGGATTTT	CTCTACAACATACACCCTTTAAATACC	+	Mm
ZFP263	NP_Mm_Zfp263_F/R	TTTATGTAATTTGAGTTAGGTTGTTT	CAAATCACAATCTCAACAATATCA	+	Mm

Tabelle 2.1-2: Adaptorsequenzen zur GS FLX-Sequenzierung, bestehend aus Adaptor, Key und 10 bp-Erkennungssequenz (MID)

Adaptor-Typ	Adaptor	Key	MID	MID #
Adaptor A (F)	cgtatgcctccctcgcgcca	tcag	TCTCTATGCG	10
Adaptor B (R)	ctatgcgccttgccagcccgc	tcag	TCTCTATGCG	
Adaptor A (F)	cgtatgcctccctcgcgcca	tcag	TCACGTACTA	16
Adaptor B (R)	ctatgcgccttgccagcccgc	tcag	TCACGTACTA	
Adaptor A (F)	cgtatgcctccctcgcgcca	tcag	TACGAGTATG	22
Adaptor B (R)	ctatgcgccttgccagcccgc	tcag	TACGAGTATG	
Adaptor A (F)	cgtatgcctccctcgcgcca	tcag	TCGTCGCTCG	25
Adaptor B (R)	ctatgcgccttgccagcccgc	tcag	TCGTCGCTCG	

Tabelle 2.1-3: Oligonukleotide zur Amplifikation genomischer DNA (qPCR). Hs: *Homo sapiens*, Mm: *Mus musculus*

Amplikon	Primer	5' Primer (5' - 3')	3' Primer (5' - 3')	Spezies
C21orf93	MD_009_Hs_F/R	AGGAGGTATTGGAGCAGGACA	GGTGCTACCGAGGTGAATCTAA	Hs
CpG-frei	MD_014_Hs_F/R	TTGTGTGGTCAGGGATTATTCAGT	AGCAGCATCTTCCTGCAACCTTT	Hs
H2BFS	MD_008_Hs_F/R	GTACAAGGTGCTGAAGCAGGTC	CAGGCATCAATTTGGCAAGTT	Hs
KRTAP10-5	MD_005_Hs_F/R	AACCAGCATGAATGACGCATCTCC	AGCTCTCTGGGCAGTCGTCCACTC	Hs
MEST	MEST_F/R	AATTGCGTCTGAAGTCTGCGG	TCCACACAACAGAGAGGCG	Hs
MRPS6	MD_015_Hs_F/R	CCTGCCACCTTAGACAGTTGCT	CTGCTGCCTGTGCCAAAAGT	Hs
SH3BGR	MD_007_Hs_F/R	CATCTCTGCACAGGTCTCCGA	CCGCTAGACAAATCCTGACAGTGA	Hs
TRPM2	MD_003_Hs_F/R	CGACTCCTGCCCTGCCCTTTC	CGCAGCTAGAGCCAAGGACAC	Hs
CpG-frei	NP-MmCsa-1-F/R	CTTGTTGTTGGGTGGAGAGAATGGAA	AACCCCTGTGCTGCTCTGTGTCATC	Mm
H19	NP-MmH19_KK_F/R	AAGGAGACCATGCCCTATTCCTTG	CCTGAGGTAAGTGAAGTGGGTGA	Mm
Igf2r	NP-MmIgf2r_KK_F/R	ACAGCCTGGATAGTGCAGAGC	AGGTATCCTGAGGGTGCAAACCT	Mm
Xist	NP-MmXist-1-F/R	GCGGATCAGTTAAAGGCGTGC	GCCATAAGGCTTGGTGGTAGGG	Mm

Tabelle 2.1-4: Plasmidspezifische Oligonukleotide zur Amplifikation von in pGEM-T-Vektor subklonierten Fragmenten (Kolonie-PCR)

Primer	5' Primer (5' - 3')	3' Primer (5' - 3')
Kol-pcr-primer-F/R	GCTATTACGCCAGCTGGCGAAAGGGGATGTG	CCCCAGGCTTTACACTTTATGCTTCCGGCTCG

2.1.9 Microarrays

Hg18 Tiling Set36 Array

Zur chromosomenweiten Analyse angereicherter methylierter DNA (MeDIP) in humanen Zelltypen wurde der *Hg18 Tiling Set36 Array* (Assembly Hg18) von Roche NimbleGen (Reykjavík, Island) verwendet. Abdeckt sind hierbei kontinuierliche Bereiche der Chromosomen 20 (50.138.864-62.432.160), 21 (9.882.592-46.941.530) und 22 (14.430.066-39.529.778) wobei die 80-mere in 100 bp-Fenstern liegen. Repetitive Bereiche sind nicht im Array-Design enthalten.

Mouse 385K RefSeq Promotor Array

Zur genomweiten Analyse angereicherter methylierter DNA (MeDIP) in murinen Zelltypen wurde der *Mouse 385K RefSeq Promotor Array (Mm8)* von Roche NimbleGen (Reykjavík, Island) verwendet. Promotoren aller RefSeq Gene (Assembly Mm8) sind durchschnittlich in einem Bereich von -2000/+500 bp um die jeweilige Transkriptionsstartstelle abgedeckt, wobei die 50-75-mere in 100 bp-Fenstern liegen. Repetitive Bereiche sind nicht im Array-Design enthalten.

HumanHT-12 v3 Expression Bead Chip

Zur Analyse der mRNA-Expression in humanen Zelltypen wurde der *HumanHT-12 v3 Expression Bead Chip* von Illumina (San Diego, USA) verwendet. Das Design umfasst 35325 proteinkodierende RefSeq Transkripte (NM, XM), 642 nicht-kodierende Transkripte (NR, XR) sowie 12837 mRNAs/UniGene, wobei jedes Transkript mit 1-2 Oligos abgedeckt ist. Es können 12 Proben parallel bearbeitet werden.

2.1.10 Eukaryotische Zellen + DNA

2.1.10.1 Humane DNA

Folgende Nukleinsäuren aus humanem Vollblut/Leukozyten standen bereits zur Verfügung:

Leukozyten 480, 676 und 682:

Institut für Humangenetik und Molekulare Zytogenetik (A. Gerlach und H. Tönnies), Charité, Berlin (verwendet in Zhang et al. 2009)

Leukozyten MC:

Pool von sieben Donoren (männlich); Biological Resources Center (BRC) "Interactions Gène-Environnement en Physiopathologie CardioVasculaire" (IGE-PCV), Nancy, Frankreich; Teil einer Kontrollgruppe der STANISLAS-Kohorte (Siest et al. 1998, Visvikis et al. 2000, Visvikis-Siest und Siest 2008)

2.1.10.2 Humane Zelllinien

Zellpellets folgender humaner Zelllinien wurden zur Verfügung gestellt:

HEK-293T:

Adhärenz wachsende humane embryonale Nierenzelllinie („*Human Embryonic Kidney*“), welche durch das eingebrachte große T-Antigen aus dem Simian virus 40 (SV40) ein stabileres Zellwachstum sowie erhöhte Adhärenz erreicht. Verwendete Zellpellets

entstammen der laboreigenen Kultur (ursprünglich zur Verfügung gestellt von der Abteilung Virologie, AG Müller-Lantzsch, der Universitätsklinik Homburg.)

HepG2:

Epithelähnliche, adhärent wachsende hepatische Karzinomzelllinie; verwendete Zellpellets wurden vom Fraunhofer-Institut für Biomedizinische Technik (IBMT), St. Ingbert, zur Verfügung gestellt (ursprüngliche Bezugsquelle: DSMZ, Braunschweig) - die Zellen wurden z.T. vom IBMT kryokonserviert (Protokoll siehe Anhang B, Abschnitt 6.2.1).

2.1.10.3 Murine Zelllinien/Gewebe

Zellpellets folgender muriner Zelllinien wurden vom Labor Schöler, MPI für Molekulare Biomedizin, Münster zur Verfügung gestellt:

EpiSC („Epiblast Stem Cell“):

Zelllinie, welche aus Stammzellen des Epiblasten (Post-Implantations-Embryo) generiert wurde (GOF18 EpiSCs aus 5,5 Tage alten Embryonen (129/Sv Weibchen x C56/BL6 und DBA/2 Hintergrund GOF18+/+ Männchen)) (Greber et al. 2010).

ESC („Embryonic Stem Cell“):

Zelllinie, welche aus Stammzellen der inneren Zellmasse des Prä-Implantations-Embryos generiert wurde (Oct4-GFP transgenische Mäuse (B6;CBA-Tg[Pou5f1-EGFP]2Mnn/J) ~ OG2 Mäuse).

GSC („Germline Stem Cell“):

Zelllinie, welche aus unipotenten Zellen der männlichen Keimbahn generiert wurde (Oct4-GFP transgenische Mäuse (B6;CBA-Tg[Pou5f1-EGFP]2Mnn/J) ~ OG2 Mäuse) (Ko et al. 2009).

gPS („germline Pluripotent Cell“):

Zelllinie, welche durch spontane Reprogrammierung aus GSC generiert wurde (Oct4-GFP transgenische Mäuse (B6;CBA-Tg[Pou5f1-EGFP]2Mnn/J) ~ OG2 Mäuse) (Ko et al. 2009).

MEF („Mouse Embryonic Fibroblast“):

Zelllinie, welche aus embryonalen Fibroblasten der Maus generiert wurde (C57BL/6J).

Zu Etablierungszwecken wurde außerdem murines Leber-Gewebe (Inzuchtstamm der Hausmaus (*Mus musculus domesticus*); C57BL/6 (B6)) von der hauseigenen Tierhaltung zur Verfügung gestellt (die Präparation erfolgte durch *Dr. Konstantin Lepikhov*).

2.1.11 Bakterienstämme

Zur laboreigenen Herstellung chemokompetenter Zellen wurden Tetracyclin-Resistente *Escherichia coli* Top10F'-Zellen (Invitrogen, Karlsruhe, Deutschland / Life Technologies) mit folgendem Genotyp herangezogen:

$F' \{lacI^q, TN10(Tet^R)\}$ *mcrA* $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi80lacZ\Delta M15$ $\Delta lacX74$ *recA1* *araD139* $\Delta(ara-leu)7697$ *galU* *galK* *rpsL* (Str^R) *endA1* *nupG*

2.1.12 Instrumente und Maschinen

Fluorometer	Qubit™-Fluorometer, Invitrogen (Life Technologies), Darmstadt, Deutschland
Geldokumentation	Genesnap, Biostep GmbH, Jahnsdorf, Deutschland
Heizblock	Heap Labor Consult, Bovenden, Deutschland
Magnet-Ständer	Dynal® MPC-S, Invitrogen (Life Technologies), Darmstadt, Deutschland DynaMag™-2 Invitrogen (Life Technologies), Darmstadt, Deutschland
	PCR-Maschinen Eppendorf Master Gradient Eppendorf, Hamburg, Deutschland MxPro-Mx3000P, Stratagene, La Jolla CA, USA
Photometer	BioMate3, Thermo Scientific, Waltham, USA
Pipetten	Eppendorf, Hamburg, Deutschland Gilson, Middleton WI, USA
Sequenzierer	Genome Sequencer FLX, Roche, Mannheim, Deutschland
Sonifizierer	Bioruptor UCD200, Diagenode, Liège, Belgien
Überkopffrotator	Heidolph, Schwabach, Deutschland
Waagen	Sartoris, Bradfaord, USA
Vakuumzentrifuge	heto hetovac VRI BMI Surplus Inc., Hanover MA, USA CVC2 vacuubrand, Wertheim, Deutschland
Zentrifugen	Eppendorf MiniSpin, Eppendorf, Hamburg, Deutschland Eppendorf Centrifuge 5415C, Eppendorf, Hamburg, Deutschland Spectrafuge 24D Labnet International Inc., Woodbridge NJ, USA Napco 2019-R, Napco/ Precision Scientific, Winchester, USA

2.2 Methoden

A. Molekularbiologische Methoden

2.2.1 Isolation und Analyse von Nukleinsäuren

2.2.1.1 Isolation genomischer DNA

Genomische DNA von Zellkultur-Zellen wurde mit Hilfe der **Phenol-Chloroform-Isoamylalkohol-Extraktion (PCI)** isoliert. Dazu wurden die Zellen in je 500 µl Lösung A und Lösung B resuspendiert und nach Zugabe von 200 µg/ml Proteinase K bei 55 °C mindestens 3 h im Heizblock geschüttelt (bis alle Zellen aufgelöst waren). Anschließend wurden je ein halbes Ansatzvolumen Phenol und Chloroform/Isoamylalkohol (24:1) dazu gegeben und die Proben 15 min auf dem Überkopfrötator durchmischt. Nach 15 minütiger Zentrifugation bei RT, 13.000 rpm, wurde die obere Phase in ein neues Gefäß überführt. Anschließend wurde der gesamte Arbeitsschritt einmal mit den gleichen Mengen Phenol und Chloroform-Isoamylalkohol sowie einmal mit einem Ansatzvolumen Chloroform-Isoamylalkohol wiederholt. Die wässrige obere Phase (~1 ml) wurde dann mit 1/10 Volumen 3 M Natriumacetat (pH 5,5) und 2,5 Volumen absolutem Ethanol (eiskalt) versetzt. Der bei Invertierung des Gefäßes ausfallende DNA-Faden wurde vorsichtig mit einer Pipettenspitze „geangelt“, in 70%-igem Ethanol gewaschen und in einem Reaktionsgefäß getrocknet. Die DNA wurde dann in 1x TE gelöst. Anschließend wurden 20 µg RNase A pro 100 µl DNA-Lösung zugegeben und 30 min bei RT inkubiert.

Im späteren Verlauf der Arbeit wurde zur Isolation genomischer DNA ein auf bindenden Silika-Säulen basierendes **kommerzielles Kit** verwendet (*GenElute Mammalian Genomic DNA Miniprep Kit*, Sigma-Aldrich). Die Extraktion wurde nach Angaben des Herstellers durchgeführt.

Hochmolekulare genomische DNA ohne substantielle Anteile von RNA oder Degradation lieferten in der zur Überprüfung durchgeführten Agarosegelelektrophorese (1,2%-iges Gel, 1x TBE, 2.2.1.4) eine deutliche Bande ohne „Schmier“. Zusätzlich wurde die Qualität der DNA durch Bestimmung der UV-Absorption im Photometer eingeschätzt und die Konzentration mit Hilfe des *dsDNA BR Assay* im Fluorometer bestimmt (2.2.1.3).

2.2.1.2 Aufreinigung von DNA-Fragmenten

2.2.1.2.1 Aufreinigung von PCR-Produkten

DNA-Fragmente aus semiquantitativen PCRs (2.2.1.5.1) wurden unter Verwendung des *Gel/PCR DNA Fragments Extraction Kit* (Avegene) direkt oder, im Falle unspezifischer Produkte bzw. vorhandener Primer-Dimere, aus Agarosegelen aufgereinigt. Dabei wurden die Angaben des Herstellers befolgt. Die Elution erfolgte mit 15-20 µl Aqua bidest.

Die Produkte wurden anschließend weiteren Arbeitsschritten wie Ligation bzw. Sequenzierung (2.2.2.2.2 und 2.2.2.2.3) zugeführt.

2.2.1.2.2 Aufreinigung nebulisierter DNA

Nebulisierte DNA (2.2.2.1.1) wurde mit Hilfe des *MinElute PCR Purification Kit* (Qiagen) unter Abwandlung der Herstellerangaben gemäß des *GS FLX Shotgun DNA Library Preparation Method Manual* (Roche, Dezember 2007) aufgereinigt. Aufgrund des erhöhten Volumens war dabei eine mehrfache Beladung von jeweils 2-3 Säulen pro Nebulisierungsansatz nötig. Nach Elution in je 25 µl 1x TE wurden diese Fraktionen wieder vereinigt.

Die Proben wurden anschließend hinsichtlich ihrer Fragmentverteilung (1,5%-iges Agarosegel, 1x TBE) sowie DNA-Konzentration (*dsDNA BR Assay* oder *dsDNA HS Assay*, Fluorometer) überprüft (2.2.1.4, 2.2.1.3) und zur Immunopräzipitation eingesetzt (2.2.2.1.2). Material einer optionalen zweiten Elution wurde lediglich für quantitative PCRs verwendet (2.2.1.5.1).

2.2.1.2.3 Aufreinigung präzipitierter DNA

Angereicherte DNA-Fragmente der Immunopräzipitation (2.2.2.1.2) wurden, in Anlehnung an die Angaben von Weber und Mitarbeitern (Weber et al. 2005), mit Hilfe eines modifizierten Protokolls der Phenol-Chloroform-Isoamylalkohol-Extraktion aufgereinigt, welches der geringen Menge an DNA angepasst war. Die in 250 µl ProteinaseK-Puffer mit 70 µg Proteinase K ÜN bei 55 °C inkubierten Proben wurden mit einem Volumen extra-reinem Phenol-Chloroform-Isoamylalkohol-Gemisch (25:24:1, v/v, Invitrogen) versetzt und 15 min über Kopf rotierend durchmischt. Nach 15 minütiger Zentrifugation (RT, 13000 rpm) wurde die obere Phase in ein neues Gefäß überführt und der Arbeitsschritt einmal mit der gleichen Menge PCI-Gemisch und einmal mit einem Volumen Chloroform-Isoamylalkohol (24:1) wiederholt. Schließlich wurde die in der wässrigen Phase enthaltene DNA durch Zugabe von 500 µl absolutem Ethanol (eiskalt), 20 µl 5 M NaCl und 1 µl Glycogen (20 µg/µl) bei -20 °C über Nacht gefällt. Die DNA wurde 30 min bei 4 °C zentrifugiert (13000 rpm)

und das so erhaltene Pellet zweimal mit 800 µl 70%-igem Ethanol bei Zentrifugation (13000 rpm, 4 °C, 30 min) gewaschen. Das getrocknete Pellet wurde schließlich in 40 µl Nuklease-freiem Wasser (Sigma) eluiert.

Die aufgereinigte DNA wurde unter Verwendung des *ssDNA Assay* im Fluorometer quantifiziert (2.2.1.3), wobei typischerweise Präzipitationsraten von 2-5% der Ausgangsmenge gemessen wurden. Anschließend wurden die Proben weiteren Arbeitsschritten des MeDIP-Protokolls zugeführt (2.2.2.1).

2.2.1.2.4 Aufreinigung WGA-amplifizierter DNA

Input- und IP-Proben wurden nach genomrepräsentativer Amplifikation (2.2.1.5.2) unter Verwendung des *GenElute™ PCR Clean-Up Kit* (Sigma) gemäß den Angaben des Herstellers aufgereinigt. Die Elution erfolgte mit 50 µl Nuklease-freiem Wasser (Sigma).

Die aufgereinigte DNA wurde im Fluorometer quantifiziert (*dsDNA BR Assay*, typischerweise 3-5 µg) (2.2.1.3), die Fragmentverteilung mittels Agarosegel-elektrophorese überprüft (1,5%-iges Gel, 1x TBE; Hauptfragmente bei 400 nt) (2.2.1.4) und weiteren Arbeitsschritten des MeDIP-Protokolls zugeführt (2.2.2.1.2).

2.2.1.2.5 Aufreinigung von RNA

RNA wurde unter Verwendung des *AllPrep DNA/RNA/Protein Mini Kit* (Qiagen) gemäß Herstellerangaben aufgereinigt und anschließend im Fluorometer quantifiziert (*RNA Assay*) (2.2.1.3). Die weitere Prozessierung erfolgte durch einen Dienstleister (MFT, Tübingen).

2.2.1.3 Bestimmung der DNA-Konzentration und -Reinheit

Der **DNA-(oder RNA-)Gehalt** einer Probe wurde standardmäßig im Fluorometer ermittelt. Durch Verwendung DNA-(oder RNA-)spezifischer Fluoreszenzfarbstoffe werden Messungen nicht durch mögliche Verunreinigungen beeinflusst und erreichen somit eine höhere Genauigkeit als Photometer-basierte Systeme. Zudem können auch geringe Mengen fragmentierter DNA hoch-sensitiv quantifiziert werden (z.B. Immunopräzipitate oder PCR-Produkte). Die Messungen wurden im Qubit™ Fluorometer (Invitrogen) unter Verwendung der Reaktionskits i) *Quant-iT™ dsDNA BR Assay Kit* (2-1000 ng Doppelstrang-DNA), ii) *Quant-iT™ dsDNA HS Assay Kit* (0.2-100 ng Doppelstrang-DNA) oder iii) *Quant-iT™ ssDNA Assay Kit* (1-200 ng Doppel- und Einzelstrang-DNA) und iv) *Quant-iT™ RNA Assay Kit* nach Angaben des Herstellers durchgeführt. Dabei wurden in der Regel 2 µl DNA-Lösung eingesetzt.

Zusätzlich wurde im Falle genomischer DNA zur Qualitätskontrolle noch die Extinktion verschiedener Wellenlängen im Photometer bestimmt. Das Verhältnis der DNA/Protein-Absorption (Quotient OD 260 nm/ 280 nm) gilt dabei als **Reinheitsgrad** und sollte zwischen 1,7 und 2 liegen. Die Messungen wurden unter Verwendung von Quarzküvetten (Thermo Scientific) im Biomate 3 Spektrometer (Thermo Scientific) durchgeführt.

2.2.1.4 Gelelektrophoretische Auftrennung von DNA-Fragmenten

Zur Größen-Auftrennung von DNA-Fragmenten wurde die Gelelektrophorese verwendet. Dabei wandert die negativ geladene DNA nach Anlegen eines elektrischen Feldes in Richtung Anode. Die Porengröße der als Trägermatrix verwendeten Agarosegele wird über die Agarosekonzentration bestimmt.

In der vorliegenden Arbeit wurde die mit 1/10 Volumen Probenpuffer versetzte DNA in 1,2%-igen bzw. 1,5%-igen Agarosegelen (in 1x TBE) elektrophoretisch (80-200 V) aufgetrennt. Nach Anfärben der DNA in einem Bad mit Ethidiumbromid-Lösung wurde ein Photo unter UV-Licht angefertigt.

2.2.1.5 Amplifikation von DNA

2.2.1.5.1 Polymerase-Kettenreaktion

Bei der Polymerase-Kettenreaktion („*polymerase chain reaction*“) handelt es sich um ein *in vitro* Verfahren zur Vervielfältigung (Amplifikation) definierter DNA-Abschnitte mit Hilfe sequenzspezifischer Oligonukleotide („*Primer*“).

PCR zur Amplifikation von Bisulfit-behandelter DNA

Zur Amplifikation CpG-haltiger Regionen Bisulfit-behandelter DNA (2.2.2.2.1) wurde eine Heißstart-PCR durchgeführt. Mit Hilfe spezieller Enzym-Systeme werden durch Inaktivierung der DNA-Polymerase bei niedrigen Temperaturen unspezifische Amplifikationen reduziert. Die entsprechenden Parameter zur Verwendung der HOT FIREPol® DNA Polymerase (Solis BioDyne) bzw. der prozessivere HotStarTaq DNA Polymerase (Qiagen) sind in Tabelle 2.2-1 und Tabelle 2.2-2 zusammengefasst.

Tabelle 2.2-1: Parameter zur Amplifikation Bisulfit-behandelter DNA unter Verwendung der HOT FIREPol DNA Polymerase (HF). Die jeweilige Anlagerungstemperatur (AT) bzw. Zyklenzahl ist Tabelle 2.2-4 zu entnehmen

PCR-Reaktionsansatz HF	HF A (30 µl)	HF B (30 µl)	Thermalprofil		
Reagenz	Volumen [µl]	Volumen [µl]	Temperatur	Zeit	
Wasser	18,3	17,9	105 °C	Deckeltemp.	
10x Puffer B (MgCl ₂ -frei)	3	3			
MgCl ₂ (25 mM)	3	3	95 °C	15 min	
dNTPs (je 2,5 mM)	2,4	2,4	95 °C	1 min	x Zyklen
Bindingprotein (10 µg/µl)	0,2	0,2	AT	1 min	
Vorwärtsprimer (10 µM)	0,3	0,5	72 °C	1,5 min	
Rückwärtsprimer (10 µM)	0,3	0,5	72 °C	10 min	
HOT FIREPol (5 U/µl)	0,5	0,5	4 °C	∞	
Bisulfit-DNA	2	2			

Tabelle 2.2-2: Parameter zur Amplifikation Bisulfit-behandelter DNA unter Verwendung der HotStarTaq DNA Polymerase (HS). Die jeweilige Anlagerungstemperatur (AT) bzw. Zyklenzahl ist Tabelle 2.2-4 zu entnehmen

PCR-Reaktionsansatz HS	HS A (30 µl)	HS B (50 µl)	Thermalprofil		
Reagenz	Volumen [µl]	Volumen [µl]	Temperatur	Zeit	
Wasser	21,5	35,5	105 °C	Deckeltemp.	
10x PCR Puffer (15 mM MgCl ₂)	3	5			
dNTPs (je 2,5 mM)	2,4	4	95 °C	5 min	
Bindingprotein (10 µg/µl)	0,2	-	94 °C	1 min	x Zyklen
Vorwärtsprimer (10 µM)	0,3	1	AT	1 min	
Rückwärtsprimer (10 µM)	0,3	1	72 °C	1 min	
HotStarTaq (5 U/µl)	0,3	0,5	72 °C	10 min	
Bisulfit-DNA	2	3	4 °C	∞	

Tabelle 2.2-3: Parameter zur Amplifikation Bisulfit-behandelter DNA unter Verwendung der HotStart-IT DNA Polymerase (HSIT). Die Anlagerungstemperatur (AT) bzw. Zyklenzahl ist Tabelle 2.2-4 zu entnehmen

PCR-Reaktionsansatz HSIT	HSIT (30 µl)	Thermalprofil		
Reagenz	Volumen [µl]	Temperatur	Zeit	
Wasser	20,4	105 °C	Deckeltemp.	
10x PCR Puffer (15 mM MgCl ₂)	3	95 °C	5 min	
dNTPs (je 2,5 mM)	2,4	95 °C	1 min	x Zyklen
Vorwärtsprimer (10 µM)	0,6	AT	1 min	
Rückwärtsprimer (10 µM)	0,6	72 °C	1 min	
HotStart-IT (1,25 U/µl)	1	72 °C	5 min	
Bisulfit-DNA	2	4 °C	∞	

Die jeweilig verwendeten Protokolle, Anlagerungstemperaturen (AT) sowie Zyklenzahlen der verschiedenen Amplikons sind Tabelle 2.2-4 zu entnehmen. Zur Ermittlung der geeigneten Anlagerungstemperatur der Oligonucleotide wurde u.U. eine Gradienten-PCR

durchgeführt, bei welcher identische Reaktionsansätze bei ATs zwischen 50 °C und 60 °C amplifiziert wurden.

Die PCR-Produkte wurde mittels Agarosegelelektrophorese (1,2%-iges Gel, 1x TBE) überprüft (2.2.1.4), aufgereinigt (2.2.1.2.1) und weiteren Teilschritten der Methylierungsanalyse zugeführt.

Tabelle 2.2-4: PCR-Bedingungen der verschiedenen Amplikons, HF: HOT FirePol, HS: HotStarTaq, HSIT: HotStart-IT, AT=Anlagerungstemperatur

Amplikon	Parameter	AT (°C)	Zyklenzahl	Amplikon	Parameter	AT (°C)	Zyklenzahl
DP_DMSO	HS B	50	39	ZFP263	HF A	55	41
DP_f48h	HS B	56	39	Cacng5	HF A	55	41
DP_f120h	HS B	55	39	Ppp2r5c	HS A	55	38
LINE1	HF A	54	32 ^a	Grin2b	HS A	55	41
HERVK	HF B	54	40	Snap25	HF A	55	41
CEBPA_1	HIT	60	40	Usp29	HS A	55	41
CEBPA_2	HF B	54	40	Dmc1h	HS A	55	41
FOS_1	HF B	52	40	Gabrg2	HS A	55	38
FOS_2	HF B	54	40	Pik3r1	HS A	55	41
IGFBP1	HF B	52	40				
SPRY4	HF B	54	40				

^a 1 min Elongation bei 72 °C (statt 1,5 min)

Kolonie-PCR

Zur Amplifikation der in pGEM-T subklonierten Bisulfit-Produkte (2.2.2.2.2) wurde der PCR-Ansatz mittels eines sterilen Zahnstochers mit Einzelkoloniematerial inokuliert.

Die Produkte wurden anschließend mittels Agarosegelelektrophorese (1,2%-iges Gel, 1x TBE) überprüft und zur Sanger-Sequenzierung verschickt (2.2.2.2.3).

Tabelle 2.2-5: Parameter zur Amplifikation von Einzelkoloniematerial

PCR-Reaktionsansatz (30µl)		Thermalprofil		
Reagenz	Volumen [µl]	Temperatur	Zeit	
Aqua bidest.	22,9	105 °C	Deckeltemperatur	
10x Puffer III ^a	3	95 °C	5 min	
dNTPs (je 2,5 mM)	2,4	94 °C	30 sec	30 Zyklen
Vorwärtsprimer (10 µM)	0,6	65 °C	30 sec	
Rückwärtsprimer (10 µM)	0,6	72 °C	1,5 min	
Taq-Polymerase ^a	0,5	72 °C	5 min	
		4 °C	∞	

^a laboreigene Herstellung

Quantitative PCR (qPCR)

Im Gegensatz zur klassischen PCR ermöglicht die Echtzeit- oder auch quantitative PCR (qPCR) eine Quantifizierung des Produkts. Dazu können u.a. interkalierende Farbstoffe wie SYBR Green genutzt werden, deren Fluoreszenz bei Bindung an doppelsträngige DNA im Verlauf der PCR erfasst werden. Zur Quantifizierung werden die während der exponentiellen Phase gemessenen Signale herangezogen (optimale Reaktionsbedingungen), deren Beginn mit Hilfe des CT-Werts bestimmt wird. Hierbei handelt es sich um jenen PCR-Zyklus („cycle“ = C), bei welchem das Signal erstmals den Hintergrund (Schwellenwert = T („threshold“)) übertrifft.

In der vorliegenden Arbeit wurde die qPCR zur Überprüfung der Immunopräzipitate (IP) (2.2.2.1.2) herangezogen. Dabei wurde die spezifische Anreicherung verschiedener Promotorregionen mit bekannten Methylierungsleveln im Vergleich zum Hintergrund CpG-freier Negativkontrollen untersucht (Anhang A, Tabelle 6.1-1 und Tabelle 6.1-2).

Zur Etablierung der Amplikons wurden mit Hilfe von Verdünnungsreihen (0,125-32 ng) zunächst Standardkurven erstellt. Die Effizienz (E), welche über die Steigung der Standardkurve ermittelt wurde, sollte dabei reproduzierbar nahe 100% liegen. Zudem wurde die Spezifität der Produkte mittels Schmelzkurven überprüft. Auf diese Weise war eine korrekte Quantifizierung in allen Konzentrationsbereichen gewährleistet.

Die verwendeten Parameter des ursprünglichen sowie des im späteren Verlaufe optimierten Protokolls sind in Tabelle 2.2-6 und Tabelle 2.2-7 zusammengefasst. Für Input und IP wurden dabei pro Amplikon Duplikate erstellt.

Tabelle 2.2-6: ursprüngliche Parameter zur Amplifikation präzipitierten Materials, verwendet während der Etablierung der Immunopräzipitation (humane DNA)

PCR-Reaktionsansatz (20µl)		Thermalprofil		
Reagenz	Volumen [µl]	Temperatur	Zeit	
Nuklease-freies Wasser	8,4	95 °C	5 min	
10x PCR Puffer (MgCl ₂ -frei)	2	95 °C	30 sec	40 Zyklen
MgCl ₂ (50 mM)	0,6	60 °C	30 sec	
dNTPs (je 2,5 mM)	1,6	72 °C	30 sec	
SYBR Green (1:5000)	1	95 °C	1 min	
ROX (0,002x)	0,3	60 °C	30 sec	
Vorwärtsprimer (10 µM)	0,5	95 °C	30 sec	
Rückwärtsprimer (10 µM)	0,5			
PlatinumTaq DNA Polymerase	0,1			
Templat ^a	5			

^a 1/30 IP bzw. 25 ng Input

Tabelle 2.2-7: optimierte Parameter zur Amplifikation präzipitierten Materials unter Verwendung eines kommerziellen PCR-Gemisches, verwendet zur Überprüfung der in Abschnitten 3.2 und 3.3 und dargestellten Proben

PCR-Reaktionsansatz (20µl)		Thermalprofil		
Reagenz	Volumen [µl]	Temperatur	Zeit	
Nuklease-freies Wasser	3,3	95 °C	3 min	
Takara PremixExTaq™	10	95 °C	10 sec	40 Zyklen
SYBRgreen (1:5000)	1	60 °C	18 sec	
ROX (0,002x)	0,3	72 °C	18 sec	
Vorwärtsprimer (10 µM)	0,2	95 °C	10 sec	
Rückwärtsprimer (10 µM)	0,2	60 °C	18 sec	
Templat (2ng)	5	95 °C	30 sec	

Zur Bestimmung der spezifischen (absoluten) Anreicherung methylierter Kontrollregionen wurden die CT-Werte der IP unter Einbeziehung der Effizienz gegen Werte der Input-Kontrolle normiert (Formel (1)). Die relative Anreicherung methylierter Regionen im Verhältnis zur unspezifischen Präzipitation CpG-freier Referenz-Regionen wurde mit Formel (2) berechnet. Diese Normierung erlaubte auch einen direkten Vergleich von Experimenten mit abweichenden Parametern. Proben ausreichender Qualität wurden weiteren MeDIP-Teilschritten zugeführt (2.2.2.1).

$$(1) \text{ absolute Anreicherung} = 2^{\Delta CT_{\text{Target}} (\text{Input-IP})}$$

$$(2) \text{ relative Anreicherung} = \frac{2^{\Delta CT_{\text{Target}} (\text{Input-IP})}}{2^{\Delta CT_{\text{Referenz}} (\text{Input-IP})}}$$

2.2.1.5.2 Genomrepräsentative Amplifikation (WGA)

Die Microarray-Prozessierung von Proben des MeDIP-Experiments (2.2.2.1) erforderte von Seiten des Dienstleisters >3 µg Material. Da in der Regel lediglich 40-100 ng DNA präzipitiert wurden, war eine genomrepräsentative Amplifikation (Faktor 100-1000) nötig („whole genome amplification“, WGA).

Zu den bekannten WGA-Techniken, welche sich in Anpassung an die Erfordernisse molekularbiologischer Hochdurchsatz-Verfahren weiterentwickelten, zählen u.a. PCR-basierte Ansätze. Diese werden vergleichsweise wenig von der DNA-Qualität beeinflusst (Dean et al. 2002, Telenius et al. 1992, Zhang et al. 1992) und eignen sich somit zur Verarbeitung fragmentierter DNA. Dabei werden u.a. Protokolle verwendet, welche durch Ligation von Adaptoren an diese DNA-Fragmente gemeinsame Primer-Bindestellen für die anschließende Amplifikation liefern.

In der vorliegenden Arbeit wurde zur Amplifikation von IP und Input (2.2.2.1.2) das auf dieser Technik beruhende Reaktionskit *GenomePlex Complete Whole Genome Amplification (WGA2) Kit* (Sigma) verwendet, welches bereits erfolgreich in Verbindung mit immunopräzipitiertem Material eingesetzt wurde (O'Geen et al. 2006, Sorensen und Collas 2009). Die Prozessierung wurde, mit folgenden Ausnahmen, nach Angaben des Herstellers durchgeführt: i) Der erste Arbeitsschritt, eine chemische Fragmentierung der DNA, wurde ausgelassen, da bereits sonifizierte DNA vorlag; ii) Es wurden jeweils 20 ng IP bzw. Input statt der im Protokoll genannten 10 ng DNA eingesetzt, um die Gefahr eines Ungleichgewichts während der Amplifikation zu reduzieren. Das Material wurde durch Eindampfen (Speedvac) auf das erforderliche Volumen von 10 µl eingestellt und gemäß Herstellerprotokoll amplifiziert.

Die WGA-Produkte wurden anschließend aufgereinigt (2.2.1.2.4) und hinsichtlich Ausbeute sowie Qualität untersucht. Dazu wurde eine fluorometrische Quantifizierung (*dsDNA BR Assay*, Ausbeute typischerweise >3 µg, 2.2.1.3), eine gelelektrophoretische Bestimmung der Fragmentgrößen (1,5%-iges Agarosegel, Hauptfragmente um 400 nt, 2.2.1.4) sowie eine Überprüfung der spezifischen Anreicherung methylierter Kontrollregionen (qPCR, 2.2.1.5.1) vorgenommen. Die Proben wurden schließlich an den Dienstleister zwecks Microarray-Prozessierung verschickt (2.2.2.1.3).

2.2.2 Analyse der DNA-Methylierung

2.2.2.1 Immunopräzipitation methylierter DNA (MeDIP)

Bei MeDIP („*Methylated DNA Immunoprecipitation*“) handelt es sich um eine Antikörper(AK)-basierte Methode zur Analyse der DNA-Methylierung. Hierbei wird zunächst hoch-molekulare genomische DNA fragmentiert, denaturiert und mit einem anti-5mC-Antikörper inkubiert. Die Anreicherung gebundener methylierter DNA erfolgt mittels eines sekundären an magnetische Partikel gekoppelten Antikörpers. Unmethylierte Fragmente verbleiben im Überstand und werden während der anschließenden Waschschriffe entfernt. Das aufgereinigte Präzipitat (IP) wird im Vergleich zu lediglich fragmentierter Ausgangs-DNA (Input) evaluiert. Einzelne Loci können mittels Standard-Methoden wie qPCR analysiert werden. Durch Ko-Hybridisierung von IP und Input auf hochauflösende Tiling-Arrays können, wie in der vorliegenden Arbeit, genom- oder chromosomenweite Methylierungsprofile erfasst werden. Da die Immunopräzipitation von Anzahl und

Verteilung verfügbarer AK-Bindestellen abhängig ist, werden hierbei qualitative Informationen generiert.

Die einzelnen Arbeitsschritte des in Anlehnung an Weber und Mitarbeiter etablierten Protokolls (Weber et al. 2005, Weber et al. 2007) sind nachstehend zusammengefasst. Dabei werden auch entsprechende Qualitätskontrollen vermerkt, welche zur kritischen Beurteilung experimenteller Parameter während der Optimierung sowie zur Überprüfung jeder Probe nötig sind (siehe auch Pälme et al. 2011).

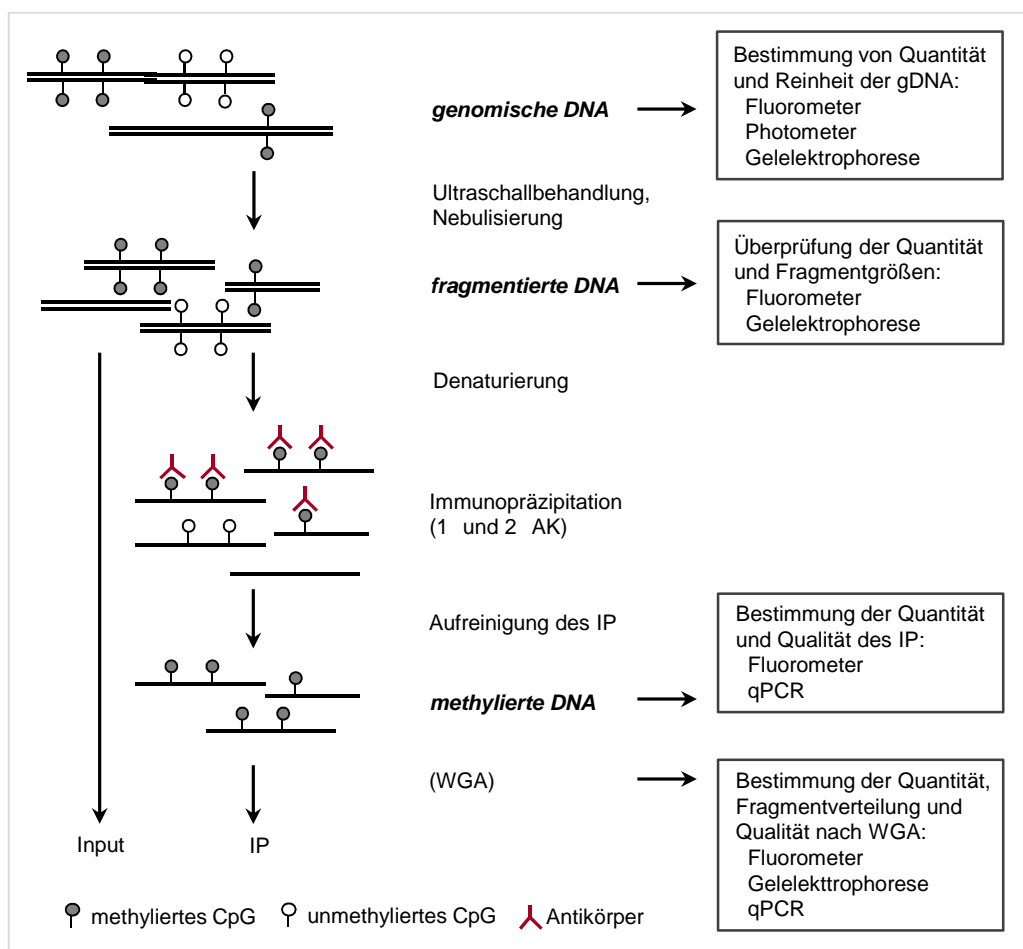


Abbildung 2.2-1: Prinzip der Immunopräzipitation methylierter DNA (MeDIP). Gezeigt sind die Teilschritte des Experimentes inklusive entsprechend durchgeführter Qualitätskontrollen

2.2.2.1.1 Fragmentierung genomischer DNA: Sonifizierung und Nebulisierung

Für eine effiziente Immunopräzipitation wird zufällig fragmentierte DNA benötigt, wobei die Fragmente, bei einer mittleren Größe von 500 bp, zwischen 250 und 1000 bp liegen

sollten (Weber 2005). Dazu wurden mechanische Verfahren wie Ultraschallbehandlung (Sonifizierung) oder Nebulisierung verwendet.

Sonifizierung

Bei der Sonifizierung erfolgt die Fragmentierung genomischer DNA mittels Ultraschallwellen in einem Wasserbad, wobei die Effizienz dieses Prozesses von mehreren Faktoren wie Temperatur, Volumen, DNA-Konzentration und DNA-Qualität beeinflusst wird.

1,5-10 µg genomische DNA wurden in 1x TE auf 75-100 µl verdünnt und im (vorgekühlten) Eiswasserbad des Bioruptors UCD-200 (Diagenode) in 15-Sekunden Zyklen fragmentiert. In Abhängigkeit der verfügbaren Ausgangsmenge und Konzentration der DNA sowie der angestrebten Anzahl an Präzipitationen (Replikate) war dabei eine Anpassung der Parameter (Tabelle 2.2-8) sowie unter Umständen eine Aufteilung auf mehrere Gefäße nötig.

Tabelle 2.2-8: Parameter der Sonifizierung. Protokoll B ist durch eine vergleichsweise niedrigere Geräteleistung in Anpassung an geringe Ausgangskonzentrationen gekennzeichnet.

		Protokoll A	Protokoll B
DNA-Menge	[µg]	9-10	1,5-6
Volumen	[µl]	100	75-80
Leistung des Geräts	[Watt]	200	160
Pulsfrequenz	[Sek]	15 An/ 15 Aus	15 An/ 15 Aus
Anzahl der Zyklen		6-9	8-14

Mit Hilfe der Agarosegelelektrophorese (2.2.1.4) (~100 ng DNA, 1,5%-iges Gel) wurden die Fragmentgrößen überprüft und u.U. die Anzahl weiterer Scherzyklen abgeschätzt. Abschließend wurde die DNA-Konzentration unter Verwendung des *dsDNA BR Assay* im Fluorometer verifiziert (2.2.1.3).

Nebulisierung

Die Nebulisierung stellt ein alternatives mechanisches Verfahren zur Fragmentierung von DNA dar, welches insbesondere im Bereich der Hochdurchsatz-Sequenzierung Anwendung findet. Bei diesem Vorgang wird durch das unter definiertem Druck zugeführte Stickstoffgas eine Flüssigkeit zerstäubt und die in ihr gelöste DNA geschert.

Die Nebulisierung wurde gemäß des *GS FLX Shotgun DNA Library Preparation Method Manual* (Roche, Dezember 2007) unter Verwendung des entsprechenden Materials des *GS DNA Library Preparation Kit* (Roche) durchgeführt. Bis zu 7 µg genomische DNA wurden mit 1x TE auf ein Volumen von 100 µl gebracht und mit 500 µl Nebulisierungspuffer (Roche)

versetzt. Die Proben wurden dann 1 min mit 3 bar Stickstoffgas in den Nebulisierungskammern fragmentiert. Etwa 85% des ursprünglichen Volumens konnten bei Entnahme der im Gefäß zerstäubten Probe wiedergewonnen werden.

Nach Aufreinigung der Proben (2.2.1.2.2) wurden abschließend die Fragmentgrößen mittels Agarosegelelektrophorese überprüft (~100ng, 1,5%-iges Gel, 2.2.1.4) und die DNA-Konzentration verifiziert (*dsDNA BR Assay*, 2.2.1.3).

2.2.2.1.2 Immunopräzipitation

Für die weiteren Arbeitsschritte wurden mindestens 5 µg fragmentierte DNA pro Ansatz benötigt, wovon 1 µg als Input-Kontrolle für spätere Prozesse zurückbehalten wurde. Für die Immunopräzipitation selbst wurden je 4 µg fragmentierte DNA in 1x TE auf ein Volumen von 450 µl verdünnt und 10 min im kochenden Wasserbad denaturiert. Auf diese Weise sollte die spätere Bindung des AK begünstigt werden, welcher eine hohe Affinität zu einzelsträngiger DNA aufweist. Die 10 min in Eiswasser abgekühlten Proben wurden mit 50 ml 10x IP-Puffer sowie 8 µg anti-5mC-AK versetzt (AK:DNA-Verhältnis 2:1) und bei 4 °C über Kopf rotierend inkubiert (ÜN).

Die Anreicherung AK-gebundener DNA-Fragmente wurde mit Hilfe eines sekundären, an magnetische Partikel gekoppelten Antikörpers (Dynabeads) erreicht. Zur Vermeidung unspezifischer Bindungen wurde der 2° AK (je 20 µl) zunächst durch zweimaliges Waschen in Blockier-Puffer vorbehandelt (je 5 min über Kopf rotierend in 800 µl Puffer) und schließlich in 20 µl 1x IP-Puffer eluiert. Die Proben wurden anschließend 2 h bei RT mit Dynabeads auf einem Überkopffrotator inkubiert, um eine Sedimentation zu verhindern. Die AK-DNA-Komplexe wurden dann mit Hilfe des Magnetständers von ungebundener DNA separiert und schrittweise gereinigt. Dazu wurden die Dynabeads drei mal 5 min in je 700 µl 1x IP-Puffer unter vorsichtiger über Kopf Rotation gewaschen und schließlich in 250 µl Proteinase-K-Puffer resuspendiert. Um die Effizienz des nachfolgenden AK-Abbaus mittels Proteinase-K zu erhöhen, wurde vor deren Zugabe (70 µg, 3,5 µl einer 20 mg/ml Stocklösung) eine zusätzliche 10-minütige Denaturierung im kochenden Wasserbad vorgenommen. Die DNA-Fragmente wurden wie in Abschnitt 2.2.1.2.3 beschrieben aufgereinigt und in 40 µl Nuklease-freiem Wasser eluiert.

Anschließend wurde die Präzipitationsrate, welche typischerweise bei 2-5% des Ausgangsmaterials lag, durch Konzentrationsmessung im Fluorometer (*ssDNA Assay*) bestimmt (2.2.1.3). Des Weiteren wurde die Qualität des Materials durch Ermittlung der spezifischen Anreicherung methylierter Regionen im Vergleich zum Hintergrund überprüft (qPCR, 2.2.1.5.1).

2.2.2.1.3 Post-Prozessierung

Bei anschließender Micorarray-Analyse wurde das Material, nach optionaler genomrepräsentativer Amplifikation (2.2.1.5.2), zur Versendung an den Dienstleister (Imagenes, Berlin) vorbereitet. Dazu wurden je ~3 µg amplifizierte bzw. ~300 ng unamplifizierte IP und Input auf 20 µl eingedampft.

Im Folgenden sind die zur Bearbeitung der in Abschnitten 3.1-3.2 dargestellten Proben verwendeten Parameter zusammengefasst (Tabelle 2.2-9). (Experimente zu Optimierungszwecken sind hierbei nicht berücksichtigt. Die verschiedenen Proben (Leukozyten (676, 682), HEK293T-Zellen, HepG2-Zellen, murine Leber) wurden dabei generell mit Sonifizierungs-Protokoll A, und Präzipitation von 4 µg DNA mit 1°AK der Firma Diagenode durchgeführt.)

Tabelle 2.2-9: Versuchsbedingungen der in der vorliegenden Arbeit dargestellten Proben

	Zelltyp	Proben-Bezeichnung	Fragmentierung	1°AK	Material in IP	WGA	Array-Design
1	Leukozyten (480, 676)	Leuko_1	Sonifizierung Protokoll A	Diagenode	4 µg	+	Hg18 Tiling Set36
	HEK293T	+WGA, -WGA			7 x 4 µg	+/-	Hg18 Tiling Set36
2	Leukozyten (MC)	Leuko_2	Sonifizierung Protokoll B	Diagenode	2 µg	+	Hg18 Tiling Set36 ^a
3	HepG2	NKtr, Ktr, DMSO, f48h, f120h	Sonifizierung Protokoll B	Diagenode	4 µg	+	Hg18 Tiling Set36
4	murine SC + MEF	GSC, gPS, ESC, EpiSC, MEF	Nebulisierung	Schübeler	2 x 4 µg	-	RefSeq Promotor (Mm8)

^a Immunopräzipitation und WGA erfolgten durch die Diplomandin *Sabrina Pichler*. In der vorliegenden Arbeit wurde die Datenauswertung des erneut hybridisierten Materials vorgenommen.

2.2.2.2 Sequenzierung Bisulfit-behandelter DNA

Zur Validierung und Ergänzung der Immunopräzipitations-basierten Daten wurden definierte CpG-haltige Regionen durch Sequenzierung Bisulfit-behandelter DNA untersucht. Dieser als Gold-Standard angesehene Ansatz ermöglicht die Erfassung des Methylierungsstatus einzelner CpG-Positionen im Sequenzkontext und stellt somit eine hoch-sensitive und quantitative Methode dar.

Die Methode basiert auf einer chemischen Konvertierung unmethylierter Cytosine (in einzelsträngiger DNA) zu Uracil, während methylierte Cytosine aufgrund der am C5-Atom gebundenen Methylgruppe vor dieser Reaktion geschützt sind (Abbildung 2.2-2 (A)).

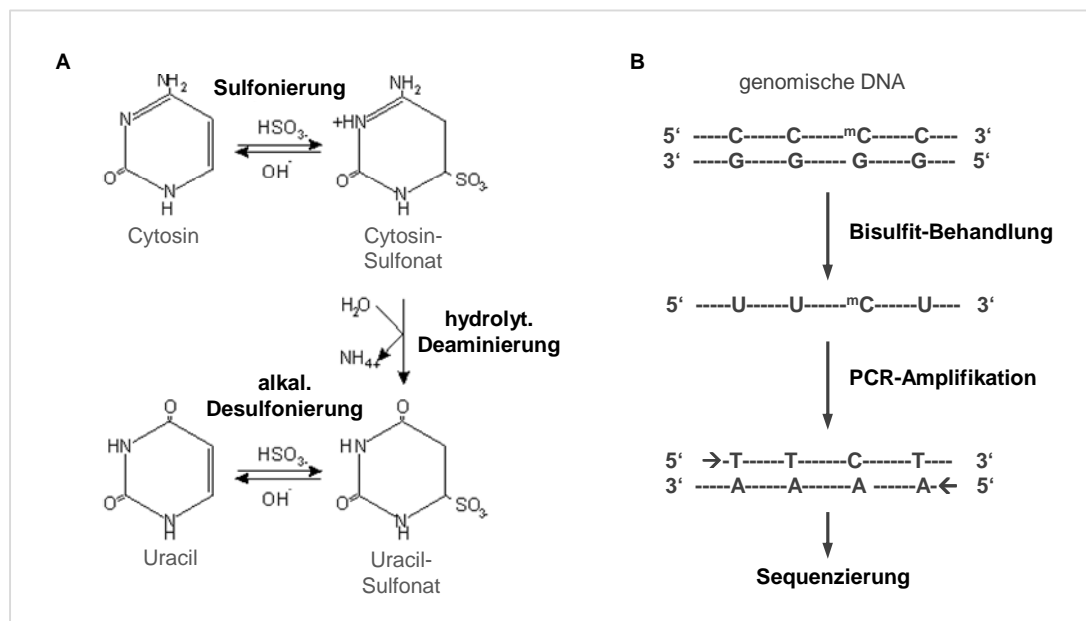


Abbildung 2.2-2: Sequenzierung Bisulfit-behandelter DNA. (A) Prinzip der chemischen Konvertierung, mit Sulfonierung von Cytosin zu Cytosinsulfonat, Deaminierung zu Uracilsulfonat und anschließender Desulfonierung zu Uracil. (B) Identifizierung des Methylierungsstatus einzelner CpG-Positionen mittels Sequenzierung PCR-amplifizierter Bisulfit-DNA. In Uracil konvertierte unmethylierte Cytosine werden in Thymin umgewandelt. C^m: methyliertes Cytosin, C: unmethyliertes Cytosin, U: Uracil, T: Thymin, A: Adenin, G: Guanin. (verändert nach http://www.methods.info/Methods/DNA_methylation/methylation_bisulphite.gif (A) und <http://media.wiley.com/CurrentProtocols/MB/mb2117/mb2117-fig-0004-1-full.gif> (B))

In einer darauf folgenden PCR-Amplifikation der Bisulfit-behandelten DNA werden Uracile in Thymin umgewandelt (Clark et al. 1994, Clark et al. 2006, Frommer et al. 1992) (Abbildung 2.2 2 (B)). Nach abschließender Sequenzierung der Amplifikate kann im Vergleich mit der Ausgangssequenz der Methylierungsstatus der einzelnen CpG-Positionen bestimmt werden.

2.2.2.2.1 Bisulfit-Behandlung genomischer DNA

Es wurden je 300-400 ng genomische DNA in einem Volumen von 20 µl mit 187 µl Natriumbisulfit-Lösung sowie 73 µl Dioxan-Radikalfänger-Lösung versetzt. Die Reaktion erfolgte im Thermocycler nach folgendem Programm (Tabelle 2.2-10):

Tabelle 2.2-10: Programm zur Bisulfit-Behandlung genomischer DNA

Thermalprofil	
Temperatur	Zeit
99 °C	15 min
50 °C	30 min
90 °C	5 min
50 °C	90 min
99 °C	5 min
50 °C	90 min

Der Ansatz wurde anschließend mit 150 µl Aqua bidest. gemischt und zur Aufreinigung auf eine DNA-bindende Säule (Millipore) gegeben. Nach 20 minütiger Zentrifugation bei 1100 rpm wurde das Material 15 min bei 1100 rpm mit 500 µl 1x TE gewaschen. Es folgte eine 10 minütige Inkubation mit 500 µl 0,3 M NaOH und nachfolgender Zentrifugation (15 min, 1100 rpm) sowie ein weiterer Waschschrift mit 500 µl 1x TE (15 min, 1100 rpm). Die DNA wurde schließlich von der umgekehrt in ein frisches Reaktionsgefäß überführten Säule mit 50 µl 1x TE (50 °C) durch 20 minütige Zentrifugation bei 3000 rpm eluiert.

Die Bisulfit-behandelte DNA wurde im Folgenden zur Amplifikation definierter Abschnitte mittels PCR (2.2.1.5.1) und späterer Sequenzierung (2.2.2.2.3) eingesetzt.

2.2.2.2.2 Subklonierung von Bisulfit-PCR-Produkten

Einige der Bisulfit-PCR-Produkte wurden zwecks späterer Sequenzierung repräsentativer Einzelfragmente (2.2.2.2.3) subkloniert.

Ligation von Bisulfit-PCR-Produkten

Aufgereinigte PCR-Produkte Bisulfit-behandelter DNA wurden unter Verwendung des *pGEM-T vector system I ligation kit* (Promega) in den Vektor pGEM-T ligiert. Der Reaktionsansatz (Tabelle 2.2-11) wurde 2 h bei Raumtemperatur oder über Nacht bei 4 °C inkubiert.

Tabelle 2.2-11: Reaktionsansatz zur Ligation von Bisulfit-PCR-Produkten in pGEM-T Vektor

Reaktionsansatz (10 µl)	
Reagenz	Volumen [µl]
Aqua bidest.	0,5
2x Ligationspuffer	5
pGEM-T Vektor	0,5
T4-Ligase (3U/µl)	1
aufgereinigtes PCR-Produkt	3

Hitzeschock-Transformation kompetenter E.coli Zellen

Es wurden 50 µl chemokompetenter Zellen (*E.coli* TOP10 F', laboreigene Herstellung) mit 5 µl Ligationsansatz gemischt und für 30 min auf Eis inkubiert. Nach einem 60 Sekunden langen Hitzeschock bei 42 °C (Heizblock) wurden 200 µl eiskaltes SOC-Medium auf die Zellen gegeben. Der Ansatz wurde 1 h bei 37 °C inkubiert und auf LB-Agar-Platten mit Ampicillin (100 µg/ml), X-Gal (2%) und IPTG (0,1 M) ausplattiert. Es folgte eine Inkubation bei 37 °C über Nacht.

Da die Integration des PCR-Produkts in den Vektor durch Unterbrechung der kodierenden Sequenz des lac-Z-Gens dessen Expression und damit die enzymatische Umsetzung des in den Platten enthaltenen X-Gal verhindert, konnten positive Klone anhand der ausbleibenden Blaufärbung erkannt werden. Je 16-24 Kolonien wurden mittels PCR näher überprüft (2.2.1.5.1). Produkte mit korrektem Insert wurden zur Sequenzierung eingesetzt (2.2.2.2.3).

2.2.2.2.3 DNA-Sequenzierung

Zur Bisulfit-basierten Analyse der DNA-Methylierung wurde die Sanger-Sequenzierung einzelner Klone sowie eine auf Pyrosequenzierung beruhende Hochdurchsatz-Sequenzierung mit Hilfe des GS FLX Systems (454/Roche) eingesetzt. Tabelle 2.2-12 gibt für die verschiedenen Regionen die verwendete Methode an.

Tabelle 2.2-12: Übersicht der pro Amplikon verwendeten Sequenzier-Technik

Amplikon	Sequenzier-Methode	Amplikon	Sequenzier-Methode
DP_DMSO	Sanger	ZFP263	Pyrosequenzierung, GS FLX
DP_f48h	Sanger	Cacng5	Pyrosequenzierung, GS FLX
DP_f120h	Sanger	Ppp2r5c	Pyrosequenzierung, GS FLX
LINE1	Pyrosequenzierung, GS FLX	Grin2b	Pyrosequenzierung, GS FLX
HERVK	Pyrosequenzierung, GS FLX	Snap25	Pyrosequenzierung, GS FLX
IGFBP1	Pyrosequenzierung, GS FLX	Usp29	Pyrosequenzierung, GS FLX
SPRY4	Pyrosequenzierung, GS FLX	Dmc1h	Pyrosequenzierung, GS FLX
CEBPA (1+2)	Pyrosequenzierung, GS FLX	Gabrg2	Pyrosequenzierung, GS FLX
FOS (1+2)	Pyrosequenzierung, GS FLX	Pik3r1	Pyrosequenzierung, GS FLX

Sanger-Sequenzierung

Nach Subklonierung der Bisulfit-PCR-Produkte (2.2.2.2.2) wurden die Amplifikate von <16 Kolonien zur Aufreinigung und Sanger-Sequenzierung an das Max Planck Institut für molekulare Genetik, Berlin, verschickt.

Pyrosequenzierung mit dem GS FLX System

Die Mehrzahl der Amplikons wurde mit dem GS FLX System, 454/Roche, analysiert. PCR-Produkte wurden dabei unter Verwendung von Fusions-Oligonukleotiden erstellt. Die darin enthaltenen Erkennungssequenzen (MID) sowie die physikalische Aufteilung der PicoTiter-Platte in 8 Regionen ermöglichte eine spätere Unterscheidung der verschiedenen Proben. Die aufgereinigten Produkte (2.2.1.2.1) wurden im Fluorometer quantifiziert (*dsDNA BR* bzw. *HS Assay*, 2.2.1.3) und in 1x TE (Sigma) auf 1×10^9 Moleküle/ μl verdünnt. Die weitere Prozessierung der Proben erfolgte durch die Technische Assistentin *Jasmin Gries* gemäß den Angaben des Herstellers (*GSFLX-Titanium emPCR Lib-A SV Protokoll*).

B. Bioinformatische Methoden und Programme

2.2.3 Analyse der MeDIP-Chip-Daten

Die vom Dienstleister (ImaGenes, Berlin) erhaltenen Rohdaten der Microarray-Hybridisierung von IP und Input wurden durch eine Abfolge verschiedener Analyseschritte ausgewertet (Abbildung 2.2-3, siehe auch Pälmeke 2010). Aufgrund der großen Datenmengen (>350000 Datenpunkte pro Farbkanal pro Array) und der daraus resultierenden komplexen Analyse hinsichtlich Statistik und Visualisierung, wurde für einige Teilschritte, insbesondere bei Qualitätskontrolle und Prä-Prozessierung der Daten, die freie Programmiersprache R (<http://www.r-project.org>, (R.D.C.Team 2010, Ihaka und Gentleman 1996)) herangezogen. Die darauf aufbauende Software Bioconductor (<http://www.bioconductor.org>, (Gentleman et al. 2004)) bietet zahlreiche Pakete, die eine Erweiterung des Funktionsumfangs zur Auswertung von Micorarrays erlauben. Zur Identifikation (differentiell) methylierter Regionen wurden in Abhängigkeit des Array-Designs zwei unterschiedliche Ansätze herangezogen, welche auf einer Analyse definierter Promotor-Regionen bzw. chromosomenweiter Abschnitte basieren. Die Analyse der finalen Regionen hinsichtlich der Anreicherung biologischer Prozesse (GO-Annotation) sowie Sequenzeigenschaften wurde hauptsächlich mit Web-basierten Programmen durchgeführt.

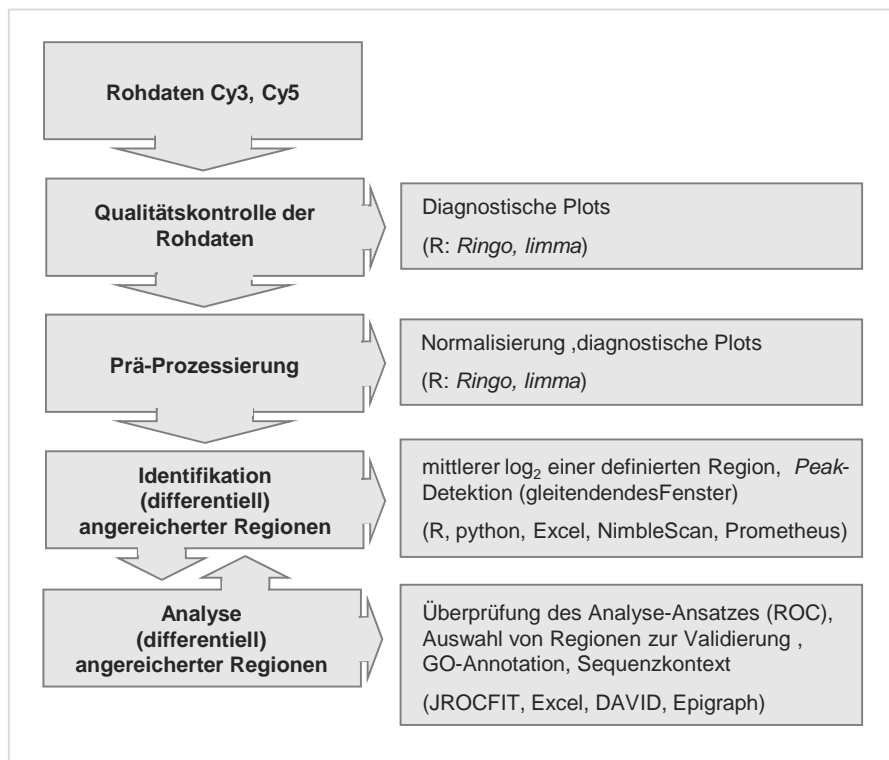


Abbildung 2.2-3: Übersicht der aufeinanderfolgenden Teilschritte zur Analyse der MeDIP-Chip-Daten mit Angabe der jeweils hauptsächlich verwendeten Programme.

2.2.3.1 Qualitätskontrolle und Prä-Prozessierung

Vor der eigentlichen Analyse (differenziell) methylierter Regionen wurden die MeDIP-Chip-Daten zunächst einer Qualitätskontrolle sowie Prä-Prozessierung unterzogen. Signalverzerrungen und Variabilitäten zwischen verschiedenen Arrays, welche z.B. auf eine unterschiedliche Effizienz der Farbstoffinkorporation oder abweichende Hybridisierungs- und Scan-Bedingungen zurückzuführen sind, sollten so identifiziert und wenn möglich durch Normalisierung korrigiert werden (2.2.3.1.2, 2.2.3.1.3) (Huber et al. 2005, Bolstad et al. 2003, Quackenbush 2002). Diagnostische Plots (2.2.3.1.1) wurden dabei als Werkzeug zur Überprüfung der Datenqualität bzw. Kontrolle der Normalisierung herangezogen.

Qualitätskontrolle und Prä-Prozessierung wurden entweder anhand der Signale der einzelnen Kanäle (Cy3, Cy5) oder, hauptsächlich, auf Ebene der \log_2 -Ratios von IP (Cy5)/Input (Cy3) vorgenommen. Dazu wurden vorwiegend R-Standardfunktionen sowie einzelne Funktionen der R-Pakete *Ringo* (Toedling et al. 2007) und *limma* (Smyth 2005, Smyth und Speed 2003) kombiniert.

2.2.3.1.1 Diagnostische Plots

Zur Qualitätskontrolle der Rohdaten (2.2.3.1.2) bzw. zur Überprüfung der Prä-Prozessierung (2.2.3.1.3) wurden hauptsächlich die im Folgenden aufgeführten diagnostischen Plots herangezogen.

Pseudobild

Zur Identifikation systematischer Hybridisierungs-Artefakte wie Kratzer oder Belichtungsfehler wurde anhand der Rohdaten eine Rekonstruktion des Array-Bilds vorgenommen, wobei die Signale der einzelnen Oligonukleotide gemäß ihrer Koordinaten dargestellt wurden. Dazu wurde die entsprechende *image*-Funktion des *Ringo*-Pakets (Toedling et al. 2007) auf das zuvor erstellte *RG*-Objekt (2.2.3.1.2) angewendet.

Signalverteilung

Mit Hilfe verschiedener Plots wurde die Verteilung der Signalintensitäten einzelner Kanäle bzw. \log_2 -Ratios visualisiert. *Boxplots* und *Dichteplots* wurden unter Verwendung von R-Standardfunktionen zum einfachen und schnellen Vergleich mehrerer Proben oder Kanäle herangezogen, während *MA-plots* Auskunft über das Verhältnis der Signalintensitäten der Kanäle zueinander gaben. Mittels der *limma*-Funktion *plotMA3by2* () wurde das Intensitäts-Verhältnis *M* gegenüber der durchschnittlichen Intensität *A* evaluiert (Formel (3) und (4)), wobei *R* und *G* die Signal-Intensitäten der roten und grünen Kanäle repräsentieren. Signalverzerrungen konnten anhand von Abweichungen der für Immunopräzipitations-basierte Daten typischen gewölbten Kurvenform identifiziert werden.

$$(3) \quad M = \log_2 (R) - \log_2 (G)$$

$$(4) \quad A = 1/2 (\log_2 (R) + \log_2 (G))$$

Korrelationsmatrix

Zur Überprüfung der Reproduzierbarkeit der Experimente wurden unter Berücksichtigung der \log_2 -Ratios aller Oligonukleotide paarweise Korrelationen zwischen Replikaten bzw. Proben bestimmt. Die zwischen den Arrays errechneten Pearson-Korrelations-Koeffizienten wurden mit entsprechenden Streudiagrammen zum direkten Vergleich in einer Korrelationsmatrix zusammengefasst. Korrelations-Koeffizienten sollten dabei über 0,8 liegen. Es wurde eine unabhängige R-Funktion, mit leichten Abwandlungen, verwendet (<http://addictedtor.free.fr/graphiques/graphcode.php?graph=137>).

Hierarchische Clusteranalyse

Zur systematischen Gruppierung von Replikaten/Proben in Abhängigkeit ihrer Ähnlichkeit zueinander wurde, auf Basis der Pearson-Korrelations-Koeffizienten, eine distanzbasierte Methode herangezogen. Die hierarchische Clusteranalyse wurde mittels R-Standardfunktionen durchgeführt und als Dendrogramm visualisiert.

2.2.3.1.2 Qualitätskontrolle

Die Signalintensitäten (Rohdaten) wurden vom Dienstleister pro Array und Farbstoff (Cy3, Cy5) in tabellierten Dateien (.pair) erhalten. Unter Verwendung der Funktion *readNimblegen ()* des *Ringo*-Paketes wurden diese Daten in R als **Objekt der Klasse *RGList*** eingelesen. Die dazu benötigten Dateien zur Spezifikation der Proben bzw. Reporter-Kategorien des Arrays wurden zuvor gemäß der Vorlage des Pakets erstellt.

Mittels der in 2.2.3.1.1 beschriebenen **diagnostischen Plots** Pseudobild, Boxplot, Dichtepplot, MA-Plot und Korrelationsmatrix wurden die Rohdaten hinsichtlich technischer Variabilitäten, systematischer Signalverzerrungen und sonstiger Artefakte untersucht. In Kombination gaben diese Plots Auskunft über Qualität und Reproduzierbarkeit der Experimente. Datensätze mit leichten Abweichungen wurden der Prä-Prozessierung zugeführt (2.2.3.1.3), während solche mit extremen Artefakten von weiteren Analysen ausgeschlossen wurden.

2.2.3.1.3 Prä-Prozessierung

Eine unterschiedliche Effizienz der Farbstoffinkorporation oder abweichende Hybridisierungs- und Scan-Bedingungen können zu Signalverzerrungen oder Variabilitäten zwischen verschiedenen Arrays führen. Um eine Detektion tatsächlich biologisch relevanter Veränderungen zwischen Proben zu erlauben, wurden Variabilitäten zwischen Arrays bzw. leichte Signalverzerrungen mittels **Normalisierung** korrigiert. Dabei wurden die Arrays zunächst basierend auf dem jeweiligen Median der \log_2 -ratios auf den gleichen Level (Null) zentriert. Für diese Intra-Array-Normalisierung (Median-Zentrierung) wurde die Funktion *preprocess ()* (Methode „nimblegen“) des *Ringo*-Paketes auf das *RG*-Objekt angewendet (unter Ausschluss der Random-Oligonukleotide). Eine weitere Inter-Array-Angleichung erfolgte durch zusätzliche Anpassung aller Quantile (Quantil-Normalisierung) mit Hilfe der *limma*-Funktion *normalizeBetweenArrays ()* (Methode „quantile“). Diese wurde auf Basis der \log_2 -Ratios aller Oligonukleotide (Matrix) durchgeführt. Weitere Normalisierungs-Ansätze wurden unter Verwendung der *limma*-Funktionen *normalizeWithinArrays ()* (Methode „loess“) bzw. *normalizeBetweenArrays ()* (Methode „quantile“) auf Basis der einzelnen Kanäle (*RG*-Objekt) getestet.

Die normalisierten Daten wurden mit Hilfe der in Abschnitt 2.2.3.1.1 beschriebenen **diagnostischen Plots** Boxplot, Dichteplot, MA-Plot, Korrelations-Plot und Cluster-Analyse überprüft. Des Weiteren wurden **Mittelwerte** technischer bzw. biologischer Replikate pro Oligo- \log_2 berechnet und die Daten schließlich in tabellarischer Form gespeichert.

Im Falle der murinen MeDIP-Chip-Daten (3.3) erfolgte die Prä-Prozessierung mit Hilfe der Software *Prometheus* (de Las Heras et al. in Vorbereitung). Über eine graphische Anwenderoberfläche wurden die Daten in das auf R basierende Programm automatisiert eingelesen und unter Spezifikation der entsprechenden Parameter normalisiert. Durch Zugriff auf das *limma*-Paket wurde eine Median-Zentrierung (Intra-Array-Normalisierung) sowie Skalierungs-Normalisierung (Inter-Array-Angleichung) vorgenommen. Die korrigierten Daten wurden zur Überprüfung mittels der in 2.2.3.1.1 beschriebenen diagnostischen Plots exportiert.

Erfolgreich prä-prozessierte Daten wurden zur Evaluierung der spezifischen Anreicherung methylierter Fragmente eingesetzt (2.2.3.2).

2.2.3.2 Identifizierung (differentiell) angereicherter Regionen

Prä-prozessierte Daten wurden zur Evaluation (differentiell) angereicherter Regionen eingesetzt. In Abhängigkeit des Array-Designs wurden dazu zwei verschiedene Ansätze genutzt, welche auf einer Betrachtung definierter Bereiche (Promotoren) (2.2.3.2.1) bzw. einer Peak-Detektion über große genomische Abschnitte eines Chromosoms (2.2.3.2.2) beruhen. Um Überinterpretationen individueller Ausreißer zu vermeiden, wurden jeweils mehrere benachbarte Oligonukleotide berücksichtigt.

2.2.3.2.1 Promotor-Analyse

Der *MM8 385K RefSeq Promotor Array* (Roche-NimbleGen) deckt durchschnittlich 2,5 kb große genomische Abschnitte um die TSS von 19489 Transkripten ab. Dabei sind jeweils etwa 25 Oligonukleotide (50-70-mere) in 100 bp-Fenstern platziert. Die Analyse der (differentiellen) Anreicherung solcher Promotor-Regionen erfolgte in Anlehnung an publizierte Ansätze (Mohn et al. 2008). In der vorliegenden Arbeit wurden dazu hauptsächlich Funktionen des Programms *Prometheus* sowie des R-Paketes *limma* genutzt. Zunächst wurde die **mittlere Anreicherung** standardisierter Promotor-Regionen auf Basis der normalisierten Daten (2.2.3.1.3) ermittelt. Um einen Informationsverlust der dynamischen Methylierungssignaturen solcher regulativen Einheiten zu minimieren, wurden mittels implementierter Funktionen der Prometheus-Software pro TSS zwei 1000 bp-große Subregionen definiert: i) -1500 bp bis -500 bp von TSS gelegen (Aufwärts-Promotor); ii) -

500 bp bis +500 bp um TSS (Kern-Promotor). Die \log_2 -Ratios der diese Regionen abdeckenden Oligonukleotide wurden durch Berechnung des Medians zusammengefasst. Die mittleren Anreicherungsgrade der Aufwärts- und Kern-Promotoren wurden im Folgenden getrennt betrachtet. Die Daten wurden zur weiteren Analyse (u.a. diagnostische Plots (2.2.3.1.1)) bzw. Visualisierung (z.B. Signalmap, siehe 2.2.7) in tabellarischer Form exportiert.

Zur Ermittlung der **differentiellen Anreicherung** zwischen verschiedenen analysierten Proben wurden eine Abfolge verschiedener Funktionen des R-Paketes *limma* genutzt. Dabei wurden mittels der empirischen Bayesschen Methode unter Berücksichtigung der Replikat-Varianz p-Werte berechnet. Auf Basis der Median- \log_2 -Werte (tabellarische Datenmatrix) wurde zunächst ein lineares Modell angepasst und eine gemäßigte T-Statistik berechnet (Funktionen *lmFit()*, *contrast.fit()*, *eBayes()*). Korrigierte p-Werte (FDR, „false discovery rate“) wurden durch multiples Testen mittels der Benjamini-Hochberg-Methode erhalten (Funktion *topTable()*). Promotor-Regionen mit statistisch signifikanter differentieller Anreicherung wurden anhand der folgenden Kriterien ermittelt: i) ein korrigierter p-Wert (FDR) < 0.05; ii) eine Median- \log_2 -Ratio >0 in mindestens einer der beiden zu vergleichenden Proben; iii) ein delta- \log_2 -Wert >0.58 (>1.5-facher Anreicherungsunterschied) und iv) eine Abdeckung von >67% des Promotors mit Oligonukleotiden. Die Resultate wurden anschließend weiteren Analysen (Abschnitte 2.2.3.3, 2.2.5) zugeführt.

2.2.3.2.2 Chromosomenweite Analyse: Detektion von Peaks

Der *Hg18 Tiling Set36-Array* (Roche-NimbleGen) deckt kontinuierliche Abschnitte der Chromosomen 20-22 ab. Diese genomischen Bereiche, in welchen 80 bp-Oligonukleotide in 100 bp-Fenstern platziert sind, werden durch nicht im Design enthaltene Regionen repetitiven Charakters unterbrochen. Zur Erstellung von Anreicherungsprofilen im Kontext unterschiedlicher funktioneller Regionen wurde das im Zusammenhang mit ChIP-Chip-Analysen gebräuchliche Konzept eines gleitenden Fensters herangezogen (cite). In abgewandelter Form und unter Einbeziehung zusätzlicher statistischer Funktionen des R-Paketes *limma* wurde dieser Ansatz ebenfalls zur Evaluierung differentieller Anreicherung verwendet.

Basierend auf den normalisierten \log_2 -Ratios (GFF-Format) (2.2.3.1.3) wurden **angereicherte Regionen** mit Hilfe einer für Methylierungsanalysen entwickelten Methode zur Detektion von Peaks identifiziert (*NimbleScan* v2.5, Roche-NimbleGen). Unter Verwendung eines modifizierten Kolmogorov-Smirnov-Tests (Scacheri et al. 2006) wurde

dabei in einem gleitenden Fenster die Wahrscheinlichkeit einer Anreicherung jedes Oligonukleotids unter Berücksichtigung der umgebenden Fragmente berechnet. Benachbarte Oligonukleotide über einem spezifischen Schwellenwert (>1.3 , entspricht $-\log_{10}$ der Wahrscheinlichkeit „p“ (<0.05)) wurden zu Anreicherungspeaks (AP) zusammengefasst und deren Koordinaten im GFF-Format ausgegeben.

Regionen **differentieller Anreicherung** wurden basierend auf $\Delta\log_2$ -Werten der zu vergleichenden Proben (GFF-Format) unter Verwendung des obigen Ansatzes bestimmt. Zur Verfeinerung der Analyse wurden unter Berücksichtigung der Replikat-Varianz korrigierte p-Werte (FDR) anhand der in 2.2.3.2.1 erläuterten Methode mit Hilfe des R-Paketes *limma* ermittelt. Die dazu benötigte mittlere Anreicherung der jeweiligen Proben in der spezifizierten Region wurde mittels maßgefertigter Funktionen zur Anwendung in Python (*Diana Santacruz*) berechnet. Dazu wurden zunächst anhand der genomischen Koordinaten die einen Peak bildenden Oligonukleotide bestimmt und der Mittelwert ihrer \log_2 -Ratios errechnet. Regionen statistisch signifikanter differentieller Anreicherung (DP: differentieller Anreicherungspeak) wurden schließlich anhand folgender Kriterien bestimmt: i) korrigierter p-Wert (FDR) <0.05 (oder nicht-korrigierter p-Wert <0.01); ii) eine mittlere \log_2 -Ratio >0 in mindestens einer der zu vergleichenden Proben; iii) ein $\Delta\log_2$ -Wert >0.5 (>1.4 -facher Anreicherungsunterschied) und iv) eine Abdeckung von $>67\%$ des Peaks mit Oligonukleotiden.

Die Resultate wurden anschließend weiteren Analysen (2.2.3.3) sowie der Visualisierung (z.B. Signalmap, siehe 2.2.7) zugeführt.

2.2.3.3 Analyse (differentiell) angereicherter Regionen

2.2.3.3.1 ROC-Analyse

Zur Evaluation des Ansatzes hinsichtlich der Identifizierung (differentiell) methylierter Regionen bzw. zum qualitativen Vergleich verschiedener Experimente wurde eine ROC-Analyse vorgenommen (ROC, „receiver operator characteristic“). Dieses Klassifizierungsmodell dient der Bestimmung von Sensitivität und Spezifität einer diagnostischen Variablen, deren graphische Darstellung über die Fläche unter der Kurve (AUC, „area under curve“) Auskunft über die Performance gibt. Dabei sind steile Kurven bei einer AUC nahe 1 ideal.

Die diagnostische Variable wurde in diesem Fall durch die An- oder Abwesenheit eines APs oder DPs (2.2.3.2.2) in einer Promotor-Region in Abhängigkeit des Methylierungsstatus repräsentiert. Dazu wurden zunächst die genomischen Positionen der Peaks unter Verwendung maßgefertigter Funktionen zur Anwendung in Python (*Diana Santacruz*) mit

>260 sequenzierten Amplikons des Chromosoms 21 (Zhang et al. 2009) abgeglichen (>1 bp Überlapp). Die Promotor-Regionen wurden entsprechend ihrer gewebespezifischen Methylierungslevel bzw. Methylierungsunterschiede anhand verschiedener Schwellenwerte (10%, 20%, 25%, 30%, 40%, 50%) kategorisiert in: i) unmethyliert und methyliert oder ii) nicht differentiell methyliert und differentiell methyliert. Die jeweiligen Positiv- und Negativ-Raten wurden unter Verwendung des Peak-Score (mittlere Anreicherungswahrscheinlichkeit der Oligonukleotide) mit Hilfe des Web-basierten Programms *JROCFIT* (<http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html>) berechnet und mittels Excel visualisiert (Verwendung angepasster Werte („fitted“)).

2.2.3.3.2 Systematische Analyse der Regionen / Gene

Verschiedene Arbeitsschritte dienten einer systematischen Analyse der APs, DPs und Promotor-Regionen im Vergleich der Proben zueinander hinsichtlich zellspezifischer Effekte und Sequenzcharakteristika.

Zur Ermittlung interessanter Effekte wurden APs und DPs anhand der genomischen Koordinaten mit Genen (+/- 1 kb um Peak) bzw. untereinander zwischen verschiedenen Proben mittels maßgeschneiderten Funktionen zur Verwendung in Python abgeglichen. Zum zellspezifischen Vergleich differentiell methylierter Promotoren konnte aufgrund der standardisierten Regionen die spezifische Kennung genutzt werden (Excel). Entsprechende Überschneidungen bzw. Gruppierungen wurden mit verschiedenen Programmen visualisiert (2.2.7). Gene von Interesse wurden des Weiteren einer funktionellen Analyse mittels Gen-Ontologie (2.2.5) zugeführt.

Eine Untersuchung von DPs in Bezug auf verschiedene Sequenzeigenschaften wurde mit Hilfe des Web-basierten Programms *EPIGRAPH* vorgenommen (Bock et al. 2009). Die Positionen der Regionen von Interesse wurden dabei gegen ein automatisch generiertes Kontrollset mit einer Reihe verfügbarer Attribute (Gene, CpG-Inseln, CpG-Dichte etc.) abgeglichen.

2.2.4 Analyse von mRNA-Expressions-Daten

Daten der *HumanHT-12 v3 Expression Bead Chips* (Illumina) wurden nach Qualitätskontrolle, Prä-Prozessierung und statistischer Analyse (ANOVA und Anpassung des p-Wertes für multiples Testen) vom Dienstleister (MFT, Tübingen) bereitgestellt. Signifikant differentiell exprimierte Gene (FDR <0.05, >2-facher Expressionsunterschied) wurden anschließend im Vergleich der Proben von Interesse systematisch gruppiert (Excel),

visualisiert (2.2.7) sowie einer funktionellen Analyse mittels Gen-Ontologie zugeführt (2.2.5).

2.2.5 Funktionelle Analyse mittels Gen-Ontologie

Zur funktionellen Analyse mittels Gen-Ontologie wurden Gene von Interesse (differentiell methylierte Promotoren, mit DPs überlappende Gene, differentiell exprimiert Gene) hinsichtlich einer Überrepräsentation biologischer Prozesse mit Hilfe des Web-basierten Programs *DAVID* (<http://david.abcc.ncifcrf.gov/home.jsp>, (Huang da et al. 2009)) untersucht. Die Gesamtheit der auf dem entsprechenden Microarray analysierten Gene wurde dazu jeweils als Referenz genutzt. Um die Redundanz der GO-Annotationen zu reduzieren und dadurch die Analyse zu fokussieren, wurde die Option „*Functional Annotation Clustering*“ gewählt, welche Kategorien ähnlicher biologischer Bedeutung in Gruppen zusammenfasst. Das Programm berechnet P-Werte („*EASE-Score*“, modifizierter *Fischer Exact Test*) der einzelnen Kategorien sowie Anreicherungsgrade der entsprechenden Gruppen (negativer Logarithmus des geometrisches Mittels der jeweiligen *EASE-Scores*). Letztere wurden in Excel visualisiert.

2.2.6 Analyse von Daten der Bisulfit-Sequenzierung

Daten der Sequenzierung Bisulfit-behandelter DNA wurden mit dem Program *BiQ-Analyzer* (Bock et al. 2005) bzw. im Falle der GS FLX-Sequenzierung mit der weiterentwickelten Version *BiQ-AnalyzerHT* (Lutsik et al. 2011) analysiert. Dabei wurden die Sequenzen automatisiert eingelesen, mit der jeweiligen Referenz abgeglichen und gefiltert (nach Sequenzidentität). Anschließend wurde der Methylierungsstatus der einzelnen CpG-Positionen bzw. die prozentuale Gesamtmethylierung ermittelt, welche sich aus dem Verhältnis methylierter Positionen im Verhältnis zur Anzahl der auswertbaren Positionen errechnet. Zusätzlich wurden von *BiQ-AnalyzerHT* pro Amplikon und Probe Methylierungskarten (Heatmaps) erstellt.

2.2.7 Liste verwendeter Datenbanken, Programme, Softwares und Web-Seiten

BDPC; Web-basiertes Programm zur Erstellung von Methylierungskarten (Heatmaps) ausgehend von .html-Dateien der *BiQ-Analyzer*-Software

<http://biochem.jacobs-university.de/BDPC/index.php> (Rohde et al. 2008)

BioVenn; Web-basiertes Programm zur Erstellung proportionaler Venn-Diagramme

<http://www.cmbi.ru.nl/cdd/biovenn/> (Hulsen et al. 2008)

Bioconductor; Version 2.2 und 2.6; auf R aufbauende Software mit Paketen zur Erweiterung des Funktionsumfangs

<http://www.bioconductor.org> (Gentleman et al. 2004))

BiQ-Analyzer / BiQ-AnalyzerHT; Software zur Auswertung von Sequenzen bisulfitierter DNA im Vergleich zur Referenzsequenz mit Berechnung der mittleren Methylierung und graphischer Darstellung

<http://biq-analyzer.bioinf.mpi-inf.mpg.de/> (Bock et al. 2005)

<http://biq-analyzer-ht.bioinf.mpi-inf.mpg.de/> (Lutsik et al. 2011)

Chromas; Software zur Auswertung von Sequenz-Chromatogrammen

<http://www.technelysium.com.au/chromas.html>

ClustalW2 (EMBL-EBI); Web-basiertes Programm für Sequenzabgleiche

<http://www.ebi.ac.uk/Tools/clustalw/index.html>

DAVID; Web-basiertes Programm für funktionelle Analysen mittel Gen-Ontologie

<http://david.abcc.ncifcrf.gov/home.jsp> (Huang da et al. 2009)

DNA Sequence Reverse And Complement; Web-basiertes Programm zur (Revers-) Komplementierung von DNA Sequenzen

http://www.cellbiol.com/scripts/complement/reverse_complement_sequence.html

Ensembl Genome Browser; Gen-Datenbank

(<http://www.ensembl.org/index.html>)

EpiGRAPH; Web-basiertes Programm zur Ermittlung von Sequenzcharakteristika für eine Gruppe von Regionen

<http://epigraph.mpi-inf.mpg.de/WebGRAPH/> (Bock et al. 2009)

Genesnap V.4, Biostep, Jahnsdorf, Deutschland; Software für UV-Fotografien von Agarosegelen

The Gene Ontology, Gen-Ontologie-Datenbank

<http://www.geneontology.org/>

gplots; Software-Paket zur Verwendung in R; Graphische Funktionen, z.B. zur Erstellung von Heatmaps

JROCFIT; Web-basiertes Programm zur ROC-Analyse

<http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html>

limma; Version 2.14.7 und 3.4.4; Software-Paket zur Verwendung in R, Prozessierung und statistische Analyse von Microarray-Daten (Smyth 2005)

MEDME; Version 1.8.0; Software-Paket zur Verwendung in R; Prozessierung von MeDIP-Chip-Daten, z.B. Glättung der Oligo- \log_2 -Ratios zur graphischen Darstellung (Pelizzola et al. 2008)

MGI (Mouse Genome Informatics), Gen- und Ontologie-Datenbank Maus

<http://www.informatics.jax.org/>

Microsoft® Office 2002, 2007, 2010

NCBI, National Center for Bioinformatical Information; Gen-und Literatur-Datenbank
<http://www.ncbi.nlm.nih.gov>

Nimblescan, NimbleGen; Software zur Analyse von Microarray-Daten, u.a. Peak-Detektions-Algorithmus

Oligo 6, Molecular Biology Insights, Inc; Software zum Design von PCR-Oligonukleotiden (Primer)

Prometheus; Software zur Analyse von 385K RefSeq Promotor-Arrays (de Las Heras et al. in Vorbereitung)

Python; Version 2.4.4; kommandobasierte Programmiersprache (van Rossum 1995)

R; Version 2.7 und 2.11.1; Kommando-basierte Programmiersprache mit statistischen und graphischen Funktionen (R.D.C.Team 2010, Ihaka und Gentleman 1996)

READSEQ Sequence Conversion; Web-basiertes Programm zur Umwandlung von Sequenzformaten
<http://www-bimas.cit.nih.gov/molbio/readseq/>

Ringo; Version 1.12.0 zbd 1.4.0; Software-Paket zur Verwendung in R; Prozessierung von Microarray-Daten (Toedling et al. 2007)

Signalmap, NimbleGen; Software zur Visualisierung von Microarray-Daten (Oligo- \log_2 -Ratios, Peaks) relativ zu genomischen Koordinaten und funktionellen Regionen (GFF-Format)

UCSC Genome Browser; Gen-Datenbank; BLAST („*Basic local alignment search tool*“) für Sequenzabgleiche, Datenbanksuche
<http://genome.ucsc.edu/>

3 Ergebnisse

3.1 Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung

Seit Einführung des MeDIP-Chips zur Analyse genomweiter DNA-Methylierungsmuster im Jahre 2005, hat diese Methode neben ähnlichen Ansätzen eine weite Verbreitung gefunden (Weber et al. 2005, Weber et al. 2007, Mohn et al. 2008, Keshet et al. 2006, Zilberman et al. 2007, Dindot et al. 2009, Farthing et al. 2008, Cortese et al. 2011, Dudzic et al. 2012). Dabei wurden der experimentelle Ablauf und die Menge des Ausgangsmaterials vielfach optimiert und den Gegebenheiten angepasst, so dass Präzipitationen derweil auch erfolgreich mit weniger als 200 ng DNA durchgeführt werden (Borgel et al. 2010, Guibert et al. 2012). Ausgehend von Studien, welche die Bedeutung gewebe- und krankheitsspezifischer Methylierung außerhalb von Promotoren verdeutlicht haben (Irizarry et al. 2009, Dindot et al. 2009, Ball et al. 2009, Suzuki und Bird 2008, Meissner et al. 2008), hat sich zusätzlich das Interesse genomweiter Methylierungsstudien von Promotor-basierten Arrays über hochauflösende „*high-tiling*“ Arrays größerer chromosomaler Abschnitte bis hin zur Sequenzierung des kompletten Methyloms (MeDIP-seq) verschoben. Daraus gingen ebenso vielfältige Ansätze zur Datenanalyse hervor, inklusive Qualitätsmanagement, Identifikation differentiell methylierter Regionen und der Umwandlung einer relativen Anreicherung in quantifizierbare Werte (Down et al. 2008, Pelizzola et al. 2008, Chavez et al. 2010, Huang et al. 2012).

Zu Beginn der vorliegenden Arbeit waren jedoch kaum Erfahrungswerte hinsichtlich experimenteller Parameter und Daten-Analyse der damals noch jungen Methode MeDIP-Chip zugänglich. Dies betraf insbesondere die Erfassung von Methylierungsmustern im Kontext eines kompletten Chromosoms, da MeDIP-Chip bis dato ausschließlich zur Analyse von Promotoren verwendet wurde. Der folgende Abschnitt geht auf wichtige Punkte und kritische Aspekte ein, die im Zuge der laborinternen Etablierung und Optimierung von MeDIP-Chip deutlich wurden.

3.1.1 Etablierung experimenteller Parameter der Immunpräzipitationsbasierten Methode

Die Methode MeDIP, kurz für „methylierte DNA Immunpräzipitation“, basiert auf einer Anreicherung methylierter DNA Sequenzen mittels spezifisch gegen 5-Methylcytosin (5mC) gerichteter Antikörper. Abbildung 3.1-1 veranschaulicht schematisch die einzelnen Schritte wie Fragmentierung und Denaturierung der DNA sowie Immunpräzipitation und Aufreinigung des Materials (siehe auch 2.2.2.1). Zusätzlich kann gegebenenfalls eine genomrepräsentative Amplifikation (WGA, „*whole genome amplification*“) des Präzipitats (IP, ca. 2-5% des Ausgangsmaterials) erfolgen. Anschließende Analysen mittels PCR, Microarray-Hybridisierung oder Sequenzierung geben Auskunft über die Anreicherung einzelner Gene oder chromosomaler Abschnitte im Vergleich zur Input Kontrolle (relative Methylierung). Eine Standardisierung der experimentellen Abläufe mit möglichst effizienter Präzipitation ist erforderlich, um reproduzierbare Daten zu erhalten und damit vergleichende Interpretationen zu erlauben.

Unter diesem Gesichtspunkt wurde MeDIP in der vorliegenden Arbeit, ausgehend vom Originalprotokoll (Weber et al. 2005) und unter Einbindung verschiedener Qualitätskontrollen, etabliert (Abbildung 3.1-1). Im Folgenden werden einige wesentliche experimentelle Teilschritte im Detail dargestellt. Dabei wird auch auf die unterschiedliche Effizienz humanen und murinen Ausgangsmaterials eingegangen.

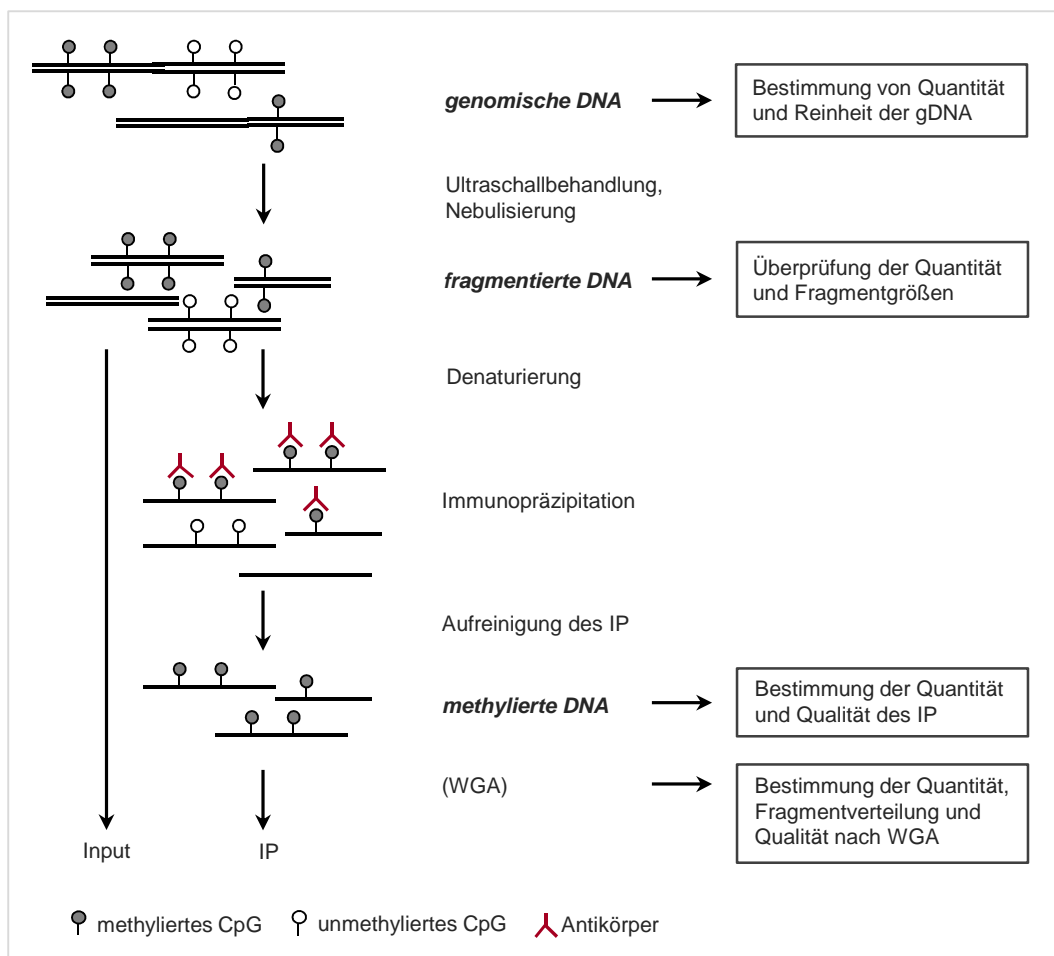


Abbildung 3.1-1: Prinzip der Immunopräzipitation methylierter DNA (MeDIP). Gezeigt sind die Teilschritte des Experimentes inklusive entsprechend durchgeführter Qualitätskontrollen (Boxen). AK: Antikörper, IP: Immunopräzipitat, WGA: genomrepräsentative Amplifikation („whole genome amplification“).

3.1.1.1 Anreicherung methylierter DNA als qualitative Information

Das Präzipitationsverhalten methylierter DNA wurde mit Hilfe der quantitativen *real time* PCR (qPCR) beleuchtet. Neben Promotoren unterschiedlicher Methylierungslevel und Sequenzeigenschaften wurden auch CpG-freie Bereiche des Genoms untersucht (Anhang A, Tabelle 6.1-1 und Tabelle 6.1-2). Die Daten wurden dabei als absolute oder relative Anreicherung betrachtet, wobei die spezifische Anreicherung (gegenüber Input) der Positiv- und Negativkontrollen getrennt oder im Verhältnis zueinander berechnet wurde (siehe auch 2.2.1.5.1., S. 38).

Letztere ist in Abbildung 3.1-2 für eine Reihe humaner Promotor-Regionen dargestellt, welche das nicht-lineare Verhältnis der Anreicherung zu Methylierungsgrad und CpG-Gehalt veranschaulicht. Die Abhängigkeit der Immunopräzipitations-basierten Methode nicht nur

von der Methylierung selbst, sondern auch von der Anzahl und Verteilung der verfügbaren AK-Bindestellen bedingt die Generierung qualitativer, nicht quantitativer, Methylierungsinformationen.

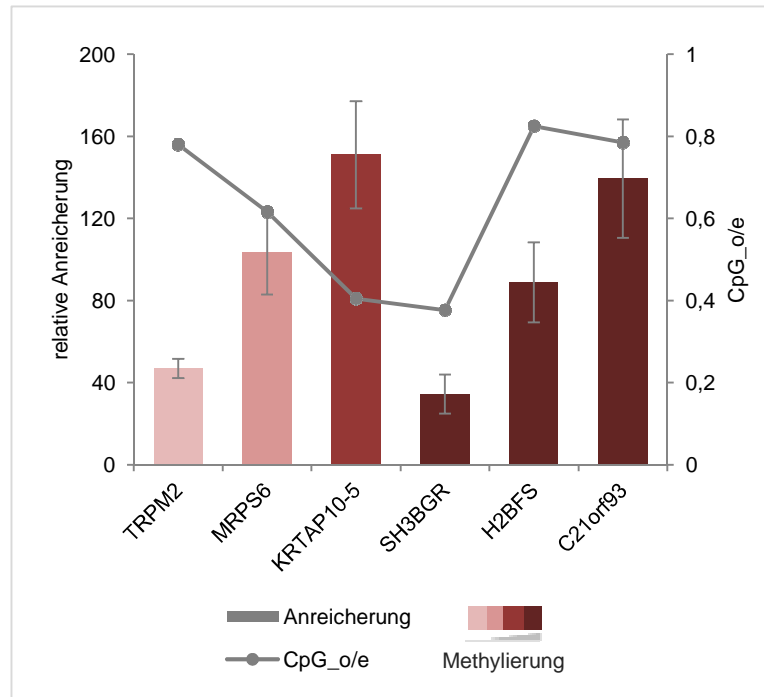


Abbildung 3.1-2: Relative Anreicherung verschiedener Kontrollregionen in Abhängigkeit des Methylierungslevels und CpG-Gehalts. Mittels qPCR wurde die Anreicherung methylierter Kontrollregionen im Verhältnis zum Hintergrund der CpG-freien Region bestimmt und im Vergleich zum jeweiligen CpG-Gehalt der untersuchten Region aufgetragen (CpG_o/e, „CpG observed/expected“). Die Regionen sind in aufsteigender Reihenfolge gemäß ihrer Methylierung (Zhang et al. 2009) dargestellt. Die gezeigten Ergebnisse entsprechen dem Mittelwert dreier unabhängiger MeDIP-Experimente an Leukozyten (682).

Die Erfassung der spezifischen Anreicherung definierter Kontrollregionen mit Hilfe der qPCR wurde im Zuge der laborinternen MeDIP-Etablierung und -Optimierung als wichtiges Mittel der Qualitätskontrolle verwendet. Dies erlaubte eine direkte Gegenüberstellung des Präzipitationsverhaltens und somit eine Bewertung der entsprechenden experimentellen Parameter sowie eine qualitative Einordnung jedes Experiments.

3.1.1.2 Randbedingungen einer erfolgreichen Präzipitation

Das Ausgangsmaterial ist entscheidend für ein erfolgreiches und reproduzierbares Experiment. Zum Einen ist die Qualität der genomischen DNA von Bedeutung, da sie nicht nur die spätere Fragmentierung sondern hauptsächlich die Effizienz und Spezifität der Antikörperbindung bestimmt. Wie in einer systematischen Versuchsreihe gezeigt werden

konnte, führen z.B. die Anwesenheit von RNA und insbesondere RNase zu erheblichen Hintergrundsignalen (Anhang A, Abbildung 6.1-1.).

Des Weiteren wird für eine effiziente Präzipitation bzw. Diskriminierung methylierter und unmethylierter Regionen zufällig fragmentierte DNA im Bereich von 250 bis 1000 nt mit einer mittleren Größe von 500 nt benötigt (Weber). Üblicherweise werden dazu Ultraschall-basierte Methoden herangezogen, deren Effizienz von mehreren Faktoren wie Temperatur, DNA-Qualität, DNA-Konzentration und Volumen abhängig ist. Ausgehend von diesen Parametern wurden in der vorliegenden Arbeit Protokolle für kleinere und größere Ausgangsmengen an DNA etabliert (Anhang A, Abbildung 6.1-2 (A), zeigt exemplarisch die zunehmende Fragmentierung in Abhängigkeit der Scherzyklen). Die Größenverteilung wurde stets mittels Gelelektrophorese überprüft und die DNA-Konzentration mit Hilfe des Qubit-Systems verifiziert.

Trotz einer standardisierten Vorgehensweise stellte sich diese Methode allerdings als sehr variabel insbesondere hinsichtlich leicht abweichender DNA-Mengen und der DNA-Qualität heraus. Dies erforderte oftmals zusätzliche Scherzyklen. Sogar Replikate derselben DNA lieferten häufig abweichende Ergebnisse (Anhang A, Abbildung 6.1-2 (B)). Daher erfolgte die Umstellung auf die robustere Nebulisierung, bei welcher das Zerstäuben einer Flüssigkeit zur Fragmentierung der in ihr gelösten DNA führt.

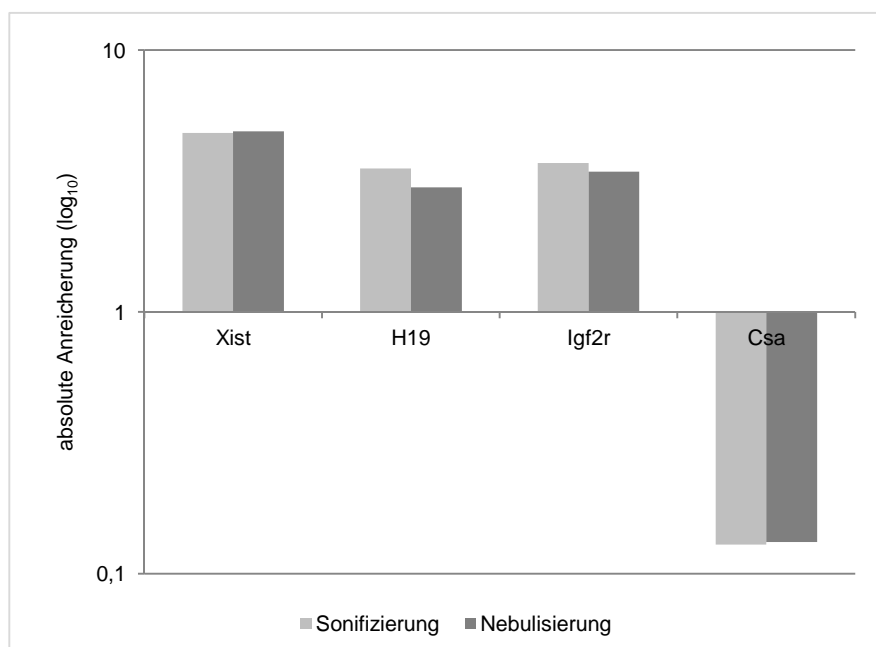


Abbildung 3.1-3: Vergleich verschiedener Fragmentierungsmethoden. Sonifizierte und nebulisierte DNA (*Mus musculus*, Leber) wurden zur Immunopräzipitation eingesetzt (gleiche Versuchsbedingungen) und die Präzipitate mittels qPCR analysiert. Dabei sind keine wesentlichen Unterschiede bzgl. der absoluten Anreicherung (logarithmische Skalierung) verschiedener Kontrollregionen zu verzeichnen.

Diese Methode lieferte stets reproduzierbare Größenverteilungen und war zudem erheblich zeiteffektiver. Im direkten Vergleich mit der Sonifizierung konnte bzgl. der spezifischen Anreicherung methylierter oder unmethylierter Kontrollregionen kein qualitativer Unterschied festgestellt werden (Abbildung 3.1-3 (B)).

3.1.1.3 Präzipitationseffizienz in Abhängigkeit des AK:DNA-Verhältnisses

Neben Reinheit der genomischen DNA und Größenverteilung der Fragmente bestimmt auch das Verhältnis der eingesetzten AK:DNA-Mengen wesentlich die Effizienz und Qualität der Präzipitation. Um ein optimales Gleichgewicht zwischen Ausbeute und Spezifität zu ermitteln, wurden unterschiedliche AK:DNA-Verhältnisse von 1:1 bis 2,5:1 an humaner DNA getestet und mittels Qubit-Messung und qPCR bewertet (Abbildung 3.1-4).

Die Präzipitationseffizienz stieg bei zunehmender Menge eingesetzten Primärantikörpers an, erreichte ihr Maximum bei einem AK:DNA-Verhältnis von 2:1 und ging dann in einen Sättigungseffekt über (Abbildung 3.1-4 (A)). Dabei konnte eine Abnahme der relativen Anreicherung getesteter Kontrollregionen verzeichnet werden, was auf die zunehmende unspezifische Präzipitation unmethylierter Regionen aufgrund des AK-Überschusses zurückzuführen war (Abbildung 3.1-4 (B) und (C.1)). Das AK:DNA-Verhältnis mit maximaler Ausbeute (2:1) wies dabei noch eine etwa 100-fache Anreicherung methylierter DNA im Vergleich zum Hintergrund auf und stellte somit einen guten Kompromiss zwischen AK-Verbrauch (Kostenpunkt), Effizienz und Qualität des Präzipitats dar. Ein vergleichbares Experiment an murinem Ausgangsmaterial führte zu ähnlichen Resultaten, so dass im Folgenden ebenfalls ein AK:DNA-Verhältnis von 2:1 gewählt wurde. Interessanterweise zeigten diese Versuche einige Unterschiede bzgl. humanen und murinen Materials auf. Zum Einen schien im Falle der murinen Versuchsreihe die Abnahme der relativen Anreicherung mit zunehmender AK-Menge weniger ausgeprägt. Des Weiteren konnte ein erheblicher, etwa zehnfacher Unterschied in der Skala der relativen Anreicherung beobachtet werden (Abbildung 3.1-4 (C.2)). Dies bleibt zu diskutieren.

Vergleichende Experimente wurden zudem für AK unterschiedlicher Herkunft anhand des festgelegten AK:DNA-Verhältnisses durchgeführt, zeigten aber keine wesentlichen Unterschiede bzgl. der getesteten Kontrollregionen auf. Zusätzliche Versuche beschäftigten sich mit einer weiteren Optimierung der Präzipitation unter Beeinflussung verschiedener Schritte des Experiments wie Denaturierung, Dauer der Inkubation mit dem Primärantikörper oder der Aufreinigung. Da diese Punkte nur zu minimalen Veränderungen führten, sollen sie nicht im Einzelnen dargestellt werden.

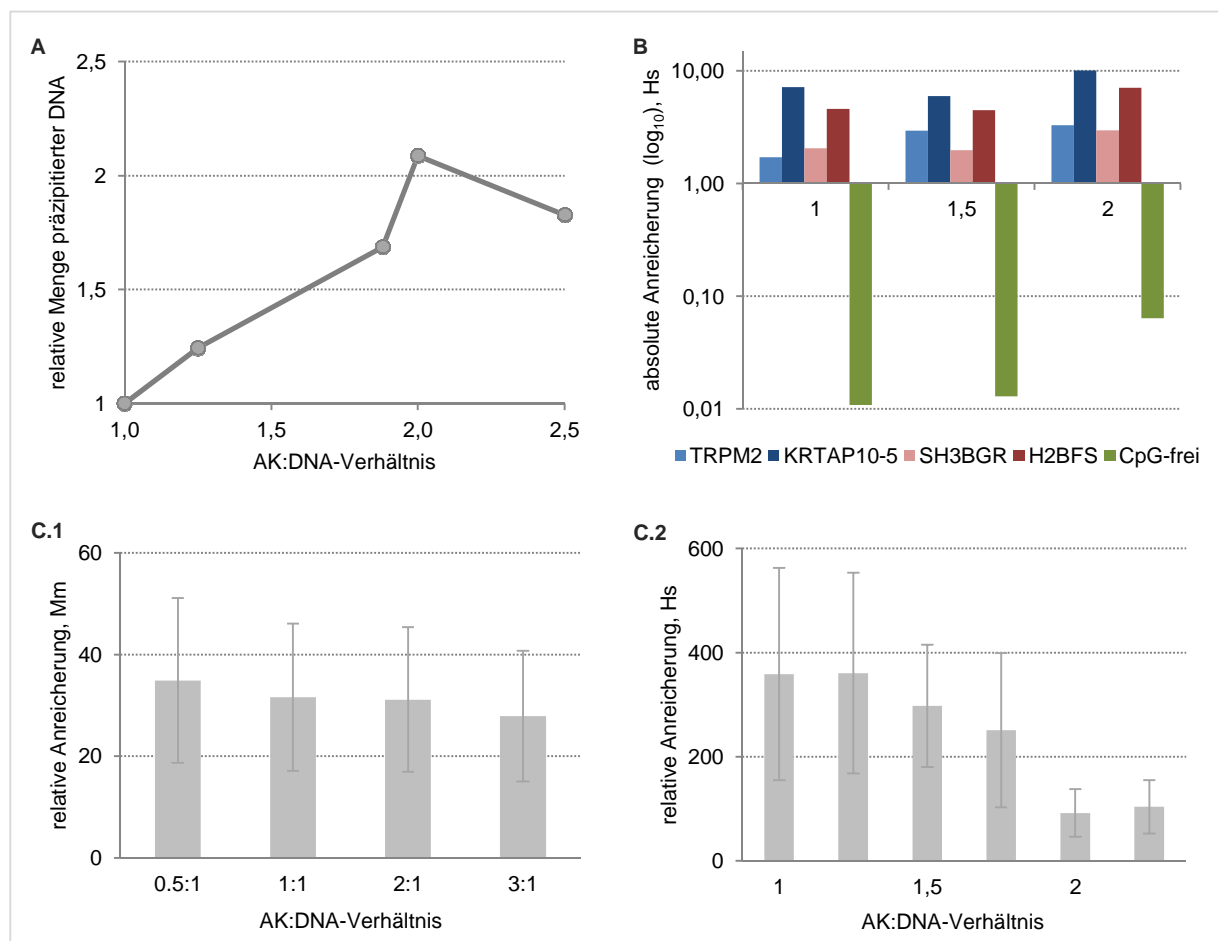


Abbildung 3.1-4: Anreicherung in Abhängigkeit des AK:DNA-Verhältnisses. Unterschiedliche AK:DNA-Ratios (0.5:1 bis 3:1) wurden für humane und murine DNA zur Immunpräzipitation eingesetzt und mittels qPCR sowie DNA-Messung evaluiert (Hs: Leukozyten (682), Mm: Leber). (A) Steigende Ausbeute mit steigendem AK:DNA-Verhältnis. Dargestellt ist jeweils die präzipitierte DNA-Menge im Vergleich zum kleinstem getesteten Verhältnis (1:1) für eine Versuchsreihe an humaner DNA. (B) Absolute Anreicherung (logarithmische Skalierung) verschiedener Kontrollregionen in Abhängigkeit des AK:DNA-Verhältnisses. (C.1) und (C.2) Mittlere relative Anreicherung methylierter Kontrollregionen in Abhängigkeit des AK:DNA-Verhältnisses in Maus (C.1, Regionen: H19, Igf2r) und Mensch (C.2; Regionen: TRPM2, KRTAP10-5, SH3BGR, H2BFS).

3.1.1.4 Einfluss der genomrepräsentativen Amplifikation (WGA) auf die Qualität des Präzipitats

Im Falle geplanter Microarray-Hybridisierungen ist in der Regel eine Amplifikation des angereicherten Materials nötig. Dienstleister wie Imagenes verlangen üblicherweise eine Menge $>3 \mu\text{g}$ DNA für Farbstoffmarkierung („labeling“) und Qualitätskontrollen. Je nach eingesetztem Ausgangsmaterial wurden allerdings pro Experiment lediglich 40-200 ng präzipitiert und ein entsprechendes Hochskalieren des Versuchsansatzes war nur selten möglich, so dass eine Amplifikation des Materials vonnöten war.

In der vorliegenden Arbeit wurde das vom Dienstleister empfohlene kommerzielle Kit (WGA2, Sigma) zur genomrepräsentativen Amplifikation (WGA, „whole genome amplification“) verwendet. Um die bei solchen Methoden üblicherweise auftretenden Verzerrungen (Jia et al. 2010, Robinson et al. 2010, Johnson et al. 2008) zu minimieren, wurden die Proben stets parallel unter Verwendung desselben Kits bearbeitet, wobei Präzipitate und Input Kontrollen gleichermaßen behandelt wurden. Alle Proben wurden stets nach der Amplifikation nochmals mittels qPCR überprüft. Dabei war in der Regel eine Verschiebung der CT-Werte hin zu späteren Zyklen zu beobachten. Dieser Effekt war in amplifizierten Input Kontrollen (+WGA_Input) und Präzipitaten (+WGA_IP) unterschiedlich stark ausgeprägt (Abbildung 3.1-5 (A)), wobei sich eine geringere Differenz der CT-Werte für Input-IP ergab und somit eine verringerte absolute Anreicherung.

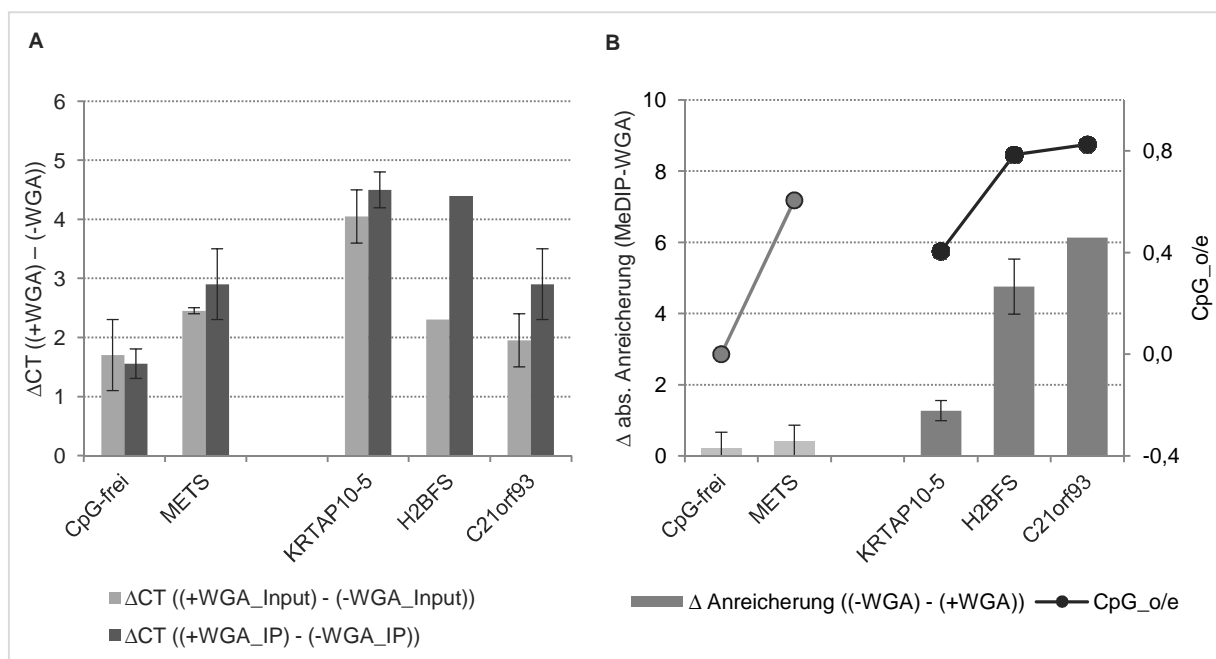


Abbildung 3.1-5: Effekt der genomrepräsentativen Amplifikation. Präzipitate wurden vor und nach WGA-Amplifikation mittels qPCR evaluiert. (A) Ungleichmäßige Verschiebung des CT-Wertes nach WGA für Input und IP in verschiedenen Negativ- und Positivkontrollen (CpG-frei und METS vs. KRTAP10-5, H2BFS und C21orf93). Dargestellt sind jeweils die mittleren delta-CT-Werte von bis zu drei Experimenten. (B) Einfluss der WGA auf die absolute Anreicherung verschiedener Negativ- und Positivkontrollen (CpG-frei und METS vs. KRTAP10-5, H2BFS und C21orf93) in Abhängigkeit des CpG-Gehalts (CpG_o/e, „CpG observed/expected“). Dargestellt sind jeweils die mittleren delta-Anreicherungswerte von bis zu drei Experimenten (Leukozyten (MC, 682), HepG2 NKtr).

Des Weiteren konnte eine Verzerrung in Abhängigkeit des CpG-Gehalts festgestellt werden, wobei CG-reiche Regionen tendenziell weniger stark amplifiziert wurden als CG-arme Regionen (Abbildung 3.1-5 (B)). Dies führte zudem zu einer verringerten spezifischen Anreicherung gegenüber dem Hintergrund (CpG-freie Regionen). Ein chromosomenweiter

Vergleich von MeDIP-Chip-Daten basierend auf amplifiziertem und nicht-amplifiziertem Material (Abschnitt 3.1.2.3, S.81) zeigte dennoch sinnvolle Resultate, so dass diese Methode trotz der festgestellten und bekannten „Mängel“ für humanes, nicht aber murines Material verwendet wurde. Bei Letzterem machten sich die Verzerrungen aufgrund der ohnehin niedrigeren Anreicherungswerte (Abschnitt 3.1.1.3, S.67) stärker bemerkbar und reduzierten die Qualität des Experiments erheblich.

3.1.2 Etablierung bioinformatischer Bausteine zur Überprüfung und Analyse der Daten aus MeDIP-Chip

Zur Generierung genomweiter bzw. chromosomenweiter Methylierungsmuster sollten in der vorliegenden Arbeit hochauflösende „*Tiling*“-Arrays (NimbleGen) verwendet werden. Dabei sind definierte Promotor-Bereiche oder kontinuierliche chromosomale Abschnitte in 100 bp-Fenstern mit Oligonukleotiden abgedeckt (siehe auch Abschnitt 2.1.9). Durch Ko-Hybridisierung von IP (Cy5, rot) und Input (Cy3, grün) kann über das \log_2 -Verhältnis ihrer Signalintensitäten die Anreicherung methylierter Fragmente pro Oligonukleotid bestimmt werden.

Die erforderlichen Schritte wie Farbstoffmarkierung und Hybridisierung der entsprechenden MeDIP-Proben wurden dabei von einem Dienstleister (ImaGenes, Berlin) vorgenommen, welcher anschließend die Rohdaten übermittelte. Die anschließende bioinformatische Analyse der Daten einschließlich Qualitätskontrolle, Prä-Prozessierung und Identifizierung differentiell angereicherter Regionen wurde, aufgrund noch nicht verfügbarer Protokolle oder Programme, selbst erarbeitet. Dazu wurden Elemente gängiger Methoden zur Auswertung von Expressions- oder ChIP-Chip-Daten kombiniert bzw. als Anregung verstanden. Quantitative Methylierungsdaten (Bisulfit-Sequenzierung) dienten dabei einer kritischen Überprüfung des Ansatzes. Der folgende Abschnitt stellt diese verschiedenen Analyse-Schritte (siehe auch Abbildung 3.1-6) unter Herausarbeitung kritischer Aspekte dar.

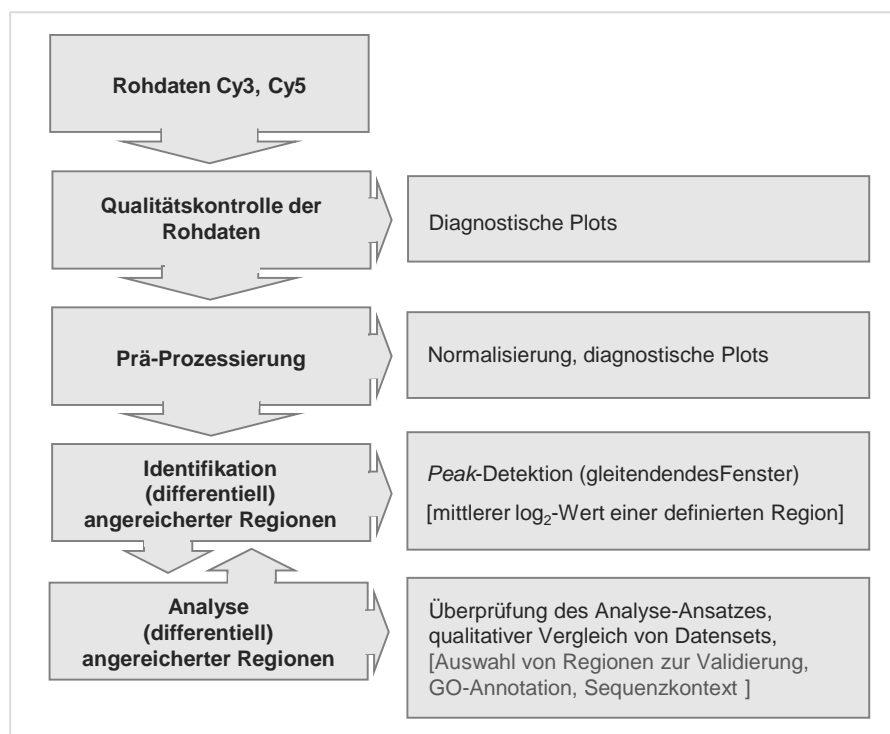


Abbildung 3.1-6: Ablauf der Datenanalyse. Übersicht über die Teilschritte der Auswertung von MeDIP-Chip-Daten. Die im Folgenden Abschnitt nicht berücksichtigten Analyse-Schritte sind durch Klammerung markiert.

3.1.2.1 Qualitätskontrollen und Prä-Prozessierung der MeDIP-Chip-Daten

Vor der tatsächlichen Auswertung der Methylierungsdaten, d.h. der Identifizierung (differenziell) angereicherter Regionen, sollte stets eine gründliche Überprüfung der Rohdaten erfolgen. Technische Variabilitäten, systematische Signal-Verzerrungen und sonstige Artefakte, welche eine sinnvolle Interpretation der Daten beeinträchtigen könnten, müssen identifiziert und gegebenenfalls korrigiert werden (Huber et al. 2005). Dazu wurde aufgrund der großen Datenmengen (>350.000 Datenpunkte pro Farbkanal pro Array) und der erforderlichen komplexen Analyse hinsichtlich Statistik und Visualisierung die freie Programmiersprache R (R.D.C.Team 2010, Ihaka und Gentleman 1996) herangezogen. Eigene R-Skripte konnten insbesondere unter Verwendung der Software-Pakete *Ringo* (Toedling et al. 2007) und *limma* (Smyth 2005, Smyth und Speed 2003) erarbeitet werden.

3.1.2.1.1 Qualitätskontrolle der Rohdaten mit Hilfe diagnostischer Plots

Mit Hilfe der entsprechenden Befehle des *Ringo*-Pakets, speziell für Datenstrukturen der auch in dieser Arbeit verwendeten NimbleGen-Arrays entwickelt, wurde R Zugriff auf die

vom Dienstleister erhaltenen Rohdaten gegeben. Diese konnten durch Erstellung einer Reihe diagnostischer Plots auf systematische Verzerrungen untersucht und zwischen verschiedenen Arrays verglichen werden. Erst in Kombination gaben die Plots ausreichend Auskunft über die Qualität und Reproduzierbarkeit der Experimente.

Zunächst wurde stets eine Rekonstruktion des Array-Bilds vorgenommen, wobei die Signale der einzelnen Oligonukleotide gemäß ihrer Koordinaten dargestellt wurden. Auf diese Weise ließen sich systematische Hybridisierungs-Artefakte wie Kratzer, Belichtungsfehler oder ungleichmäßige Hybridisierung erkennen. Größere Defekte wie in Abbildung 3.1-7 (A) wurden dabei nicht toleriert, wohingegen kleinere Unregelmäßigkeiten in der Regel vernachlässigt werden konnten.

Die Verteilung der Signalintensitäten einzelner Kanäle bzw. \log_2 -Ratios wurde mit Hilfe verschiedener Plots vor und nach der Prä-Prozessierung visualisiert. Boxplots (Abbildung 3.1-7(B)) und Dichteplots (Abbildung 3.1-7 (C)) wurden dabei für einfache und schnelle Vergleiche mehrerer Proben oder Kanäle herangezogen, um so abweichende Verteilungen der Signalintensitäten bzw. \log_2 -Werte zwischen Proben oder Replikaten identifizieren zu können. Kleinere Unterschiede, wie in Abbildung 3.1-7 (B) zwischen Replikaten 1.a und 1.b zu sehen, traten üblicherweise im Rahmen technischer Variabilität auf und konnten mittels Normalisierung ausgeglichen werden.

Auskunft über das Verhältnis der Signalintensitäten der einzelnen Kanäle zueinander gaben zusätzlich MA-Plots (Abbildung 3.1-7 (D)). Auffällig war hierbei die gewölbte Kurvenform, welche sich grundsätzlich von den eher flachen, um die Null-Linie zentrierten MA-Plots publizierter Zwei-Kanal-Arrays (z.B. Expressionsarrays) unterschied. Diese Form war jedoch typisch für alle durchgeführten MeDIP-Chips und ergab sich aus der unterschiedlichen Zusammensetzung der ko-hybridisierten angereicherten Probe (IP) vs. Kontrollprobe (Input), was sich bereits im Dichteplot an den verschobenen Intensitätsbereichen der Kanäle zeigt. Die spezifische Ausprägung der Kurve allerdings divergierte teilweise erheblich zwischen verschiedenen Experimenten und ließ Signalverzerrungen aufgrund abweichender experimenteller Parameter oder technischer Fluktuationen erkennen (siehe auch Abschnitt 813.1.2.3, S.81). Solche Abweichungen führten beim paarweisen Vergleich der Oligo- \log_2 -Werte ebenfalls zu niedrigen Pearson-Korrelationskoeffizienten ($<0,7$), welche in der Regel über 0,9 zwischen Replikaten und 0,8 zwischen unterschiedlichen Proben lagen.

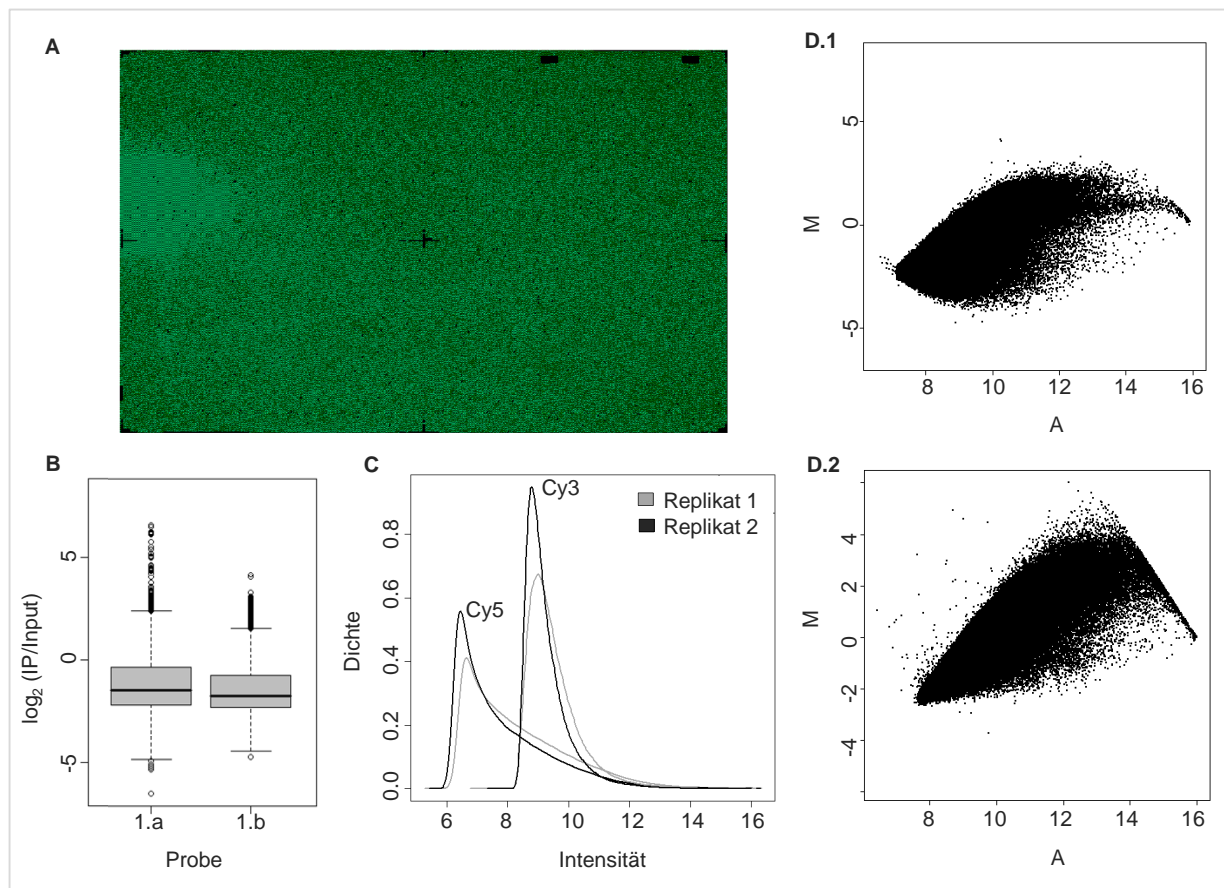


Abbildung 3.1-7: Diagnostische Plots der Qualitätskontrolle. Rohdaten der Cy3- und Cy5-Kanäle wurden mit Hilfe der Programmiersprache R überprüft. (A) Rekonstruktion eines Array-Bilds, wobei Cy3-Signale der einzelnen Oligonukleotide gemäß ihrer Position auf dem Array dargestellt sind. Zu erkennen ist eine größere Unregelmäßigkeit in Form eines Fingerabdrucks am linken Rand des Bilds sowie eine leichte vertikale Ungleichmäßigkeit bzgl. der Signalintensität (von oben nach unten abnehmend). (B) Boxplot der \log_2 -Ratios technischer Replikate 1.a und 1.b einer HepG2-Probe (Ktr). (C) Dichteplot der Signalintensitäten von Cy5- und Cy3-Kanälen derselben Replikate. Die Signale der Cy5-Kanäle (IP) bewegen sich in einem anderen Intensitätsbereich als die Signale der Cy3-Kanäle (Input). (D.1) MA-Plot der HepG2-Probe 1.b weist die übliche geschwungene Form der MeDIP-Chip-Daten auf, während die Kurve einer HEK293T-Probe auf Signalverzerrungen hinweist (D.2).

3.1.2.1.2 Prä-Prozessierung: Vergleich verschiedener Normalisierungsansätze

Die Normalisierung stellt einen bedeutenden Schritt der Datenanalyse dar. Dabei sollen Signalverzerrungen und systematische technische Unterschiede auf und zwischen Chips korrigiert werden (Bolstad et al. 2003, Quackenbush 2002). Abweichungen in z.B. der Effizienz der Farbstoffinkorporation oder Hybridisierungs- und Scan-Bedingungen beeinflussen dabei die Anreicherungswerte. Erst eine Standardisierung der Daten durch Anpassung oder Minimierung solcher nicht-biologischer Variabilität in gemessenen Signalintensitäten erlaubt die Detektion tatsächlicher biologisch relevanter Veränderungen zwischen den Proben.

In der vorliegenden Arbeit wurde eine Auswahl der zur Anwendung in R verfügbaren Algorithmen getestet und deren Effekt zum besseren Verständnis visualisiert. Auf diese Weise sollte eine Normalisierungs-Strategie ausgewählt werden, welche hinsichtlich der MeDIP-Chip-Daten ein Gleichgewicht zwischen Korrektur abweichender Signalverteilungen und der Erzeugung zusätzlicher artifizieller Verzerrungen bot. Einige dieser Ergebnisse sind in Abbildung 3.1-8 exemplarisch anhand technischer Replikate dreier verschiedener Proben mit Hilfe von Dichteplots, Boxplots und MA-Plots gezeigt.

Die dargestellten Methoden führten zu unterschiedlichen Angleichungen der Daten bzgl. der Signalintensitäten der einzelnen Kanäle bzw. deren \log_2 -Ratios. Die Median-Normalisierung beispielsweise resultierte in einer globalen Angleichung der \log_2 -Ratios, wobei die Signalverteilungen der Arrays lediglich zentriert und skaliert wurden, so dass eine verbesserte Vergleichbarkeit der Proben gegeben war (Abbildung 3.1-8 (A), Boxplot in (B)). Diese Prozessierung hatte allerdings keine Auswirkung auf weitere nicht-biologische Abweichungen, wie z.B. zwischen Replikaten a und b der Probe 3. Solche stärkeren Effekte konnten in Kombination mit der Quantil-Normalisierung, welche zusätzlich die Daten innerhalb aller Quantile angleicht, korrigiert werden. Dies galt insbesondere bei Durchführung der Normalisierung anhand ausschließlich der \log_2 -Ratios (Abbildung 3.1-8, Boxplot in (B)). Wurde das gesamte Datenobjekt unter Einbeziehung der einzelnen Kanäle berücksichtigt, führte dies zu einer schwächeren Inter-Array-Normalisierung bei gleichzeitiger Angleichung der Cy3- und Cy5-Signalintensitäten. Letztere bewirkte außerdem eine Verzerrung der MeDIP-spezifischen Kurvenform des MA-Plots (Abbildung 3.1-8, MA-Plots in (B)). Ein ähnlicher Effekt wurde durch die üblicherweise für Expressions-Arrays anerkannte Methode der LOWESS-Regression („*locally weighted linear regression*“) (Cleveland und Devlin 1988, Yang et al. 2002) hervorgerufen. Dieser insbesondere zur Intra-Array-Normalisierung geeignete Ansatz wird hauptsächlich zur Korrektur stärkerer Verzerrungen oder lokaler Effekte eingesetzt, welche mit Hilfe der oben beschriebenen Median- oder Quantil-Normalisierung nicht erreicht werden können. Die Intensitäts-abhängige Anpassung nicht-linearer Effekte mit Hilfe einer glättenden Funktion ergab im MA-Plot die nahezu symmetrische Signalverteilung um Null (Abbildung 3.1-8, (B)).

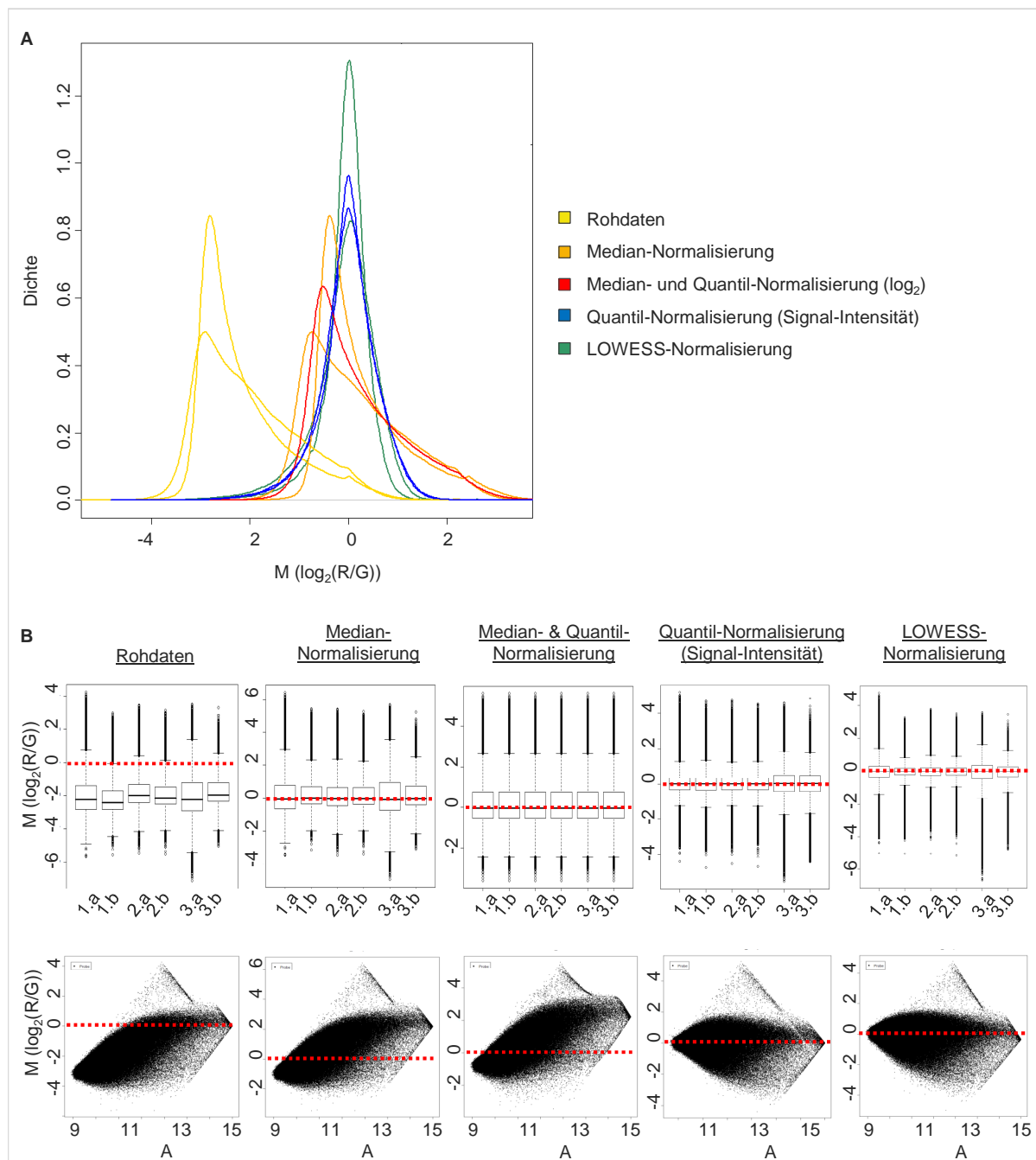


Abbildung 3.1-8: Effekt verschiedener Prä-Prozessierungs-Ansätze, dargestellt mittels Dichteplots (A), Boxplots und MA-Plots (B) im Vergleich zu Rohdaten. Insbesondere die verschiedenen Kurvenformen der MA-Plots (B) visualisieren den Effekt der jeweiligen Normalisierungsmethode auf das Verhältnis der Signalintensitäten der Kanäle zueinander. Dichteplots und MA-Plots basierend auf Leukozyten (MC), Boxplots basierend auf Leukozyten (MC, 1.a+b und MHT, 2.a+b), sowie HepG2 (NKtr, 3.a+b).

Anhand der dargestellten Ergebnisse wird deutlich, dass gängige Methoden zur Normalisierung von Zwei-Kanal-Arrays nur bedingt für MeDIP-Chip-Daten Verwendung finden können. Aufgrund der unterschiedlichen Komposition von IP und Input war eine

künstliche Angleichung der Cy3- und Cy5-Kanäle eines Arrays wie im Falle der LOWESS- oder Quantil-Normalisierung nicht sinnvoll. Wie auch im Vergleich zu quantitativen Bisulfit-Daten ersichtlich war (Anhang A, Tabelle 6.1-3), führten die erheblichen Datenmanipulationen eher zu einer Verschlechterung der Datenqualität. Daher wurden im Folgenden lediglich Inter-Array-Korrekturen auf Basis der \log_2 -Ratios mit Hilfe der Median- und Quantil-Normalisierung (im späteren Verlauf der Arbeit auch Skalierungs-Normalisierung) durchgeführt.

3.1.2.2 Identifizierung methylierter oder differentiell methylierter Regionen und Evaluierung im Vergleich zu quantitativen Methylierungsdaten

Erst auf Basis der erfolgreich normalisierten Daten können signifikant angereicherte Regionen analysiert und zwischen mehreren Proben verglichen werden. Die verschiedenen hierbei möglichen Ansätze, welche sich aus Array-Design und biologischer Fragestellung ergeben, berücksichtigen dabei in der Regel mehrere benachbarte Oligonukleotide, um Fehlinterpretationen aufgrund einzelner Ausreißer zu vermeiden. Dies kann zudem durch die Verwendung biologischer oder technischer Replikate unterstützt werden, da auch deren Standardabweichung Auskunft über nicht-signifikante Variationen geben können.

Die Identifizierung methylierter bzw. differentiell methylierter Regionen im Falle der in Abschnitt 3.3 verwendeten Promotor-Arrays war dabei aufgrund verfügbarer Anhaltspunkte aus der Literatur, entsprechender Programme sowie des einfachen Prinzips problemlos durchführbar (Weber et al. 2007, Weber et al. 2005, Mohn et al. 2008, de Las Heras et al. in Vorbereitung). \log_2 -Ratios eines definierten Abschnittes um die Transkriptionsstartstelle (TSS) wurden gemittelt und erlaubten somit durch Definition spezifischer Schwellenwerte die Unterscheidung in unmethylierte, methylierte oder differentiell methylierte Regionen (siehe Teil 3.3).

In der vorliegenden Arbeit war allerdings ebenfalls die Erfassung von Methylierungsmustern über kontinuierliche Genom-Abschnitte, genauer gesagt komplette Chromosomen, von Interesse (3.2). Da die Relevanz solcher Analysen erst durch neuere Publikationen, welche die Bedeutung CpG-armer Ausläuferregionen der CpG-Inseln sowie der Genkörper selbst bzgl. gewebe- oder krankheitsspezifischer Methylierungsmuster hervorhoben (Ball et al. 2009, Dindot et al. 2009, Irizarry et al. 2009), erkannt wurde, war die Erarbeitung einer geeigneten Strategie zur Erhebung solcher Profile erforderlich. Quantitative Methylierungsdaten von >260 Regionen aus Bisulfit-Sequenzierung (Zhang et al. 2009) erlaubten dabei eine kritische Evaluierung des Ansatzes sowie der Daten-Qualität.

3.1.2.2.1 Identifizierung methylierter Regionen

Zur Identifikation methylierter Regionen wurden Peak-Detektions-Algorithmen herangezogen. Solche Analysen wurden bereits für ähnliche Anwendungen wie Chip-ChIP-Experimente (Johnson et al. 2008) beschrieben und neben den dazu vielfach verfügbaren Programmen bietet der Dienstleister (NimbleGen) sogar eine für Methylierungsanalysen geeignete Version. Basierend auf den \log_2 -Ratios wird in einem gleitenden Fenster die Wahrscheinlichkeit einer Anreicherung jedes Oligonukleotids unter Berücksichtigung der umgebenden Fragmente berechnet. Benachbarte Oligonukleotide über einem spezifischen Schwellenwert werden dann zu Anreicherungspeaks (AP) zusammengefasst.

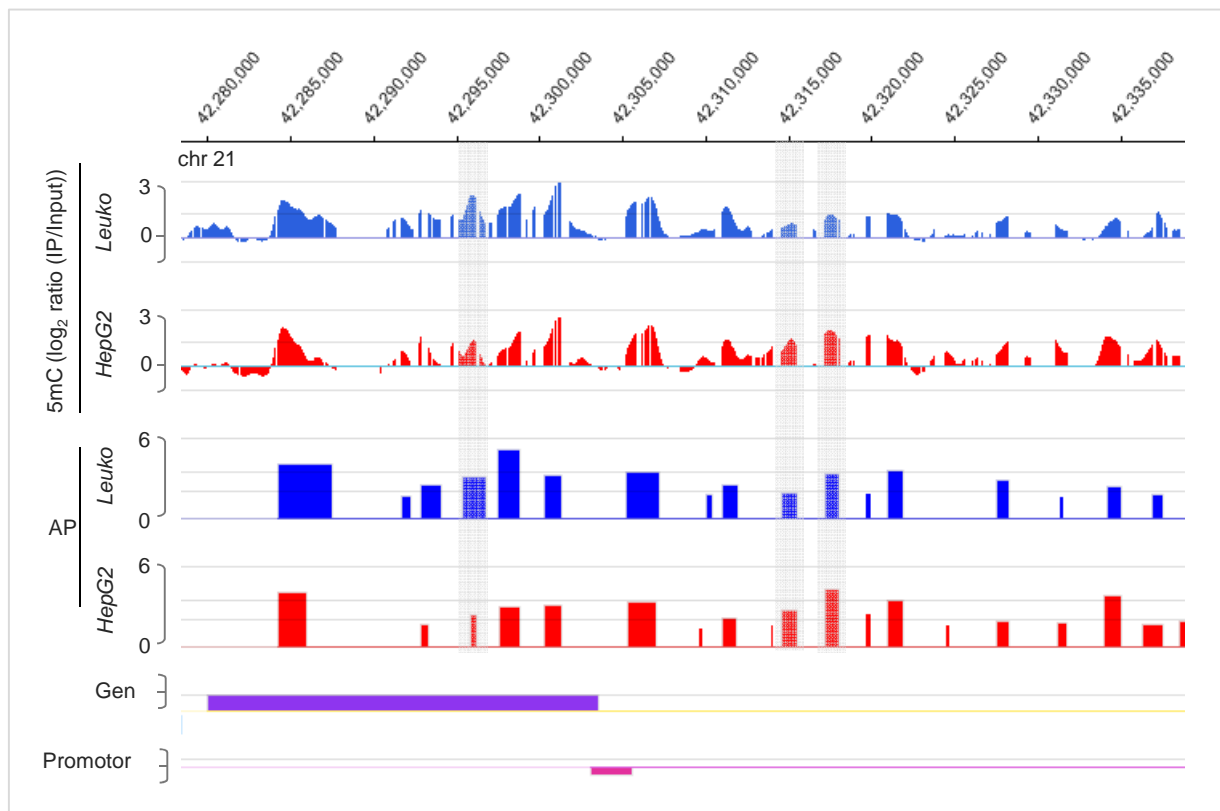


Abbildung 3.1-9: Graphische Darstellung der Anreicherungsprofile zweier humaner Zelltypen. \log_2 -Ratios (geglättet) und Anreicherungspeaks zweier Zelltypen (Leukozyten (MC) und HepG2-Zellen *NKtr*) sind relativ zu ihren genomischen Koordinaten sowie funktionellen Regionen für einen Ausschnitt des Chromosoms 21 dargestellt (Signalmap, NimbleGen). Grau markiert sind Abschnitte differenzieller Anreicherung, welche durch Evaluierung von APs nicht erfasst werden können.

Abbildung 3.1-9 veranschaulicht exemplarisch die Profile der Oligonukleotid- \log_2 -Ratios zweier verschiedener Gewebe über einen Abschnitt des Chromosoms 21 mit den entsprechend identifizierten angereicherten Regionen in Form von Peaks.

Zur Evaluierung dieses Peak-Algorithmus, welcher den Ausgangspunkt der nachfolgenden Identifikation differentiell methylierter Regionen darstellen sollte, wurden bereits publizierte gewebespezifische Methylierungsdaten aus Bisulfit-Klonierung und Sequenzierung von >260 Promotor-Regionen des Chromosoms 21 herangezogen (Zhang et al. 2009). Dazu wurde eine ROC-Analyse (ROC, „receiver operator characteristic“) vorgenommen, bei welcher Sensitivität und Spezifität einer diagnostischen Variablen, welche in diesem Fall durch die An- oder Abwesenheit eines APs in einer Promotor-Region bekannten Methylierungsstatus repräsentiert wurde, eingeschätzt werden können. Dazu wurden zunächst die genomischen Positionen der entsprechenden APs mit den sequenzierten Promotor-Regionen abgeglichen, welche in Abhängigkeit der bekannten Methylierungslevel kategorisiert wurden (methyliert, unmethyliert). Unter Verwendung des Peak-Scores (mittlere Anreicherungswahrscheinlichkeit der Oligonukleotide) als ordinale Beurteilungsskala wurden mit Hilfe des Web-basierten Programms JROCFIT Positiv- und Negativraten für verschiedene Schwellenwerte ermittelt und graphisch aufgetragen. Steile Kurven und damit größere Flächen unter der Kurve (AUC, „area under the curve“) stehen dabei für eine bessere Performance.

Abbildung 3.1-10 (A) zeigt ROC-Kurven verschiedener Gewebe unter Verwendung derselben experimentellen und diagnostischen Parameter. Für beide Proben war, bei einem Kategorisierungs-Schwellenwert von durchschnittlich 25% Methylierung, eine angemessene Identifikation methylierter Regionen mit AUCs >0,8 zu verzeichnen (Abbildung 3.1-10 (A) und (B)). Eine zunehmende Genauigkeit des Experiments aufgrund steigender Sensitivität konnte bei erhöhter Methylierung (in Abhängigkeit des zur Definition von Methylierung verwendeten Schwellenwerts) beobachtet werden (Abbildung 3.1-10 (C)).

Diese Ergebnisse zeigten, dass der verwendete Peak-Algorithmus ein geeignetes Werkzeug zur Identifizierung von Regionen mit >25% Methylierung darstellt und als Grundlage für die Analyse differentiell angereicherter Regionen dienen kann.

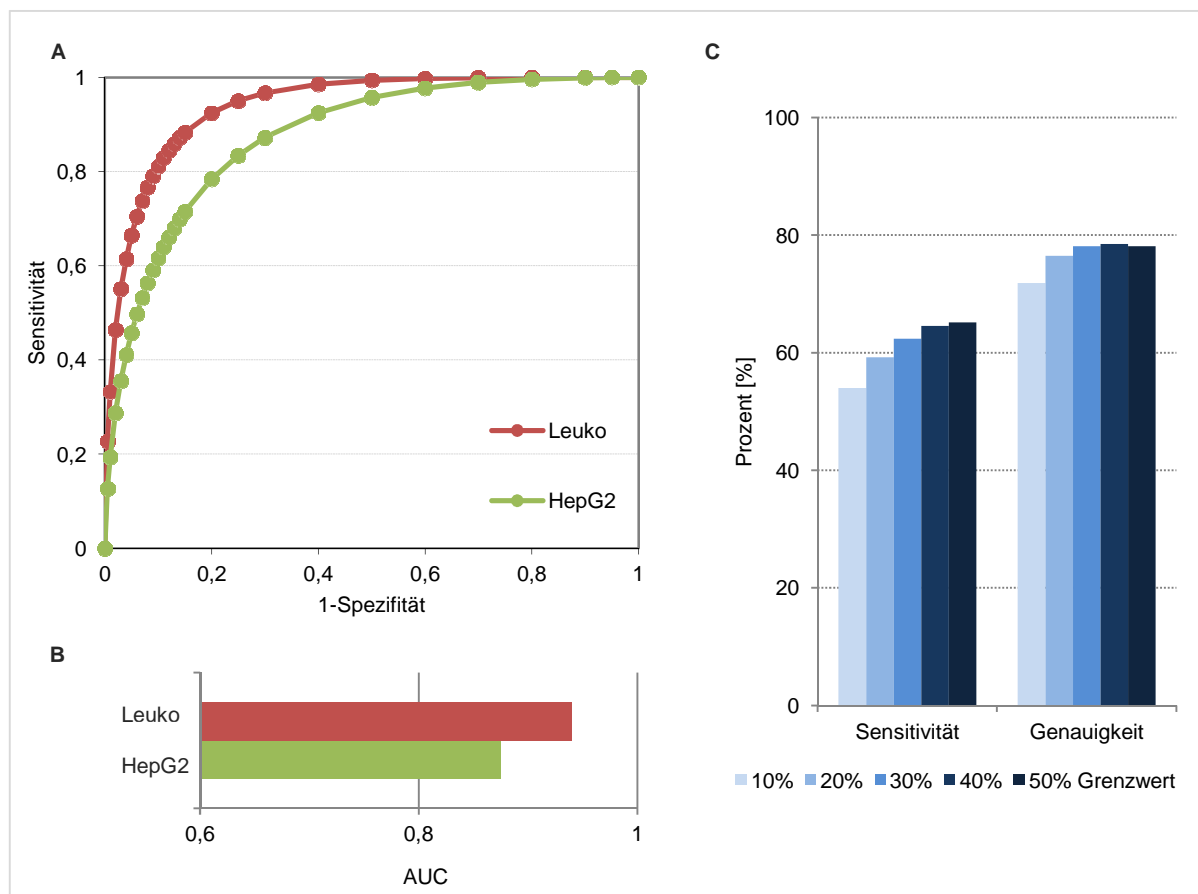


Abbildung 3.1-10: ROC-Analyse zur Überprüfung der Peak-Detektion. APs wurden mit quantitativen Methylierungsdaten für >260 Promotor-Regionen in Leukozyten (MC) und HepG2-Zellen (NKtr) abgeglichen und mit Hilfe des Programms JROCFIT evaluiert. (A) ROC-Kurven („fitted“) und (b) AUCs („area under curve“: Fläche unter der Kurve) bei einem Schwellenwert von 25% Methylierung weisen auf eine gute Performance hin. (C) Sensitivität und Genauigkeit der Analyse in Abhängigkeit des zur Definition von Methylierung verwendeten Schwellenwerts.

3.1.2.2 Identifizierung differentiell methylierter Regionen

Eine Identifizierung differentiell methylierter Regionen zwischen zwei Proben durch direkten Vergleich der entsprechenden APs wäre nur bedingt sinnvoll. Hierbei könnte nur die An- oder Abwesenheit der APs innerhalb zuvor definierter Genom-Abschnitte evaluiert werden, wodurch sich lediglich absolute Informationen bzgl. der Methylierung (Schwarz-Weiß-Bild) ergeben würden. Geringere Unterschiede methylierter Regionen, für welche APs in beiden zu vergleichenden Proben identifiziert wurden, könnten hingegen nicht erfasst werden (Abbildung 3.1-9). Zur Generierung entsprechend sensitiverer Profile wurde daher der zuvor beschriebene Peak-Detektions-Algorithmus auf Basis von $\Delta\log_2$ -Werten verwendet.

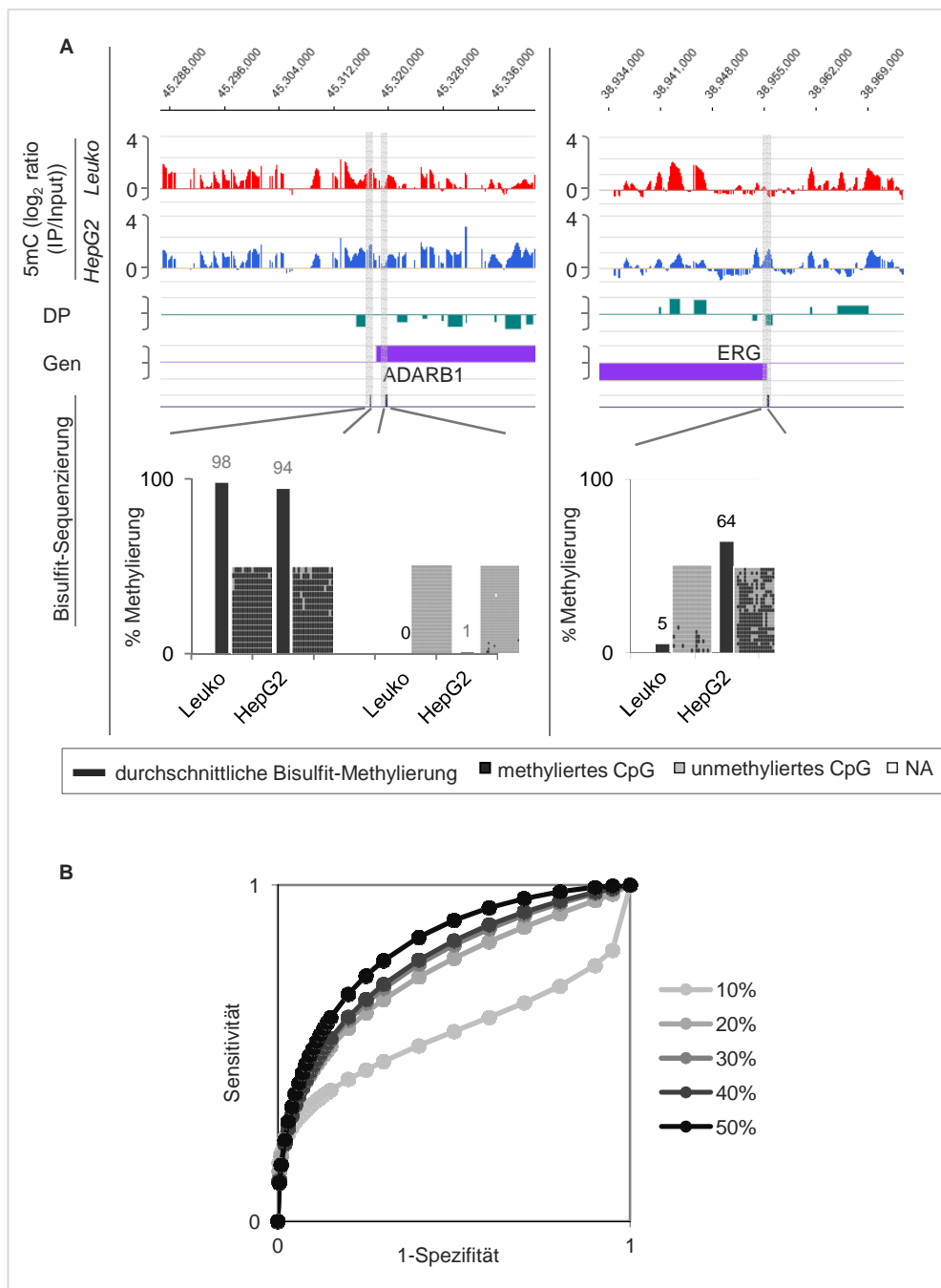


Abbildung 3.1-11: Überprüfung der Peak-Analyse zur Detektion differentieller Anreicherung. (A) Ausschnitte von Anreicherungsprofilen im Vergleich mit quantitativen Methylierungsdaten aus Bisulfit-Sequenzierung (Zhang 2009) für Leukozyten und HepG2-Zellen. Log₂-Ratios (geglättet), DPs, Gene (Ensembl) und Bisulfit-Amplikons sind relativ zu genomischen Koordinaten dargestellt (Signalmap, NimbleGen). Mittlere Methylierungslevel (Bisulfit-Sequenzierung) dreier Regionen mit unterschiedlichem Methylierungsstatus (methyliert, unmethyliert, mosaikal methyliert) sind darunter mit den entsprechenden Methylierungsmustern abgebildet, wobei individuelle Sequenzen in der Vertikalen und einzelne CpG-Positionen in der Horizontalen repräsentiert sind. (B) Evaluierung der DP-Analyse mittels ROC-Kurven („fitted“, berechnet mit JROCFIT) für verschiedene Schwellenwerte zur Definition differentieller Methylierung. DPs wurden dazu mit Methylierungsdaten aus Bisulfit-Sequenzierung von >260 Promotor-Regionen (Zhang et al. 2009) abgeglichen.

Detektierte Peaks repräsentierten somit Regionen differentieller Anreicherung benachbarter Oligonukleotide (DP: differentieller Anreicherungs-Peak). Um die Verlässlichkeit der erhaltenen Resultate zu erhöhen und eine „angemessene“ Diskriminierung von technischer oder biologischer Variabilität zu ermöglichen, wurden die DPs zusätzlich anhand verschiedener Kriterien und unter Einbeziehung der Replikat-Varianz gefiltert (2.2.3.2.2). Die finalen Regionen wiesen dann einen signifikanten, mindestens 1,4-fachen Anreicherungsunterschied zwischen den zu vergleichenden Proben auf. Entsprechende Beispiele im Vergleich zu bekannten Methylierungsdaten zweier Promotor-Regionen sind in Abbildung 3.1-11(A) dargestellt.

Zur Überprüfung des Ansatzes wurde, wie bereits für die APs beschrieben, eine ROC-Analyse vorgenommen. Dabei wurden die DPs unter Verwendung des Peaks-Scores als ordinale Bewertungsskala gegenüber bekannten gewebespezifischen Methylierungsdaten aus Bisulfit-Sequenzierung (Zhang et al. 2009) evaluiert. Sensitivität und Spezifität in Abhängigkeit des zur Kategorisierung verwendeten Schwellenwerts (differentiell methyliert, nicht differentiell methyliert) wurden in Abbildung 3.1-11 graphisch dargestellt. Die Kurven verdeutlichen, dass die identifizierten DPs mit relativ hoher Effizienz Methylierungsunterschiede ab 20% repräsentieren (AUC(20%): 0,74). Diese Resultate weisen auf das Auflösungsvermögen der verwendeten Methode MeDIP-Chip bzw. des gewählten Peak-Algorithmus hin und geben einen Eindruck der zu erwartenden Ergebnisse.

3.1.2.3 Einfluss experimenteller Parameter auf Qualität und Vergleichbarkeit der MeDIP-Chip-Daten

Durch die oben beschriebene ausführliche Evaluierung der MeDIP-Chip-Daten konnten kritische Teilschritte der gesamten Pipeline bestimmt werden, welche die Qualität und insbesondere die Vergleichbarkeit unterschiedlicher Experimente beeinflussen. Neben der bereits in 3.1.1.4 beleuchteten WGA spielen offensichtlich auch Hybridisierungseffekte oder eine Kombination verschiedener Parameter eine Rolle.

3.1.2.3.1 Einfluss der genomrepräsentativen Amplifikation

In Abschnitt 3.1.1.4 wurde die genomrepräsentative Amplifikation (WGA) bereits auf Ebene der qPCR untersucht und die Anwendbarkeit in Hinblick auf eine globalere Auswirkung hinterfragt. Um diese Problematik auch auf Array-Ebene zu analysieren, wurden Präzipitate der humanen Zelllinie HEK293T mit und ohne vorherige Amplifikation hybridisiert (+WGA bzw. -WGA, Abbildung 3.1-12 (A)).

Generell war auf Oligonukleotid-Ebene zunächst eine starke Ähnlichkeit aller Proben zu verzeichnen. Die im MA-Plot ersichtlichen Signalverteilungen der Kanäle im Verhältnis zueinander wiesen dabei keine wesentlichen Unterschiede auf (Daten nicht gezeigt). Auch der paarweise Vergleich aller Proben ließ mit Pearson-Korrelations-Koeffizienten $>0,93$ eine hohe Reproduzierbarkeit der Daten erkennen. Jedoch deuteten die gemessen an den Replikat-spezifischen Werten geringeren Behandlungs-spezifischen Koeffizienten bereits auf einen gewissen WGA-spezifischen Effekt hin (-WGA_1 vs. -WGA_2: 0,97; +WGA_1 vs. +WGA_2: 0,97; -WGA vs. +WGA: 0,93-0,94). Dies wurde durch die hierarchische Clusteranalyse verdeutlicht, bei welcher die Arrays nicht gemäß der Replikate sondern der Behandlung diskriminiert wurden (Abbildung 6.1-3, Anhang A).

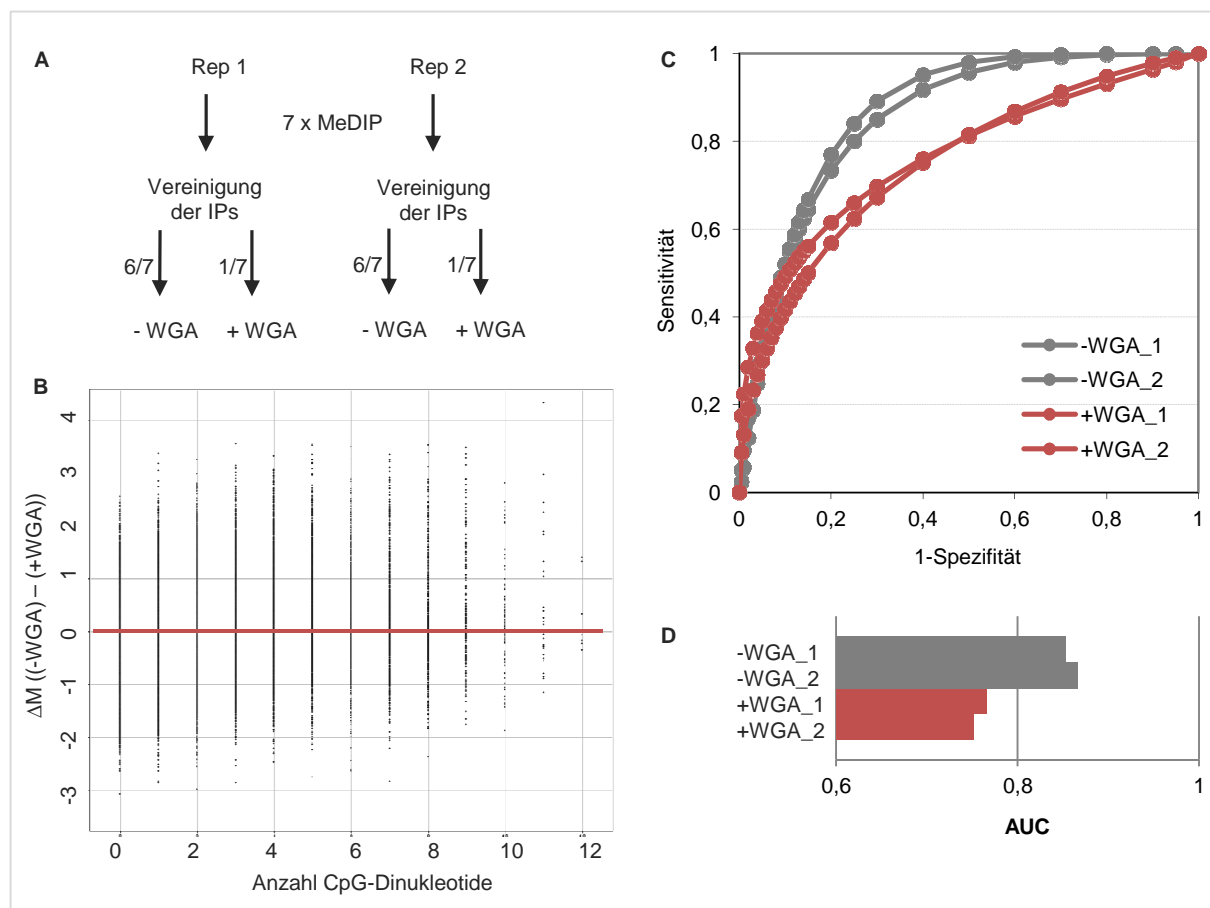


Abbildung 3.1-12: Einfluss der genomrepräsentativen Amplifikation auf die Anreicherung methylierter Regionen. (A) Die Präzipitate biologischer Replikate (Rep 1 und 2) von HEK293T-Zellen wurden mit und ohne genomrepräsentativer Amplifikation hybridisiert (+WGA bzw. -WGA). (B) Anreicherungsunterschied pro Oligonukleotid ($\Delta \log_2$) zwischen -WGA und +WGA in Abhängigkeit der jeweiligen CpG-Anzahl des Oligonucleotids. (C) ROC-Kurven und AUCs (D) zur Evaluierung der APs deuten auf eine verminderte Performance in +WGA vs. -WGA-Proben hin.

Unter Einbeziehung entsprechender Sequenzinformationen der einzelnen Oligonukleotide sollte die in 3.1.1.4 festgestellte ungleichmäßige Amplifikation verschiedener Genom-Bereiche in Abhängigkeit des CG-Gehalts näher analysiert werden. Auf Basis der jeweiligen $\Delta\log_2$ -Werte aller Oligonukleotide wurden in diesem Zusammenhang bis zu 6-fache Anreicherungsunterschiede zwischen -WGA und +WGA festgestellt. Insbesondere CpG-reiche Fragmente wiesen dabei systematisch geringere Anreicherungen der amplifizierten Proben auf (Abbildung 3.1-12 (B)). Um eine direkte Auswirkung dieser Verzerrung auf die Identifikation methylierter Regionen und damit die Interpretation der Daten zu ermitteln, wurde mit Hilfe quantitativer Methylierungsdaten eine ROC-Analyse, wie in Abschnitt 3.1.2.2.1 beschrieben, vorgenommen. Dabei ergab sich für +WGA mit $>0,76$ gegenüber -WGA mit $>0,86$ zwar eine etwas geringere, jedoch vertretbare AUC. Die Ähnlichkeit der Replikat-Kurven wies zudem auf eine Reproduzierbarkeit der amplifizierten Proben hin (Abbildung 3.1-12 (C) und (D)).

3.1.2.3.2 Vergleichbarkeit unabhängiger Experimente

Abgesehen von einem Einfluss der WGA konnten noch weitere Effekte auf die Qualität und Vergleichbarkeit verschiedener Datensets festgestellt werden. Dabei waren teilweise erhebliche Abweichungen zwischen unabhängigen Experimenten bzw. Hybridisierungen zu verzeichnen. Dies wird in Abbildung 3.1-13 anhand zwei verschiedener Leukozyten-Proben verdeutlicht, welche im Kontext zeitlich auseinanderliegender Versuchsreihen bearbeitet wurden.

Die spezifischen Kurvenformen der MA-Plots wiesen dabei auf Unterschiede in der Signalverteilung hinsichtlich der Intensitäten der einzelnen Kanäle im Verhältnis zueinander hin (Abbildung 3.1-13 (A) und (B)). Da ein direkter Einfluss auf die Güte der Dateninterpretation vermutet wurde, war eine Evaluierung der identifizierbaren methylierten Regionen nötig. Bei der hierzu durchgeführten ROC-Analyse konnte durch eine vergleichsweise geringere AUC von Leuko_1 (0,82; Leuko_2: 0,94) ein qualitativer Unterschied der betrachteten Experimente weiter herausgearbeitet werden (Abbildung 3.1-13 (C) und (D)). Während im Kontext der jeweiligen Versuchsreihen sinnvolle Resultate erhalten werden konnten, wurden keine weiteren vergleichenden Analysen zwischen den verschiedenen Experimenten durchgeführt, um Fehlinterpretationen zu vermeiden.

Die Ergebnisse dieses Abschnitts verdeutlichen nochmals die Bedeutung einer ausreichenden Überprüfung jedes Datensets, zur Erfassung biologisch relevanter Daten.

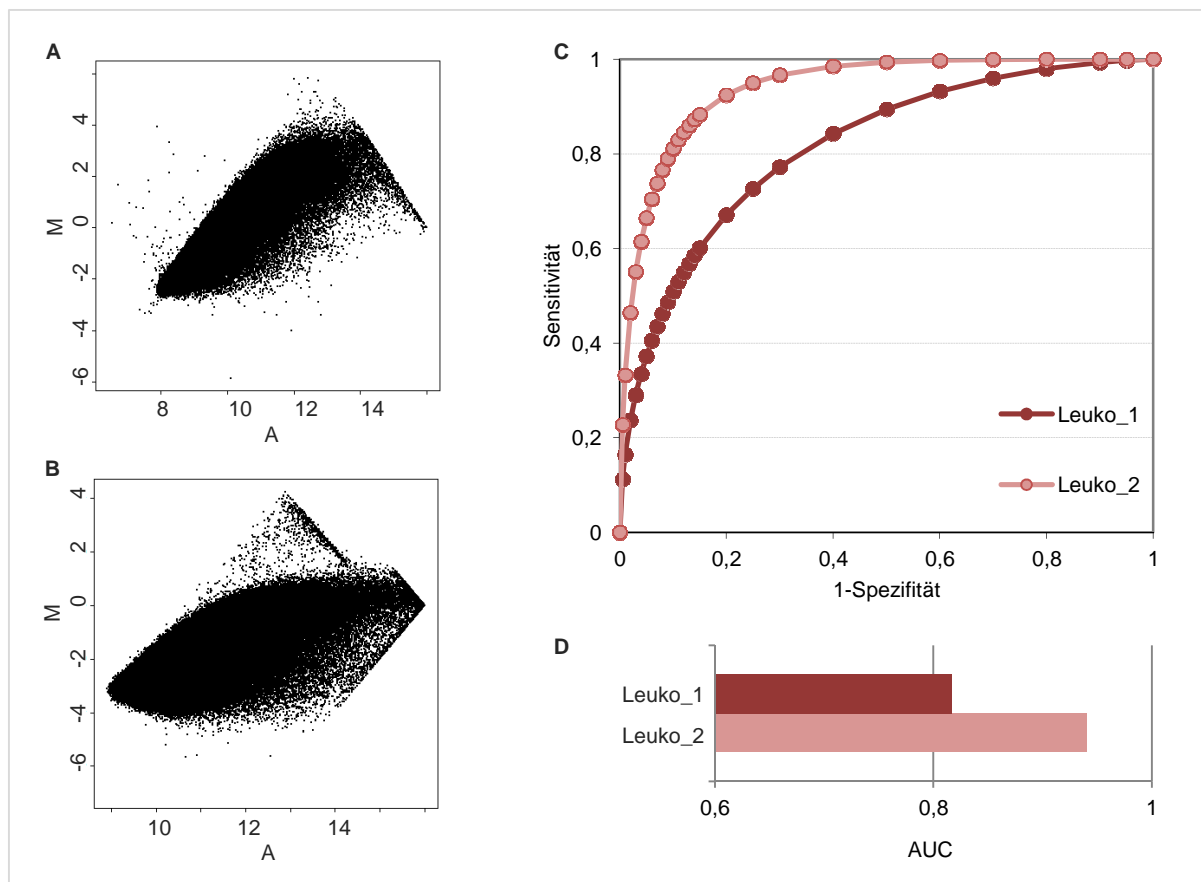


Abbildung 3.1-13: Vergleich der MeDIP-Chip-Daten unabhängiger Experimente des gleichen Zelltyps (Leukozyten). (A) und (B) MA-Plots der Experimente an Leukozyten-Proben 1 (676) und 2 (MC) deuten auf eine Signalverzerrung in Leuko_1 hin. (C) und (D) ROC-Analyse der Experimente an Leukozyten-Proben 1 und 2 deuten auf eine verminderte Identifizierung methylierter Regionen in Leuko_1 hin (Schwellenwert 25% Methylierung).

3.2 Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryopreservation

Kryopreservation biologischen Materials stellt ein verbreitetes Verfahren zur Konservierung lebender Zellen und Gewebe dar. Während ständig Fortschritte hinsichtlich der Reduktion physikalischen und biochemischen Stresses und damit einer Erhöhung der Viabilität des Materials erzielt werden, ist bisher wenig über potentielle sublethale Effekte bekannt, welche zu Langzeit-Veränderungen der Zellen führen könnten. Ausgehend von der aktuellen Literatur stellt DNA-Methylierung in diesem Zusammenhang einen mutmaßlichen Angriffspunkt dar (Wang et al. 2010, Milroy et al. 2011). Eine Schädigung dieser kovalenten DNA-Modifikationen durch während der Kryopreservation vermehrt gebildete Sauerstoffradikale (ROS, „reactive oxygen species“) (Niu et al. 2010, Rauen et al. 1999)

könnte vererbare epigenetische Veränderungen verursachen. Hierbei ist eine mögliche Verbindung zu deregulierter Expression verschiedener Gene denkbar, welche bei Kälte- und Hitzeschock beobachtet wurde (Liu et al. 2000, Sonna et al. 2006, Sonna et al. 2010). Auch kryoprotektive Substanzen wie DMSO selbst wurden bereits, im Falle längerer Kultivierung bei niedrigen Dosen, mit Änderungen epigenetischer Signaturen in Verbindung gebracht (Iwatani et al. 2006, Thaler et al. 2012). Eine Analyse möglicher stabil vererbbarer phänotypischer Veränderungen hinsichtlich Epimutagenese durch abiotischen Stress ist daher von großer Bedeutung, um die biologische Integrität konservierten Materials zu gewährleisten.

In der vorliegenden Arbeit sollten daher chromosomen- bzw. genomweite DNA-Methylierungsmuster und mRNA-Expressionslevel kryokonservierten Materials untersucht werden. Dazu wurde die hepatische Karzinomzelllinie HepG2 als Model-System herangezogen, da diese bereits früher mit Veränderungen der mRNA-Expression durch verschiedene Stressoren wie Hypothermie, Hyperthermie und Hypoxie in Verbindung gebracht wurde (Sonna et al. 2010, Sonna et al. 2003).

3.2.1 Kryokonservierung von HepG2 Zellen: experimentelles Design

Die Kryokonservierung wurde vom Fraunhofer-Institut für Biomedizinische Technik (IBMT), St. Ingbert, gemäß etablierter Standards durchgeführt, d.h. schrittweises Einfrieren bis -80°C in 10% DMSO-haltigem Kulturmedium (Anhang B, 6.2.1). Die Zellen wurden mindestens über Nacht im Endzustand belassen und nach dem Auftauen für einige Zeit zur „Erholung“ kultiviert. Dabei wurden Zeiträume von 48 h und 120 h gewählt (Proben *f48h* und *f120h*), da eine Manifestation möglicher Änderungen der DNA-Methylierung nicht unmittelbar nach der Kryopreservation erwartet wurde. Um mögliche auf das Einfrieren/Auftauen zurückzuführende Effekte von einem Einfluss des verwendeten Kryoprotektivums unterscheiden zu können, wurden zusätzlich DMSO-behandelte Zellen ohne voriges Einfrieren für 48 h kultiviert (Probe *DMSO*). Als Negativkontrolle dienten unbehandelte HepG2-Zellen, welche zu zwei verschiedenen Zeitpunkten des Dichtewachstums (nicht-konfluent und konfluent, Proben *NKtr* und *Ktr*) geerntet wurden. Dadurch sollten zusätzlich mögliche Wachstumseffekte nach Kryopreservation berücksichtigt werden.

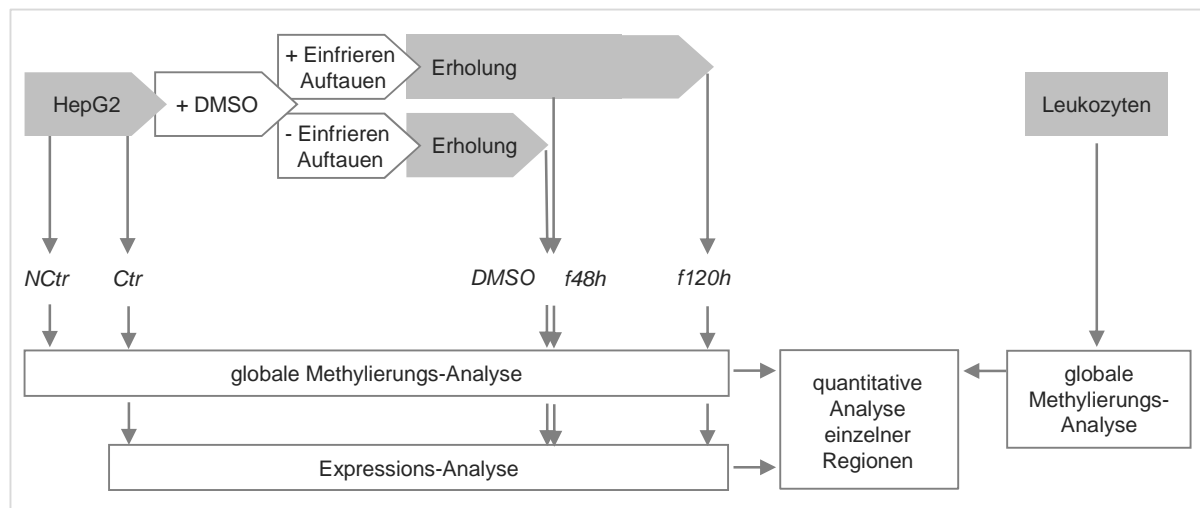


Abbildung 3.2-1: Schematische Darstellung des Kryopreservations-Experiments, inklusive Vorbehandlung mit dem Kryoprotektivum DMSO (10%, 30 min) und schrittweisem Einfrieren. Proben wurden nach 48- oder 120-stündiger Kultivierung (Erholungsphase) nachfolgend DMSO-Behandlung (Probe *DMSO*) bzw. Einfrieren/Auftauen (*f48h*, *f120h*) entnommen. Kontroll-Zellen nicht-konfluenter (*NKtr*) und konfluenter (*Ktr*) Kulturen wurden zusätzlich vor jeglicher Behandlung gesammelt. Chromosomenweite Methylierungsanalysen wurden für HepG2-Zellen und eine Positivkontrolle (Leukozyten *MC*) mit MeDIP-Chip durchgeführt. Proben *Ktr*, *DMSO*, *f48h* und *f120h* wurden zusätzlich hinsichtlich ihrer mRNA-Expression untersucht. Einzelne Zielregionen sowie repetitive Elemente wurden ferner mittels Bisulfit-Sequenzierung überprüft.

Abbildung 3.2-1 gibt einen Überblick über die Probenentnahme sowie anschließende Analyse der mRNA-Expression und DNA-Methylierung. Letztere wurde mit Hilfe der bereits beschriebenen Methode MeDIP-Chip durchgeführt, welche in Kombination mit einem Chromosomen-basierten Array-Design (Chromosom 21, Teile der Chromosomen 20 und 22, siehe auch 1.6.1) die Erfassung von Methylierungsmustern verschiedener funktioneller Bereiche, wie Promotoren, intra- und intergenischer Regionen, erlaubte. Des Weiteren wurden MeDIP-Chip-Daten für Leukozyten sowie zuvor publizierte gewebespezifische Methylierungsdaten aus quantitativer Bisulfit-Sequenzierung für >260 Promotoren des Chromosoms 21 (Zhang et al. 2009) zur Überprüfung der Ergebnisse herangezogen. Die Analyse einzelner Zielregionen sowie repetitiver Elemente mittels Bisulfit-Sequenzierung diente ferner einer Bewertung der Resultate im genomweiten Kontext.

3.2.2 Immunpräzipitation und Qualitätskontrolle vor/nach der Amplifikation

Die Immunpräzipitation der HepG2 Proben *NKtr*, *Ktr*, *DMSO*, *f48h* und *f120h* wurde gemäß dem etablierten Protokoll in technischen Replikaten (Duplikate) durchgeführt. Dazu wurden jeweils 4 µg sonifizierte DNA zur Inkubation mit Primär- und Sekundärantikörper eingesetzt und die aufgereinigten Präzipitate, sowie die Input DNA, anschließend einer

Amplifikation zugeführt. Alle Proben wurden den Standardkontrollen vor bzw. nach Amplifikation unterzogen, um die erforderliche Qualität des Materials für eine erfolgreiche Hybridisierung zu gewährleisten (siehe auch 1.6.1). Dabei lagen die Ausbeute sowie die Verteilung der Größenfragmente innerhalb der Erfahrungswerte (Anhang B, Tabelle 6.2-1).

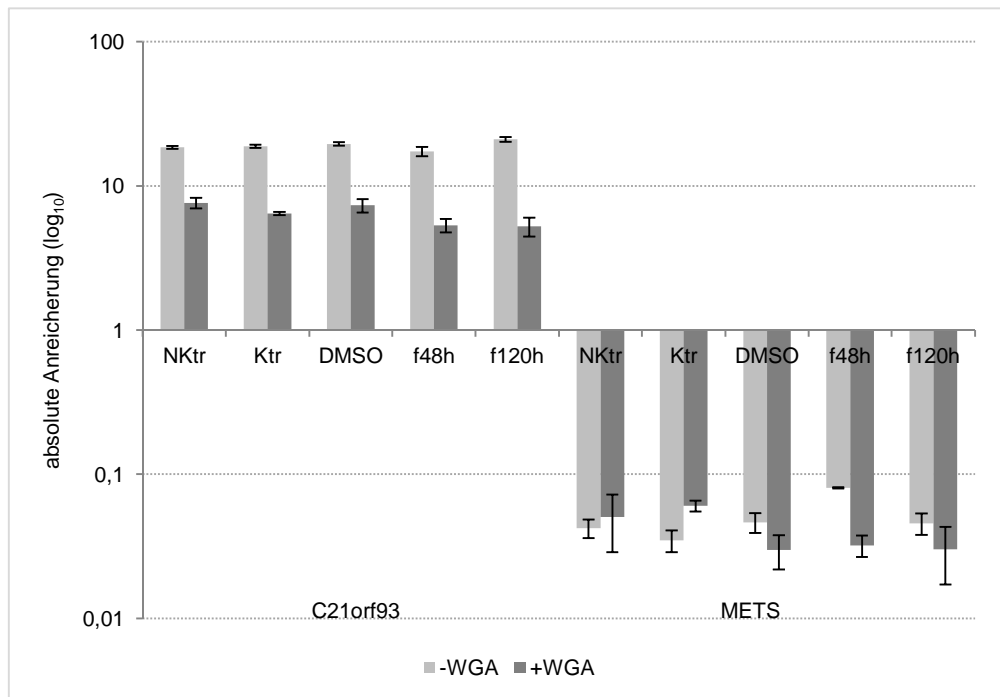


Abbildung 3.2-2: Überprüfung der HepG2-Präzipitate vor und nach WGA. Dargestellt ist die absolute Anreicherung (logarithmische Skalierung) je einer Positiv- und Negativkontrolle (C21orf93 bzw. METS) der HepG2-Präzipitate vor und nach genomrepräsentativer Amplifikation (-WGA bzw. +WGA).

Auch die mittels qPCR überprüften Kontrollregionen wiesen keine wesentlichen Unregelmäßigkeiten im Vergleich zu früheren Experimenten auf. Die methylierte Region C21orf93 zeigte dabei eine deutliche Anreicherung im Vergleich zum Hintergrund der getesteten Negativkontrolle, wobei keine beträchtlichen Unterschiede zwischen den Replikaten oder Proben zu verzeichnen waren (Abbildung 3.2-2).

Diese Ergebnisse wiesen auf eine ausreichende Qualität der Proben in Hinblick auf eine spätere Vergleichbarkeit der Chip-Daten hin, so dass die amplifizierten Präzipitate und Input-DNAs zur Hybridisierung eingesetzt werden konnten.

Die Immunopräzipitation und Qualitätskontrolle der in 3.2.1 erwähnten Leukozyten (MC) erfolgte durch die Diplomandin *Sabrina Pichler* gemäß dem etablierten Protokoll für limitiertes Probenmaterial (2 µg sonifizierte DNA zur Präzipitation, 20 ng für WGA). Diese Proben werden erst im nächsten Abschnitt näher berücksichtigt.

3.2.3 Analyse chromosomenweiter DNA-Methylierungsmuster nach Kryokonservierung mit MeDIP-Chip

Amplifizierte IP- und Input-DNAs der HepG2-Proben *NKtr*, *Ktr*, *DMSO*, *f48h* und *f120h* wurden zur weiteren Prozessierung und Hybridisierung an den Dienstleister übergeben (ImaGenes). Unter Verwendung der in Abschnitt 3.1.2.1 erläuterten Standardprotokolle wurden Qualitätskontrollen und Prä-Prozessierungen der Rohdaten vorgenommen, wobei technische Variabilitäten mittels Median-Zentrierung und Quantil-Normalisierung korrigiert wurden. Zur Vermeidung der in Abschnitt 3.1.2.3.2 dargestellten möglichen Verzerrungen in Abhängigkeit der Array-Prozessierung, wurde *NKtr* erneut parallel mit den später bearbeiteten Leukozyten-Proben hybridisiert. Diese Daten wurden getrennt von den übrigen HepG2-Proben analysiert und dienten lediglich als Positivkontrolle der experimentellen und bioinformatischen Abläufe.

3.2.3.1 Korrelation und Clustering

Zunächst wurde die technische Variabilität des MeDIP-Chips unter Berücksichtigung der jeweiligen Replikat pro Probe untersucht. In einem paarweisen Vergleich aller Datenpunkte konnte dabei eine hohe Reproduzierbarkeit der Duplikate festgestellt werden (Pearson-Korrelations-Koeffizienten: Leukozyten $r=0,92$; HepG2 $r=0,94-0,98$). Während die Proben im gewebespezifischen Vergleich eine verringerte Korrelation gegenüber den Replikaten aufwiesen (Leukozyten vs. HepG2 *NKtr*, $r=0,80$), führte die hohe Ähnlichkeit der Anreicherungswerte für HepG2 Proben *NKtr*, *Ktr*, *DMSO*, *f48h* und *f120h* untereinander ($r=0,96-0,98$) zu keiner klaren Trennung derselben in einer hierarchischen Clusteranalyse (Abbildung 3.2-3 (A) und (B)). Diese ersten Daten deuten bereits darauf hin, dass im Falle des Kryopreservations-Experiments - im Gegensatz zum gewebespezifischen Vergleich - eher geringe Veränderungen der Methylierung im Bereich technischer Variabilität zu erwarten sind.

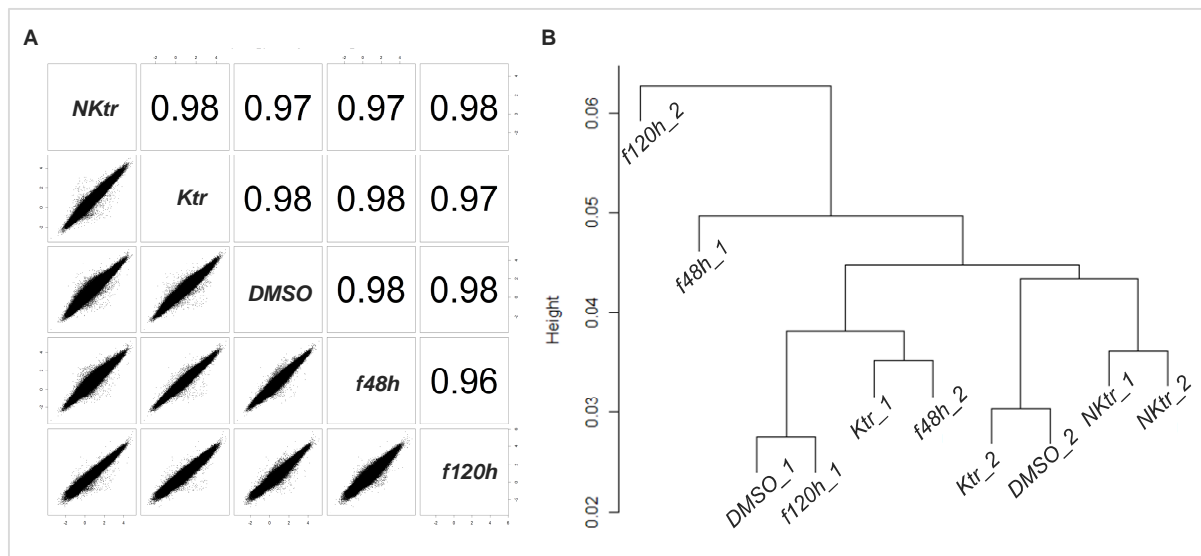


Abbildung 3.2-3: Vergleich der HepG2-Proben untereinander auf Basis der Oligo-log₂-Ratios. (A) Korrelationsmatrix mit Pearson-Korrelations-Koeffizienten, basierend auf den log₂-Ratios der gemittelten technischen Replikate. **(B)** Hierarchische Clusteranalyse führt nicht zur Diskriminierung der Proben anhand ihrer Replikate.

3.2.3.2 Differentielle Anreicherung behandelter HepG2 Proben

In einem zweiten Schritt sollten die Anreicherungsprofile der unbehandelten und behandelten Proben verglichen werden. Die bereits bei der Korrelation festgestellte Ähnlichkeit dieser Proben wird bei einfacher Betrachtung verschiedener chromosomaler Abschnitte, wie in Abbildung 3.2-4, verdeutlicht, welche nahezu identische Profile aufweisen. Der gewebespezifische Vergleich hingegen lässt wie zu erwarten verschiedene Bereiche differentieller Anreicherung zwischen Leukozyten und HepG2 Zellen erkennen (Abbildung 3.1-11, S.64).

Mit Hilfe des in Abschnitt 3.1.2.2 vorgestellten Ansatzes eines gleitenden Fensters sollte eine detaillierte chromosomenweite Analyse differentieller Methylierung zwischen den Kontrollen (*Ktr_av*: *NKtr* und *Ktr* gemittelt) und behandelten Proben (*DMSO*, *f48h*, *f120h*) erfolgen. Dabei wurden, basierend auf den delta-log₂-Werten der zu vergleichenden Proben, benachbarte Oligonukleotide mit signifikanter differentieller Anreicherung zu „Peaks“ (DP) zusammengefasst (siehe auch Abschnitt 3.1.2.2.2).

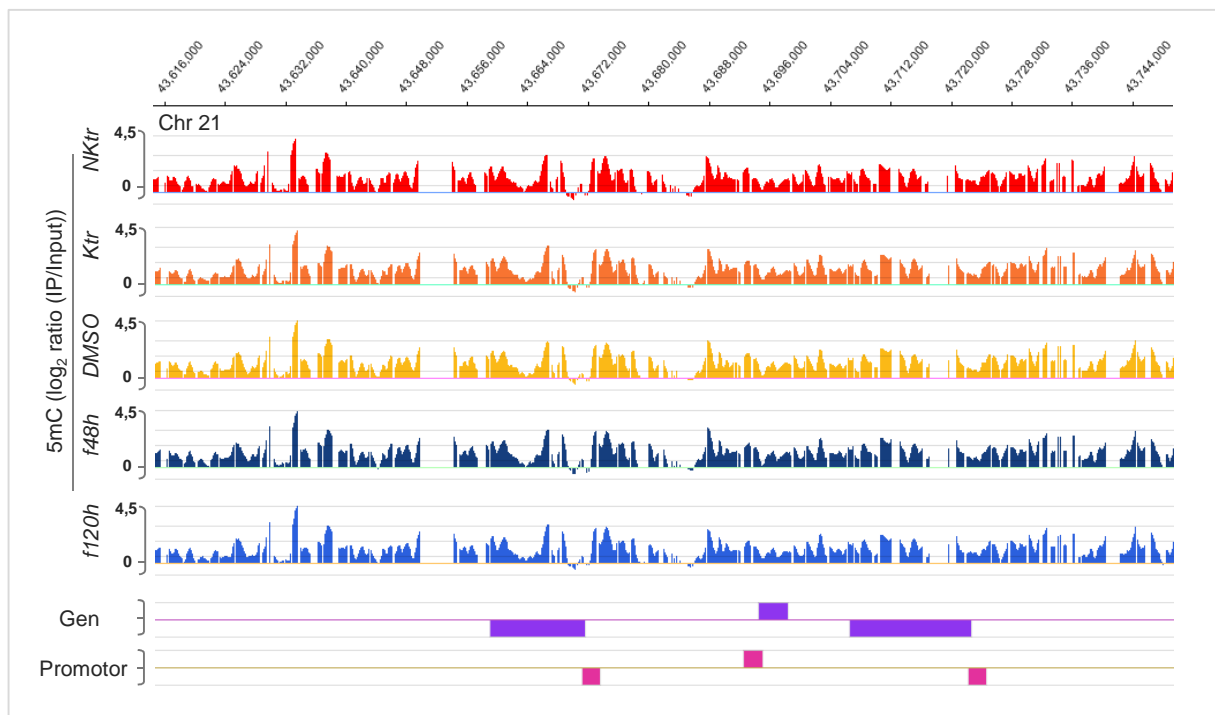


Abbildung 3.2-4: Anreicherungsprofile der HepG2-Proben. Dargestellt sind Oligo- \log_2 -Ratios (geglättet) der DMSO-behandelten, kryokonservierten Proben im Vergleich zu unbehandelten Kontrollen für einen Ausschnitt des Chromosoms 21, relativ zu den genomischen Koordinaten sowie Genen und Promotoren.

Im Gegensatz zum gewebespezifischen Vergleich konnten zwischen den HepG2 Proben selbst keine differentiell angereicherten Regionen unter der erwarteten Wahrscheinlichkeit falsch-positiver Resultate („*false discovery rate*“, $FDR < 0.05$) identifiziert werden. Lediglich bei Anwendung weniger strenger Kriterien (keine Korrektur des p-Werts für multiples Testen, $p < 0.01$) war, abhängig von der Behandlung, eine Bestimmung von 39-61 DPs möglich (*Ktr_av* vs. *DMSO*: 61 DPs, *Ktr_av* vs. *f48h*: 56 DPs, *Ktr_av* vs. *f120h*: 39 DPs). Diese Regionen wiesen ausschließlich geringe, zwischen 1,4- und 2-fache, Anreicherungsunterschiede auf und umspannten meist nicht mehr als 200 nt, was bis zu zwei benachbarten Oligonukleotiden entspricht. Damit hoben sie sich deutlich von den gewebespezifischen DPs ab, welche hauptsächlich größere Abschnitte von 500-1000 nt umfassten und in etwa der Hälfte der Fälle 2-8-fache Anreicherungsunterschiede aufwiesen. Jedoch lagen sie hinsichtlich der genannten Eigenschaften im Rahmen der (zwischen den Kontrollen *NKtr* und *Ktr* beobachteten) biologischen Variabilität. Allerdings zeigten die DPs der jeweiligen behandelten Proben eine tendenzielle Hypermethylierung, während zwischen den Kontrollen selbst eher ein Gleichgewicht zwischen Hypo- und Hypermethylierung zu verzeichnen war. Abbildung 3.2-5 veranschaulicht diese Ergebnisse

in Form von Heatmaps, welche die differentielle Anreicherung der jeweiligen DPs farblich darstellt.

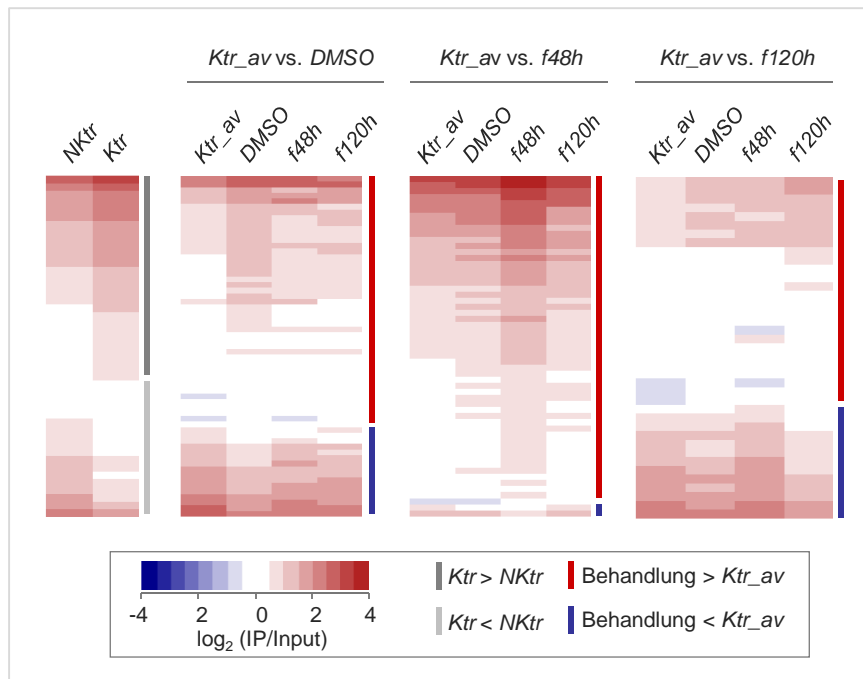


Abbildung 3.2-5: Anreicherungsunterschiede nachfolgend Kryopreservation von HepG2-Zellen. Heatmaps der mittleren \log_2 -Anreicherung über DPs zwischen Kontrollen (*Ktr_av*, gemittelte *NKtr* und *Ktr*) und den jeweiligen Behandlungen deuten auf eine leichte Hypermethylierung nach Kryopreservation hin.

Weitere Analysen der identifizierten Regionen hatten zum Ziel, diese im Verhältnis der Proben zueinander spezifischen Effekten oder Charakteristika zuordnen zu können. Dabei war von Interesse, ob eine differentielle Anreicherung die gleichen genomischen Regionen oder Gene/Promotoren in den behandelten Proben betraf und somit auf systematische Veränderungen hinweisen könnte. Unter Berücksichtigung von jeweils 1 kb-großen Abschnitten um die DPs selbst bzw. die entsprechend betroffenen Gene konnten keine signifikanten Überschneidungen zwischen den verschiedenen Behandlungen festgestellt werden (Abbildung 3.2-6). Ferner war bei systematischer Analyse der mit DPs überlappenden Gene (Anhang B, Tabelle 6.2-2 bis Tabelle 6.2-4) keine Anreicherung spezifischer biologischer Prozesse (GO-Annotation) zu verzeichnen. Diese Ergebnisse deuten somit auf stochastische Ereignisse hinsichtlich einer Veränderung der DNA-Methylierung in Zusammenhang mit Kryokonservierung hin.

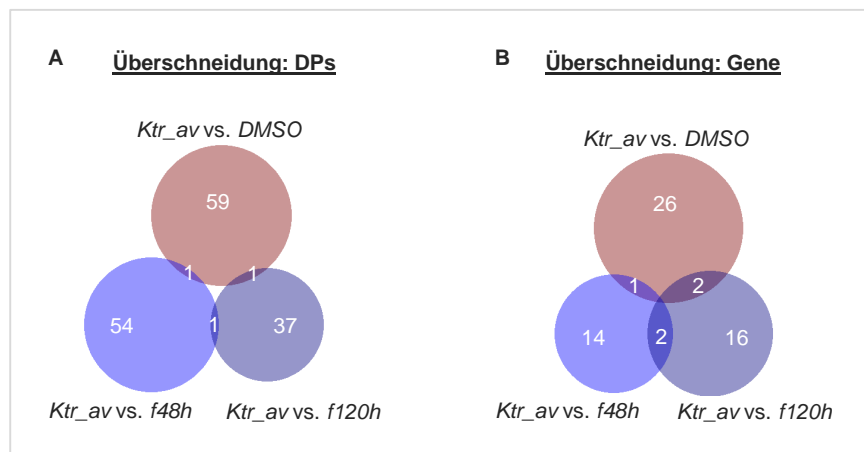


Abbildung 3.2-6: Unsystematische Anreicherungsunterschiede nachfolgend Kryopreservation von HepG2-Zellen. Venn-Diagramme weisen wenig Überschneidung der DPs (A) oder betroffenen Gene/Promotoren (B) (jeweils +/-1 kb) zwischen den verschiedenen Behandlungen auf.

Des Weiteren wurden die DPs mit Hilfe der *EPIGRAPH*-Anwendung in Bezug auf verschiedene Sequenzeigenschaften untersucht. Obwohl die geringe Anzahl zu testender Regionen fundierte, statistisch abgesicherte Aussagen beeinträchtigte, konnten zumindest gewisse Tendenzen herausgearbeitet werden. Dabei schienen DPs im Vergleich zu Kontrollregionen eher in intergenischen Regionen als Transkripten oder regulatorischen Regionen zu liegen. Zudem waren tendenziell eine geringere G- und CpG-Dichte sowie eine Überschneidung mit eher schwächeren CpG-Inseln zu verzeichnen.

3.2.4 Genomweite mRNA-Expression in Zusammenhang mit Kryopreservation

Da in der Literatur bereits mehrfach eine Veränderung der Genexpression durch abiotischen Stress wie Hypothermie berichtet wurde und DNA-Methylierung einen wichtigen Bestandteil der Genregulation u.a. auf Transkriptionsebene darstellt, sollten zusätzlich genomweite mRNA-Expressionslevel in Verbindung mit Kryokonservierung analysiert werden. Die Bearbeitung der entsprechenden RNA-Proben *Ktr*, *DMSO*, *f48h* und *f120h* einschließlich Hybridisierung auf *HT12 beadchips* (Illumina) und statistischer Analyse (ANOVA) erfolgte dabei durch einen Dienstleister.

Insgesamt konnten für 530 der 37.880 auf dem Array abgefragten Gene bzw. transkribierten Abschnitte (RefSeq und UniGene) statistisch signifikante, über zwei-fache Änderungen der Expression in jeglichen behandelten Proben festgestellt werden. Um eine mögliche Verbindung der Genexpression mit DNA-Methylierung herstellen zu können, wurden diese Ergebnisse, zunächst unabhängig des beobachteten Effekts bzgl. der

Behandlung und damit des potentiellen Stressors (DMSO und/oder Einfrieren/Auftauen), mit den korrespondierenden MeDIP-Chip-Daten verglichen.

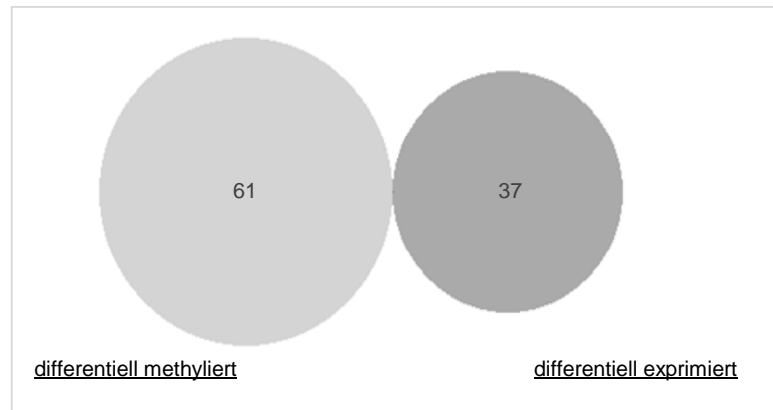


Abbildung 3.2-7: Änderungen der DNA-Methylierung bzw. mRNA-Expression nachfolgend Kryopreservation betreffen unterschiedliche Gene. Venn-Diagramm zeigt keine Überschneidung der mittels MeDIP-Chip ermittelten Gene differentieller Anreicherung mit differentiell exprimierten Genen der Chromosomen 20, 21 und 22.

Keines der in der Nähe eines DPs lokalisierten Gene wies dabei signifikante Veränderungen der mRNA-Expressionslevel auf und umgekehrt (Abbildung 3.2-7). Somit konnte für die exemplarische Analyse von Genen der Chromosomen 20-22 zu den untersuchten Zeitpunkten kein direkter Zusammenhang zwischen DNA-Methylierung und Expression hinsichtlich Kryopreservation festgestellt werden.

In Erwägung eines indirekten Zusammenhangs durch die veränderte Expression von Schlüsselkomponenten zur Aufrechterhaltung genomweiter DNA-Methylierungsmuster (Iwatani), wurden zusätzlich die DNA-Methyltransferasen (DNMT1/3a/3b/3L) genauer betrachtet. Diese wiesen allerdings keine signifikanten Abweichungen der Expressionslevel auf.

Da keine offensichtliche Verbindung zwischen DNA-Methylierung und Expression ersichtlich war, wurden die differentiell exprimierten Gene in Abhängigkeit des beobachteten Effekts über die Proben systematisch gruppiert, um somit einen näheren Einblick zu erhalten (Abbildung 3.2-8).

Bei 21% der betroffenen Gene (112 Gene, 33 hochreguliert, 79 herunter reguliert) war nach 48-stündiger Kultivierung in DMSO-behandelten und kryokonservierten Proben (*f48h*) ein ähnliches Muster zu verzeichnen, was auf einen DMSO-induzierten Einfluss unabhängig des Einfrierens oder Auftauens hinweist. Die Mehrheit dieser Gene wies dabei lediglich eine

transiente Herunter-Regulierung der Expression auf, da der spätere Kultivierungszeitpunkt (*f120h*) mRNA-Level nahe der Kontroll-Probe erreichte.

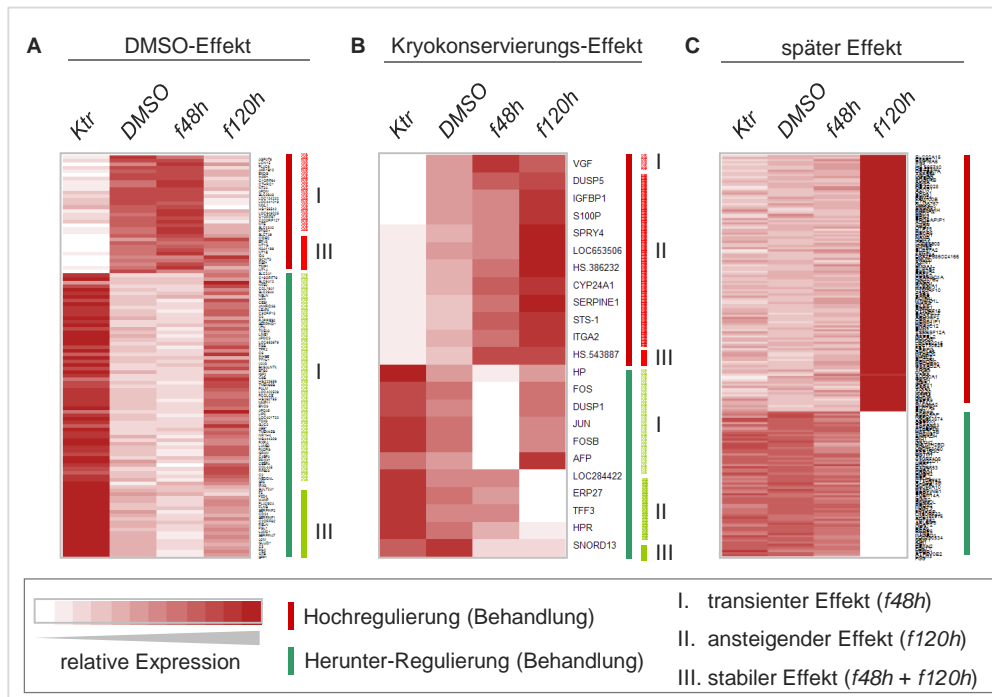


Abbildung 3.2-8: Änderungen der Expression nach DMSO-Behandlung bzw. Kryopreservation. Gene mit statistisch signifikanten Expressionänderungen zwischen Kontrolle und behandelten Proben wurden gemäß dem über die Proben beobachteten Effekt gruppiert. Dabei sind unterschiedliche Tendenzen auffallend: I. transiente Effekte mit sichtbaren Expressionänderungen nach 48-stündiger Kultivierung und normalen Leveln nach 120 h; II. aufbauende Effekte mit größter Expressionsänderung nach 120-stündiger Kultivierung; III. verbleibende Effekte während 48- und 120-stündiger Kultivierung. Heatmaps zeigen die relative Expression über die Proben, wobei rot die Proben mit den höchsten und weiß die Proben mit den niedrigsten Expressionsleveln pro Gen bezeichnen.

Eine zweite kleinere Gruppe von Genen (4%, 23 Gene hochreguliert, 12 Gene herunter reguliert) hingegen ließ zusätzlich auf potentielle Auswirkungen der Hypothermie bzw. des Erwärmens schließen, da die kryokonservierten Proben (*f48h* und *f120h*) sich in ihren Expressionsmustern deutlich von der Kontrolle als auch den DMSO-behandelten Zellen unterschieden. Während für hochregulierte Gene vorwiegend ein ansteigender Effekt mit höchstem Expressions-Level nach 120-stündiger Kultivierung beobachtet werden konnte, wiesen herunter regulierte Gene auch transiente Effekte auf. Ein drittes Gen-Cluster, dessen Expressions-Änderungen erst in der späten Kultivierungs-Phase nachfolgend Kryopreservation (*f120h*) auftraten (36 %, 191 Gene, 122 hochreguliert, 69 herunter reguliert), spiegelt vermutlich Effekte in Zusammenhang mit der Kultivierung selbst bzw. Fortschreiten des Zell-Zyklus oder aber eine verspätete Auswirkung des abiotischen

Stressors wider. Weitere 204 Gene wiesen lediglich schwache (2-fache) Änderungen der Expression auf und ließen keine eindeutige Abgrenzung der Stressoren zu (Daten nicht gezeigt).

Unter Verwendung der Web-basierten *DAVID*-Software (Huang da et al. 2009) wurde eine funktionelle Analyse mittels Gen-Ontologie durchgeführt, wobei Gene der oben beschriebenen Cluster auf überrepräsentierte biologische Prozesse hin untersucht wurden. Die Gesamtheit der auf dem Microarray analysierten Gene wurde dazu als Referenz genutzt. Um die Redundanz der GO-Annotationen zu reduzieren und dadurch die Analyse zu fokussieren, wurden Kategorien ähnlicher biologischer Bedeutung in Gruppen zusammengefasst. Diese wurden anhand von Oberbegriffen in Abbildung 3.2-9 übersichtlich dargestellt. Die vollständigen Ergebnisse dieser Analyse, sowie die entsprechenden Gene der identifizierten Gruppen, befinden sich im Anhang (Anhang B, Tabelle 6.2-11 bis Tabelle 6.2-14 sowie Tabelle 6.2-5 bis Tabelle 6.2-10).

Gene, welche in der späten Kultivierungsphase Änderungen der Expression aufwiesen, waren hinsichtlich verschiedener „Haushaltsprozesse“ wie Aminosäure-Metabolismus, -Transport und -Aktivierung (vermehrte Expression) bzw. Lipid-Metabolismus (verminderte Expression) angereichert (Abbildung 3.2-9 (C) und (D)). Im Gegensatz dazu konnten die herunterregulierten Gene der DMSO- und Kryopreservations-Kategorien insbesondere verschiedenen Stress-Reaktionen zugeordnet werden (Abbildung 3.2-9 (A) und (B)). Hierzu zählten interessanterweise nicht jene Stress-induzierten Gene, welche bereits in der Literatur in Zusammenhang mit Hypothermie genannt wurden, wie Chaperone (HSP), Zell-Zyklus-Regulatoren p53 und p21 oder Kälte-induzierte RNA-binde Proteine CIRP und RBM3.

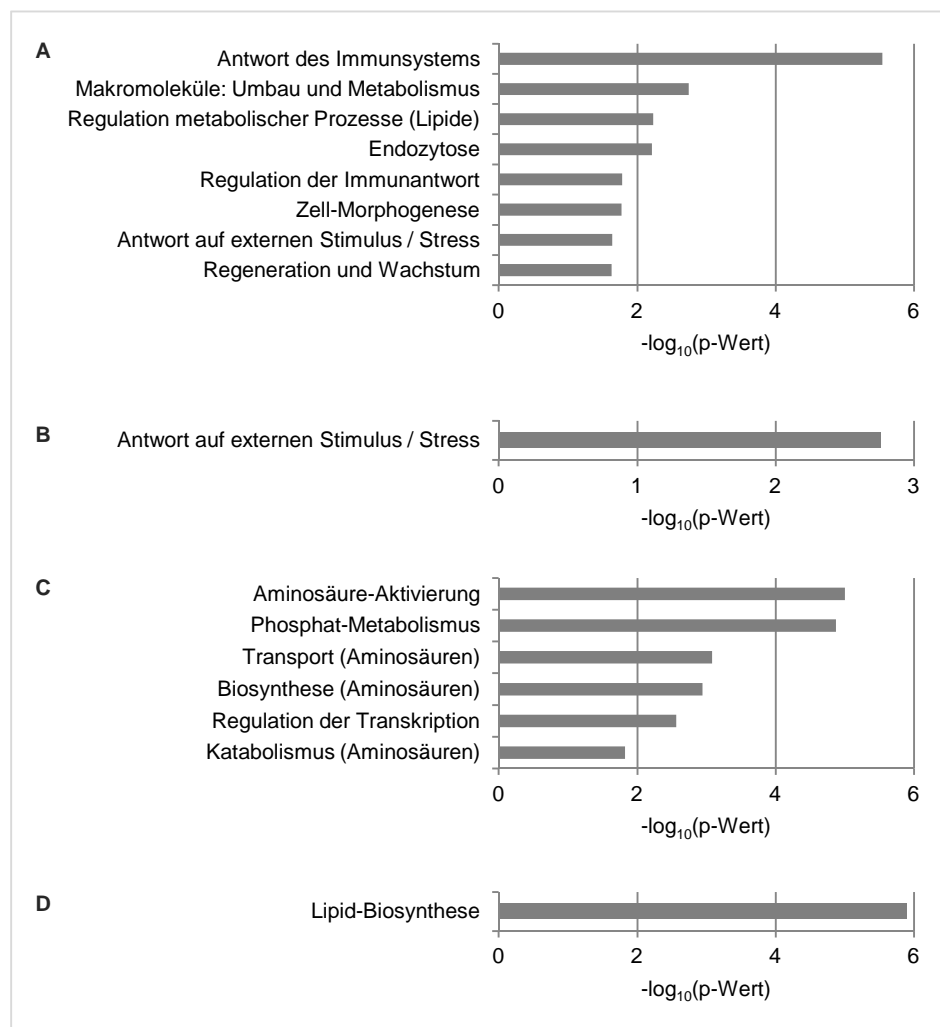


Abbildung 3.2-9: Überrepräsentierte biologische Prozesse (Gen-Ontologie) differenziell exprimierter Gene nach Kryopreservation. Gene mit signifikant deregulierter Expression in *DMSO*, *f48h* bzw. *f120h* wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im Array abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln, für (A) Herunterregulierte Gene mit *DMSO*-Effekt; (B) Herunterregulierte Gene mit Kryokonservierungs-Effekt; (C) Hochregulierte Gene mit spätem Effekt und (D) Herunterregulierte Gene mit spätem Effekt. Dargestellt sind die Oberbegriffe der in Gruppen zusammengefassten Prozesse ähnlicher biologischer Bedeutung.

3.2.5 Quantitative Methylierungsanalysen ausgesuchter Zielregionen

Zur weiteren Charakterisierung von DNA-Methylierungsprofilen in Zusammenhang mit Kryopreservation wurden quantitative Analysen einzelner Zielregionen mittels Sequenzierung Bisulfit-behandelter DNA vorgenommen, wobei einzelne CpG-Positionen im Sequenzkontext der Region von Interesse betrachtet werden können.

3.2.5.1 Validierung einzelner DPs der MeDIP-Chip-Analyse

Zur Überprüfung der MeDIP-Chip-Ergebnisse und Erfassung quantitativer Methylierungsunterschiede zwischen den entsprechenden Proben wurden einzelne DPs (jeweils ein DP pro behandelter Probe *DMSO*, *f48h*, *f120h*) mittels Bisulfit-Sequenzierung analysiert. Es wurden dazu DPSs gewählt, welche in unterschiedlichen genomischen Regionen, wie dem 5' bzw. 3'-Abschnitt eines Gens (DP Ktr_av vs. *f48h* bzw. DP Ktr_av vs. *DMSO*) sowie einem intergenischen Bereich (DP Ktr_av vs. *f120h*) liegen.

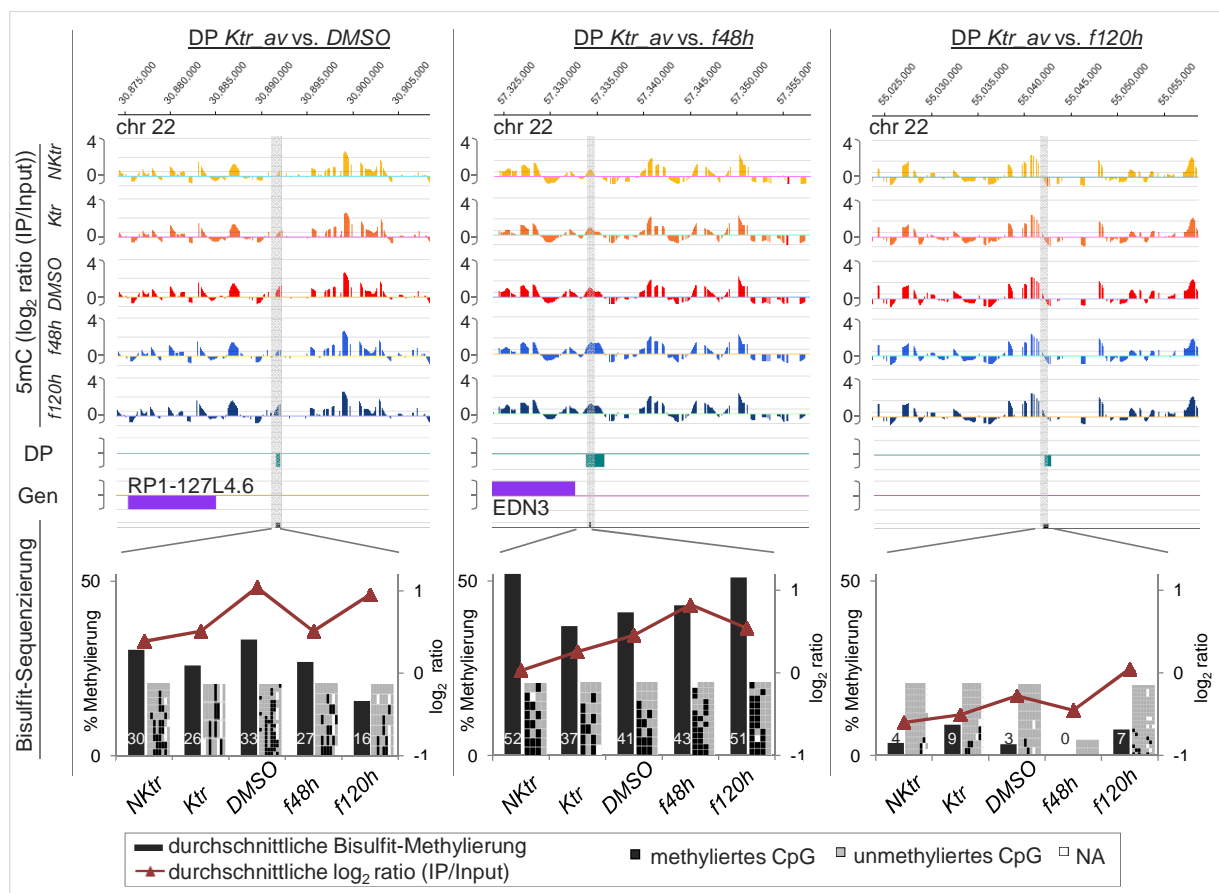


Abbildung 3.2-10: Anreicherungsprofile ausgewählter DPs im Vergleich zu quantitativen Methylierungsdaten. Dargestellt sind \log_2 -ratios (geglättet), DPs und Amplikons der Bisulfit-Sequenzierung relativ zu genomischen Koordinaten und Genen (Ensembl). Mittlere \log_2 -Ratios der DPs (sekundäre Y-Achse) und mittlere Methylierungslevel aus Bisulfit-Sequenzierung (Y-Achse) sind darunter zusammen mit den Methylierungsmustern gezeigt, wobei individuelle Sequenzen in der Vertikalen und einzelne CpG-Positionen in der Horizontalen abgebildet sind.

In allen drei analysierten Regionen konnten dabei nur geringe Abweichungen der Methylierung (<15%) zwischen Kontrollen und jeglichen behandelten Proben festgestellt werden. Dies lag im Bereich biologischer Variabilität, welche zwischen den Kontrollen *NKtr*

und *Ktr* selbst zu verzeichnen war. Angedeutete Tendenzen bzgl. Hypo- oder Hypermethylierung der entsprechenden Proben reflektierten zudem nicht die im MeDIP-Chip beobachteten Muster. Abbildung 3.2-10 veranschaulicht die Resultate im Vergleich zu Anreicherungsprofilen der MeDIP-Chip-Analyse.

3.2.5.2 Analyse differentiell exprimierter Kandidatengene

Um einen möglichen Zusammenhang zwischen DNA-Methylierung und mRNA-Expression in Verbindung mit abiotischem Stress näher zu untersuchen, wurden vier differentiell exprimierte Gene mittels Bisulfit-Hochdurchsatz-Sequenzierung analysiert. Dabei wurden, basierend auf Hinweisen aus der Literatur bezüglich einer möglichen DNA-Methylierungs-abhängigen Regulation der Expression, repräsentativ für die zuvor identifizierten Effektgruppen (Abschnitt 3.2.4) folgende Gene gewählt: i) CEBPA (Lu et al. 2010), transient verminderte Expression / DMSO-Effekt; ii) FOS (Choi et al. 1996), transient verminderte Expression / Kryokonservierungs-Effekt; iii) IGFBP1 (Ibanez de Caceres et al. 2006), erhöhte Expression / Kryokonservierungs-Effekt und iv) SPRY4 (Wang et al. 2006), erhöhte Expression / Kryokonservierungs-Effekt. In keinem der analysierten Amplikons konnten Methylierungsdifferenzen über 5,5% zwischen jeglichen Proben festgestellt werden (Anhang B, Abbildung 6.2-1).

3.2.5.3 Analyse repetitiver Elemente

Kälteschock-Experimente an Pflanzen deuten eine spezifische Demethylierung repetitiver Elemente an (Steward et al. 2002, Steward et al. 2000). Um diesem Hinweis nachzugehen und gleichzeitig potentielle Methylierungsänderungen infolge Kryopreservation auf genomweiter Ebene zu erfassen, wurde eine exemplarische Analyse dieser im MeDIP-Chip nicht berücksichtigten Regionen mittels Bisulfit-Hochdurchsatz-Sequenzierung vorgenommen. Dazu wurde die retrotransposable LINE-Familie („*long interspersed nuclear elements*“) sowie der endogene Retrovirus HERVK gewählt, welche etwa 20% bzw. 8% des menschlichen Genoms ausmachen. Die Sequenzdaten wiesen dabei Methylierungsdifferenzen von bis zu 6% zwischen jeglichen Proben auf (Abbildung 3.2-11).

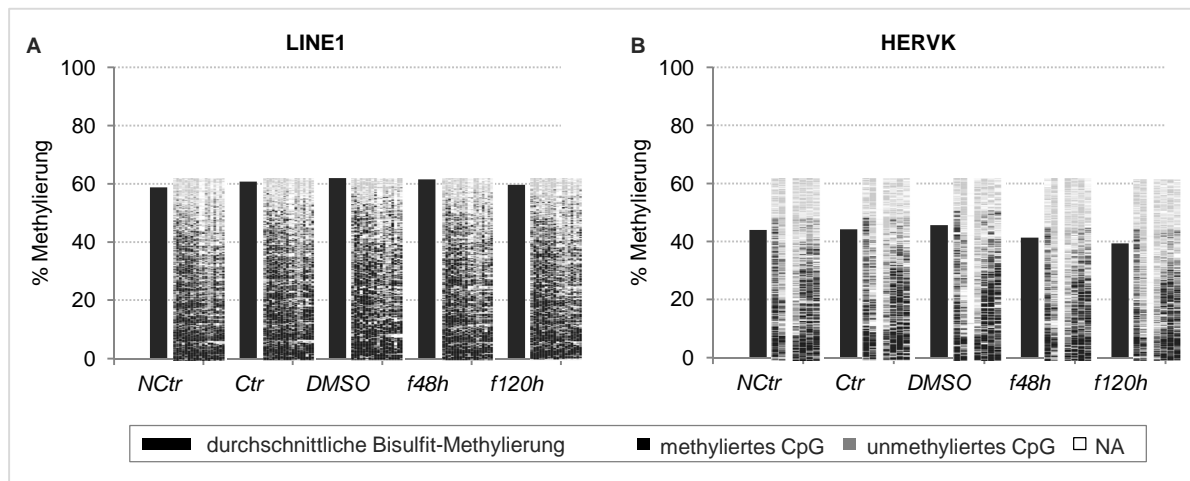


Abbildung 3.2-11: DNA-Methylierung repetitiver Elemente. Dargestellt sind die mittleren Methylierungslevel aus Bisulfit-Hochdurchsatz-Sequenzierung, sowie die jeweiligen Methylierungsmuster der untersuchten Amplikons mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen für (A) LINE1 und (B) HERVK.

3.3 Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen

Bei Stammzellen handelt es sich um Zellen, welche sich in Abhängigkeit ihres Potentials in festgelegte Gewebetypen oder sogar jegliches Gewebe ausdifferenzieren können. Während pluripotente Zellen der inneren Zellmasse in der Lage sind, alle drei Keimblätter des sich entwickelnden Organismus zu bilden, besitzen post-embryonale bzw. adulte Stammzellen ein eingeschränktes, determiniertes Entwicklungspotential (Hochedlinger und Plath 2009). Verschiedene in den vergangenen Jahren generierte Stammzelllinien ermöglichen eine Erforschung des Differenzierungspotentials verschiedener Entwicklungsstadien und erweitern damit das Grundverständnis entwicklungspezifischer Vorgänge.

In der vorliegenden Arbeit sollten verschiedene murine Stammzelltypen bzgl. ihrer Methylierungsprofile charakterisiert werden. Die hierzu verwendeten Zelllinien repräsentierten Zelltypen unterschiedlichen ontogenetischen Alters bzw. Differenzierungspotentials. Neben ausdifferenzierten Fibroblasten (MEF: „*mouse embryonic fibroblast*“) und unipotenten Keimbahnstammzellen (GSC: „*germline stem cell*“) waren verschiedene pluripotente Zelltypen von Interesse. Dazu zählten neben embryonalen Stammzellen selbst (ESC: „*embryonic stem cell*“) auch durch Reprogrammierung der GSC erhaltene ESC-ähnliche Zellen (gPS: „*germline pluripotent cell*“) sowie Stammzellen des Epiblasten (EpiSC: „*Epiblast stem cell*“), welche ein

späteres, „vorgeplantes“ Pluripotenzstadium repräsentieren. Durch eine vergleichende Analyse der Promotor-Methylierung mittels MeDIP-Chip sollte ein besseres Verständnis der zugrundeliegenden Entwicklungsstadien erreicht werden.

3.3.1 Immunpräzipitation und Qualitätskontrolle

Zur Untersuchung der DNA-Methylierung wurden Zellpellets in biologischen Replikaten zur Verfügung gestellt (Kooperationspartner MPI für molekulare Biomedizin, Münster).

Aufgrund der in Abschnitt 3.1.1.4 dargestellten WGA-Problematik sollte bei dieser Versuchsreihe auf eine spätere genomrepräsentative Amplifikation verzichtet werden. Aus diesem Grund wurden pro Replikat zwei Immunpräzipitationen durchgeführt, um durch Vereinigung der IPs genügend Material für die Microarray-Prozessierung zu erhalten.

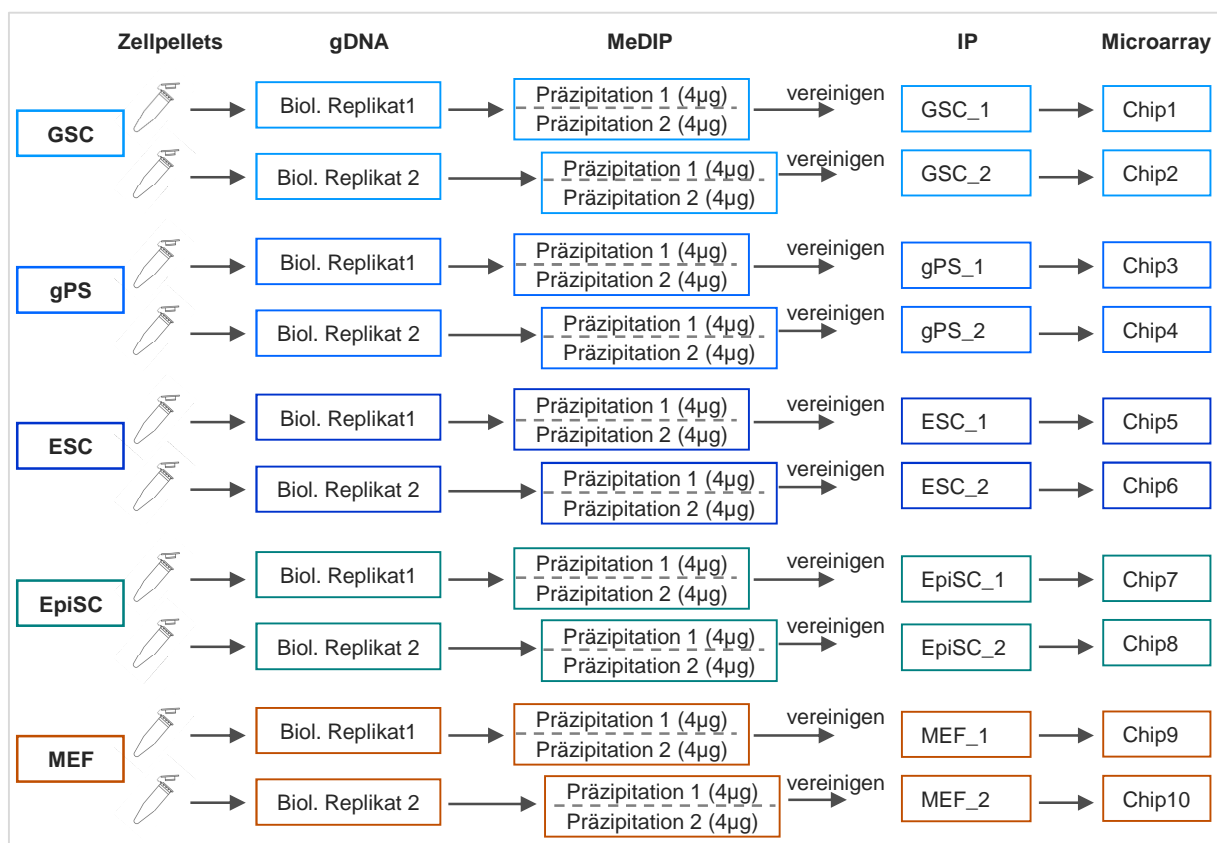


Abbildung 3.3-1: Schematische Darstellung zur Bearbeitung der murinen Proben. Zellpellets der Zelltypen GSC („*germ line stem cells*“), gPS („*germ line pluripotent cells*“), ESC („*embryonic stem cells*“), EpiSC („*epiblast stem cells*“) und MEF („*mouse embryonic fibroblasts*“) wurden in biologischen Replikaten mit je zwei Immunpräzipitationen bearbeitet, welche zwecks Microarray-Hybridisierung vereinigt wurden.

Die Experimente wurden gemäß Standardprotokoll ausgehend von je 4 µg nebulisierter DNA durchgeführt (Abbildung 3.3-1), wobei die biologischen Replikate zeitlich versetzt bearbeitet wurden. Die Präzipitate wurden anschließend im Fluorometer gemessen und mittels qPCR überprüft.

Nach Vereinigung der technischen Replikate wurden DNA-Mengen von 373-645 ng gemessen, was einer Präzipitationsrate von 4,7 - 8,1 % entspricht. Dabei waren teilweise erhebliche Schwankungen zwischen den verschiedenen Zelllinien, nicht jedoch zwischen den biologischen Replikaten - mit Ausnahme der MEFs - zu verzeichnen (Abbildung 3.3-2 (A)). Mit Hilfe der qPCR wurden zusätzlich die spezifische und unspezifische Präzipitation und damit die Qualität der IPs überprüft (Abbildung 3.3-2 (B)).

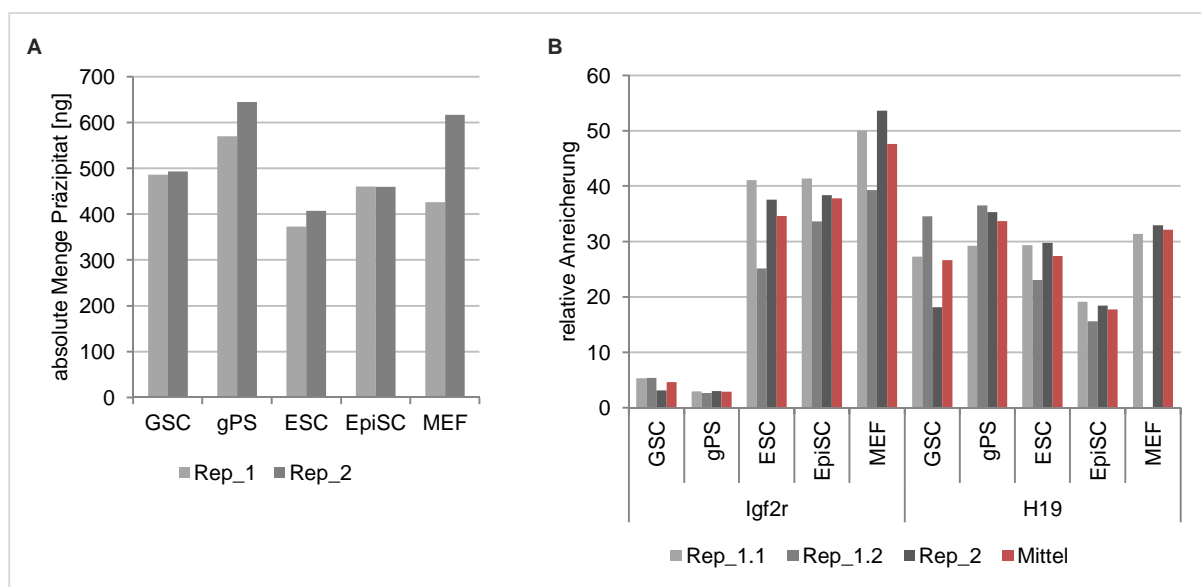


Abbildung 3.3-2: Überprüfung der Immunpräzipitation muriner Proben. (A) Quantifizierung absoluter Präzipitattmengen der jeweiligen Zelltypen (Fluorometer) für die biologischen Replikate 1 und 2 (Rep_1, Rep_2). (B) Die relative Anreicherung der Kontrollregionen Igf2r und H19 im Vergleich zum Hintergrund (CpG-frei) wurde mittels qPCR für die biologischen Replikate bzw. technische Replikate bestimmt. Rep_1.1, Rep_1.2: technische Replikate des biologischen Replikats 1; Rep_2: biologisches Replikat 2; Mittel: mittlere Anreicherung pro Zelltyp errechnet aus den Replikatanreicherungen.

Wie erwartet konnte dabei eine deutliche Anreicherung der Positivkontrollen Igf2r und H19 gegenüber der abgereicherten CpG-freien Region (Hintergrund) festgestellt werden. Die zu beobachtenden Schwankungen zwischen technischen und biologischen Replikaten lagen hierbei im Erfahrungsbereich. Die aus der Literatur bekannten androgenetischen Methylierungsmuster der keimbahnassoziierten Zellen GSC und gPS (Ko et al. 2009) konnten im Falle des maternal reprimierten Gens Igf2r aufgrund der unwesentlichen Anreicherung deutlich, im Falle des paternal reprimierten Gens H19 z.T. aufgrund technischer

Schwankungen allerdings nur unzureichend erkannt werden. Insgesamt wiesen die Proben eine ausreichende Qualität für weitere Analysen auf, so dass jeweils ~300 ng IP und Input-DNA zur Microarray-Prozessierung an den Dienstleister übergeben wurden.

3.3.2 Analyse der DNA-Promotormethylierung mittels MeDIP-Chip - Gruppierung der Zelltypen zueinander

Zur Erstellung genomweiter Methylierungsprofile wurden Promotor-Arrays der Firma NimbleGen genutzt, welche durchschnittlich einen Bereich von 2 kb um die jeweiligen Transkriptionsstartstellen (TSS) abdecken. Um Methylierungsunterschiede zwischen TSS-proximalen und aufwärts liegenden Regionen erfassen zu können, wurde eine Unterteilung in so genannte Kern- und Aufwärts-Promotoren vorgenommen (+/-500 bp um TSS bzw. -500 bis -1000 bp). Auf Basis der prä-prozessierten Daten wurden mit Hilfe der *Prometheus*-Software (de Las Heras et al. in Vorbereitung) die \log_2 -Werte aller Oligonukleotide dieser definierten Bereiche unter Berechnung des Median zusammengefasst.

Zunächst sollte eine generelle Einordnung der Daten durch paarweisen Vergleich der Zelltypen, basierend auf den Median-Werten aller Promotor-Regionen (mit Ausnahme der Geschlechts-Chromosomen), erfolgen. Dabei ergaben sich hohe Pearson-Korrelationskoeffizienten von $r > 0,87$ sowohl zwischen biologischen Replikaten ($r = 0,91-0,97$) als auch zwischen den verschiedenen Proben ($r = 0,87-0,95$). Mit Hilfe der hierarchischen Clusteranalyse konnten entsprechende Beziehungen zwischen den Zelltypen durch Gruppierung derselben erkannt werden (Abbildung 3.3-3 (A)). Dabei ergaben sich zwei interessante Cluster. Zum Einen gruppieren die keimbahnassoziierten Stammzellen GSC und gPS zusammen mit den ESC. EpiSC bildeten mit MEF ein zweites, davon klar abgegrenztes Cluster.

Bei getrennter Betrachtung der Kern- und Aufwärts-Promotoren wurde anhand der niedrigen Korrelationskoeffizienten zwischen diesen Regionen ($r < 0,5$) deutlich, dass deren Methylierungsmuster keinen Zusammenhang aufwiesen (Abbildung 3.3-3 (B)). Dabei konnten keine systematischen Unterschiede zwischen den Promotorklassen hinsichtlich Zelltyp oder Differenzierungspotential erkannt werden. Tatsächlich ließ sich sowohl für Kern- als auch Aufwärtspromotoren die bereits zuvor beschriebene Beziehung der Zelltypen untereinander feststellen (Abbildung 3.3-3 (C)). Auch war keine generelle Hypo- oder Hyper-Methylierung einer Klasse ersichtlich - dies galt sowohl für pluripotente als auch ausdifferenzierte Zellen (z.B. ESC und MEF, Abbildung 3.3-3 (D)).

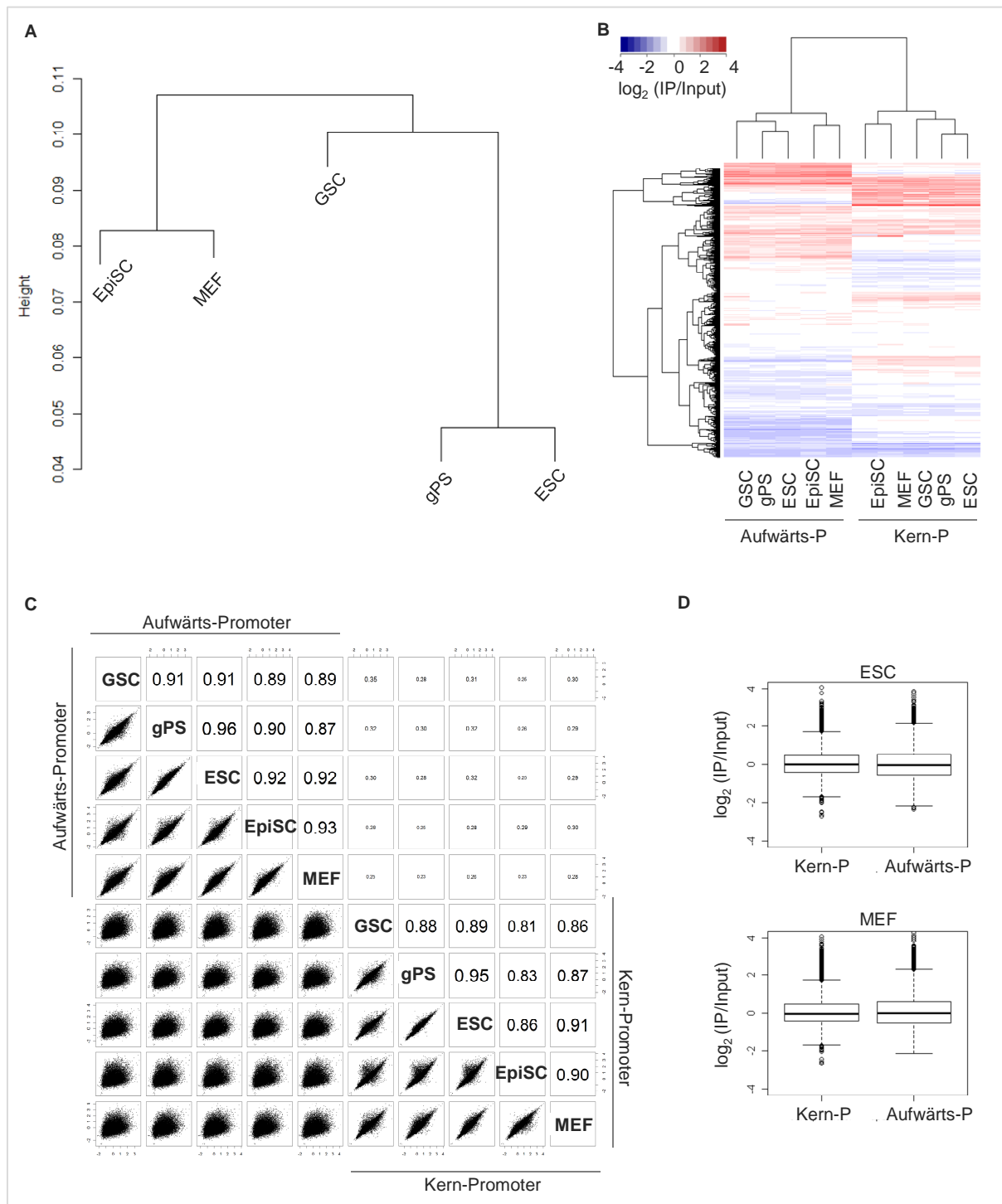


Abbildung 3.3-3: Überblick über MeDIP-Chip-Daten muriner Zelltypen. (A) Mittels hierarchischer Clusteranalyse erstelltes Dendrogramm, (B) Gegenüberstellung der Anreicherungsrechte von Aufwärts- und Kern-Promotoren aller Zelltypen in einer Korrelationsmatrix, (C) Exemplarische Darstellung der Anreicherungsrechte von Aufwärts- und Kern-Promotoren des Chromosoms 7 als Heatmap, (D) Anreicherungslevel der Kern- und Aufwärts-Promotoren zweier Zelltypen: pluripotente ESC und terminal differenzierte MEF.

3.3.3 Analyse der keimbahnassoziierten Stammzellen: gPS gegenüber GSC und ESC

Ausgehend von den in Abschnitt 3.3.2 identifizierten Beziehungen der Zelltypen zueinander wurden detaillierte Analysen differentieller Methylierung vorgenommen. In einem ersten Schritt wurden unterdies die keimbahnassoziierten Zellen im Vergleich zu ESC näher beleuchtet.

3.3.3.1 Genomweite Methylierungsmuster (MeDIP-Chip)

Basierend auf den zuvor ermittelten Median- \log_2 -Werten wurden differentiell methylierte Promotor-Regionen identifiziert. Unter Berücksichtigung der Replikat-Varianz (empirischer Bayes'scher Ansatz, *limma* (Smyth 2004)) wurden statistisch signifikante, >1,5-fache Anreicherungsunterschiede zwischen den verschiedenen Proben berechnet.

Ein erster Einblick in die Entwicklung zellspezifischer Methylierungsmuster der durch Reprogrammierung von unipotenten GSC erhaltenen gPS im Vergleich zu ESC konnte durch Erstellung eines Venn-Diagramms erhalten werden (Abbildung 3.3-4 (A)). Während die pluripotenten Zelltypen (gPS und ESC) untereinander kaum Anreicherungsunterschiede aufwiesen (648 Regionen), differierten diese jeweils erheblich von unipotenten GSC (gPS vs. GSC: 2447, ESC vs. GSC: 2276), wobei eine beträchtliche Überschneidung entsprechend betroffener Regionen zu verzeichnen war.

Durch eine Gruppierung der Regionen in Abhängigkeit systematischer Effekte zwischen den Zelltypen ergaben sich detailliertere Resultate (Abbildung 3.3-4 (B) und (C)).

Bei einem Großteil der 648 identifizierten DMRs zwischen pluripotenten SC war aufgrund der Anreicherungsweite eine Hypomethylierung der gPS zu verzeichnen (Abbildung 3.3-4 (B)). In Abhängigkeit der für GSC im Vergleich zu den obigen Zelltypen berechneten Daten, konnten verschiedene Untergruppen erkannt werden. Etwa 1/3 dieser Promotor-Regionen wiesen dabei gPS-spezifische Muster auf (B1.1, B2.1), welche durch eine inverse Methylierung der reprogrammierten Zellen gegenüber GSC und ESC gekennzeichnet waren. Des Weiteren waren u.a. auch keimbahnspezifische Muster, welche gPS und GSC gemeinsam gegenüber ESC aufzeigten (B1.2, B2.2), zu verzeichnen. Sowohl die gPS- als auch keimbahnspezifischen Methylierungsunterschiede konnten hauptsächlich Aufwärts-Promotoren zugeordnet werden (Anhang C, Abbildung 6.3-1).

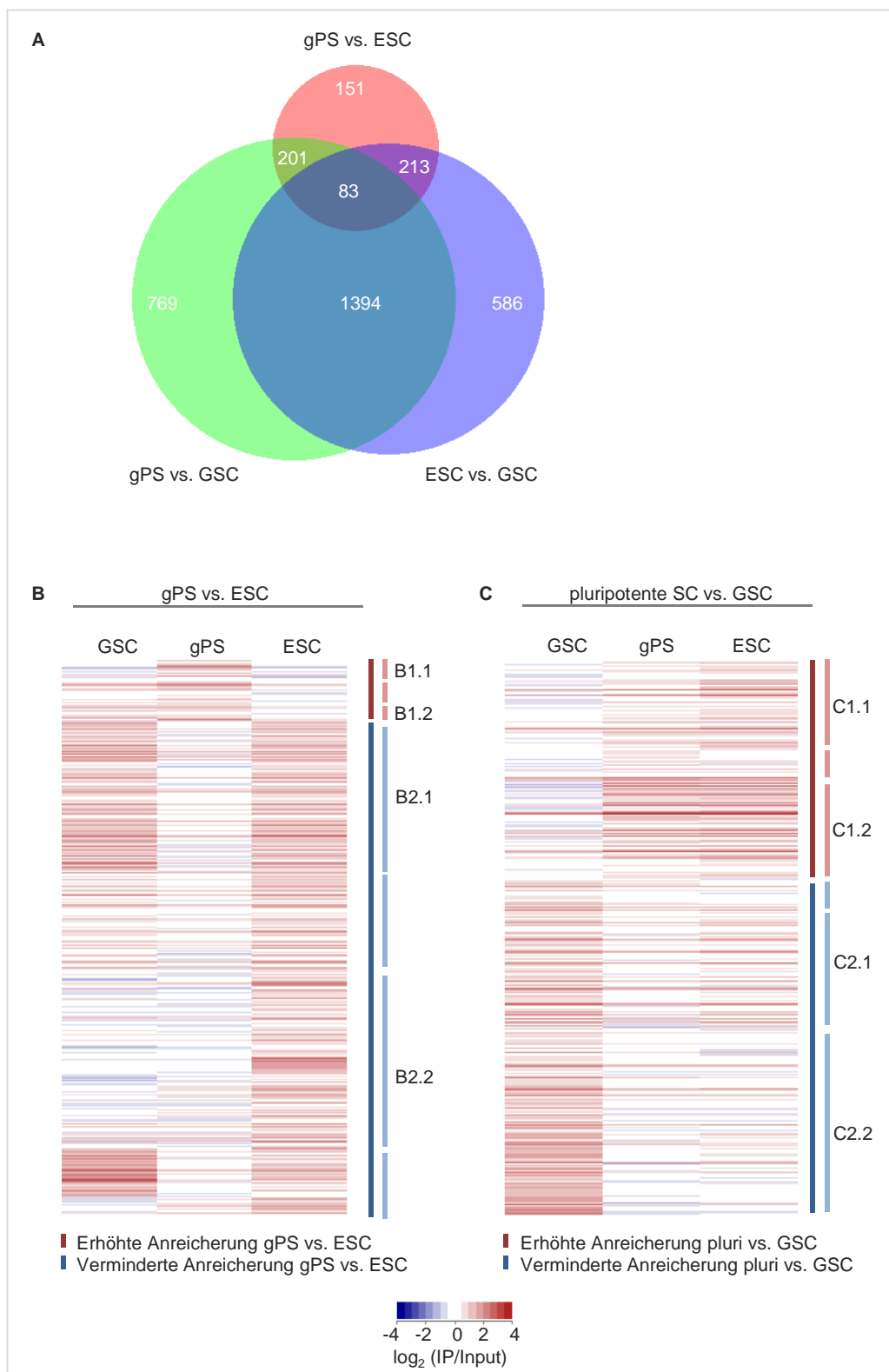


Abbildung 3.3-4: Differenziell methylierte Promotor-Regionen in gPS vs. ESC und GSC. (A) Venn-Diagramm der zwischen den drei Zelltypen differenziell angereicherten Promotor-Regionen weist auf eine größere Ähnlichkeit der gPS mit ESC als mit GSC hin. (B) und (C) Systematische Gruppierung der differenziell angereicherten Promotor-Regionen in Heatmaps zur Veranschaulichung zelltypspezifischer Muster. (B) Promotor-Regionen mit signifikant erhöhter oder verminderter Anreicherung der gPS im Vergleich zu ESC/GSC. (C) Promotor-Regionen mit signifikant erhöhter oder verminderter Anreicherung pluripotenter Stammzellen (pluri SC, gPS/ESC) im Vergleich zu GSC.

Im Gegensatz dazu konnten 2749 DMRs zwischen unipotenten GSC und den pluripotenten Stammzellen (gPS/ESC) festgestellt werden (Abbildung 3.3-4 (C)). Letztere ließen eher Hypomethylierung erkennen. In etwa 1/3 der Fälle lagen dabei die Anreicherungswerte der ESC zwischen denen der GSC und gPS, wobei die Unterschiede zu keiner der beiden Zelllinien signifikant waren (C1.1). Eine ähnliche Beobachtung konnte für ½ der hypermethylierten Gene bzgl. gPS als „Zwischenstufe“ gemacht werden (C1.2). Insgesamt waren die beobachteten Methylierungsunterschiede mit etwa gleicher Häufigkeit in Kern- und Aufwärts-Promotoren zu verzeichnen.

3.3.3.2 Funktionelle Analyse identifizierter Gen-Gruppen

Unter Verwendung der *DAVID*-Software (Huang da et al. 2009) wurde eine funktionelle Analyse mittels Gen-Ontologie durchgeführt, wobei Gene der oben beschriebenen Cluster auf überrepräsentierte biologische Prozesse hin untersucht wurden. Die Gesamtheit der auf dem Microarray analysierten Gene wurde dazu als Referenz genutzt. Wie in Abschnitt 3.2.4 bereits erläutert, wurden Prozesse ähnlicher biologischer Bedeutung in Gruppen zusammengefasst und anhand von Oberbegriffen in Abbildung 3.3-5 und Abbildung 3.3-6 übersichtlich dargestellt. Die vollständigen Ergebnisse dieser Analyse, sowie die entsprechenden Gene der identifizierten Gruppen, befinden sich im Anhang (Anhang C, Tabelle 6.3-1 bis Tabelle 6.3-5 und Tabelle 6.3-12 bis Tabelle 6.3-15).

Die im Vergleich der pluripotenten Zellen als in gPS hypomethyliert identifizierten Gene (Abbildung 3.3-5 (B)) waren u.a. hinsichtlich metaboler und kataboler Prozesse angereichert (Abbildung 3.3-5 (A)). Bei einer Verfeinerung der Analyse durch Ausschluss des keimbahnspezifischen Clusters (Abbildung 3.3-5 (B)) konnten interessanterweise auch Vorgänge der Gewebe-Entwicklung herausgearbeitet werden. Das keimbahnspezifische Cluster selbst hingegen enthielt wie zu erwarten eine Reihe imprinteter Gene, welche aufgrund fehlender GO-Annotationen jedoch nicht bzgl. ihrer Signifikanz erfasst werden konnten und z.T. mit anderen Genen dieses Clusters grob als katabole Prozesse eingeteilt wurden (Abbildung 3.3-5 (C)). In gPS hypermethylierte Gene lieferten aufgrund der geringen Gen-Anzahl keine signifikanten Resultate.

Jene Gene, welche in pluripotenten gPS und ESC im Vergleich zu unipotenten GSC eine Hypermethylierung aufwiesen (Abbildung 3.3-4 (C)), zeigten eine klare Zuordnung zu Prozessen der sexuellen Reproduktion (Abbildung 3.3-6 (A)). Hypomethylierte Gene hingegen waren hinsichtlich einer Aktivierung oder Regulation der Immun-Antwort angereichert (Abbildung 3.3-6 (B)).

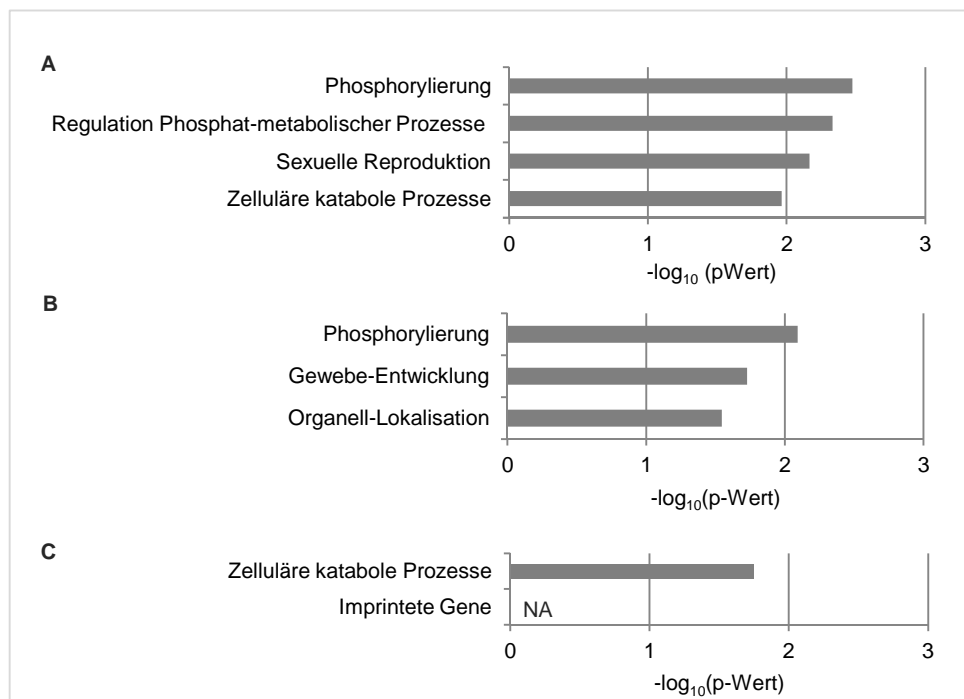


Abbildung 3.3-5: Überrepräsentierte biologische Prozesse für hypo- und hypermethylierte Gene der gPS. Gene mit statistisch signifikanter Hypo- und Hypermethylierung der gPS im Vergleich zu ESC/GSC wurden mittels der DAVID-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um angereicherte biologische Prozesse (GO-Ontologie) zu ermitteln, für (A) hypomethylierte Gene der gPS, (B) hypomethylierte Gene der gPS ohne keimbahn-spezifische Muster, (C) keimbahn-spezifisch hypomethylierte Gene der gPS. Dargestellt sind die Oberbegriffe der in Gruppen zusammengefassten Prozesse ähnlicher biologischer Bedeutung.

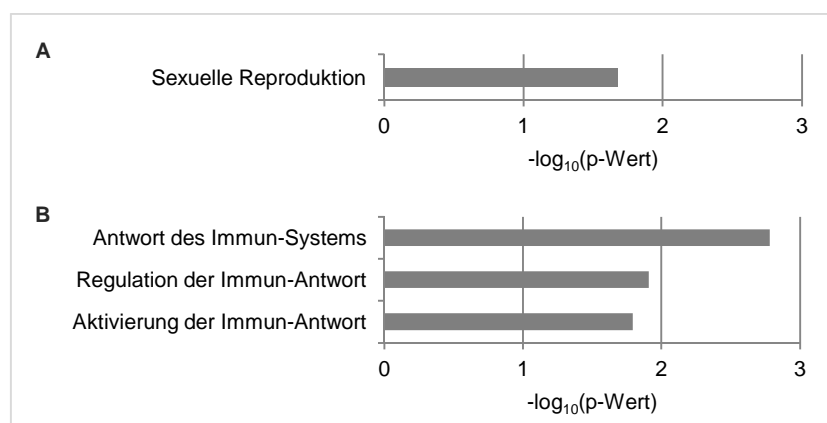


Abbildung 3.3-6: Überrepräsentierte biologische Prozesse für hypo- und hypermethylierte Gene der pluripotenten Stammzellen. Gene mit statistisch signifikanter Hypo- und Hypermethylierung der pluripotenten ESC/gPS im Vergleich zu GSC wurden mittels der DAVID-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um angereicherte biologische Prozesse (GO-Ontologie) zu ermitteln, für (A) hypermethylierte Gene der pluripotenten Zellen, (B) hypomethylierte Gene der pluripotenten Zellen. Dargestellt sind die Oberbegriffe der in Gruppen zusammengefassten Prozesse ähnlicher biologischer Bedeutung.

3.3.4 Analyse späterer Pluripotenz-Stadien: EpiSC gegenüber MEF und ESC

Basierend auf den in Abschnitt 3.3.2 angedeuteten Beziehung der Zelltypen zueinander, wurde eine detaillierte Analyse der durch EpiSC repräsentierten vorgeplanten Pluripotenz im Vergleich zu naiv-pluripotenten und ausdifferenzierten Stadien (ESC bzw. MEF) vorgenommen.

3.3.4.1 Genomweite Methylierungsmuster (MeDIP-Chip)

Differentiell methylierte Promotor-Regionen wurden, wie bereits zuvor beschrieben (Abschnitt 3.3.3.1), ausgehend von den Median- \log_2 -Werten berechnet. Die ermittelten Regionen wiesen signifikante, >1,5-fache Anreicherungsunterschiede zwischen den betrachteten Proben auf.

Mit Hilfe eines Venn-Diagramms konnte ein erster Eindruck der zellspezifischen Methylierungsmuster in Abhängigkeit des Differenzierungsstatus erhalten werden (Abbildung 3.3-7 (A)). Dabei zeigten die drei Zelltypen EpiSC, ESC und MEF untereinander jeweils eine erhebliche Anzahl differentiell angereicherter Regionen (EpiSC vs. ESC: 1881, EpiSC vs. MEF: 1849, ESC vs. MEF: 1950), wobei entsprechend beträchtliche Überschneidungen zu verzeichnen waren. Wie bereits zuvor in der hierarchischen Cluster-Analyse angedeutet, teilten EpiSC hierbei mehr Anreicherungsmuster mit MEF gegenüber ESC als mit ESC gegenüber MEF (EpiSC und ESC vs. MEF: 575; EpiSC und MEF vs. ESC: 699).

Durch eine Gruppierung der identifizierten Regionen in Abhängigkeit systematischer Effekte zwischen den Zelltypen konnten umfassendere Ergebnisse gewonnen werden (Abbildung 3.3-7 (B) und (C)).

Generell waren etwa gleich viele Regionen mit deutlicher Hyper- und Hypomethylierung der EpiSC im Vergleich zu ESC und/oder MEF zu verzeichnen (962 hypermethyliert, 867 hypomethyliert, Abbildung 3.3-7 (B)). Dabei konnten verschiedene Untergruppen identifiziert werden. Etwa die Hälfte der hypermethylierten Promotor-Regionen wiesen „gewebespezifische“ Muster auf, welche durch ähnliche Anreicherungswerte der EpiSC und MEF im Vergleich zu ESC gekennzeichnet waren (462 DMRs, Abbildung 3.3-7 (B1.1)). Ein weiteres großes Cluster hingegen zeigte EpiSC-spezifische Profile (343 DMRs, Abbildung 3.3-7 (B1.2)). Im Falle der hypomethylierten Regionen konnten neben letztgenanntem Effekt (212 DMRs, Abbildung 3.3-7 (B2.1)) vor allem stammzellspezifische (SC) Muster - übereinstimmende Anreicherungswerte der EpiSC und ESC im Vergleich zu MEF - detektiert werden (B2.2, 418 DMRs).

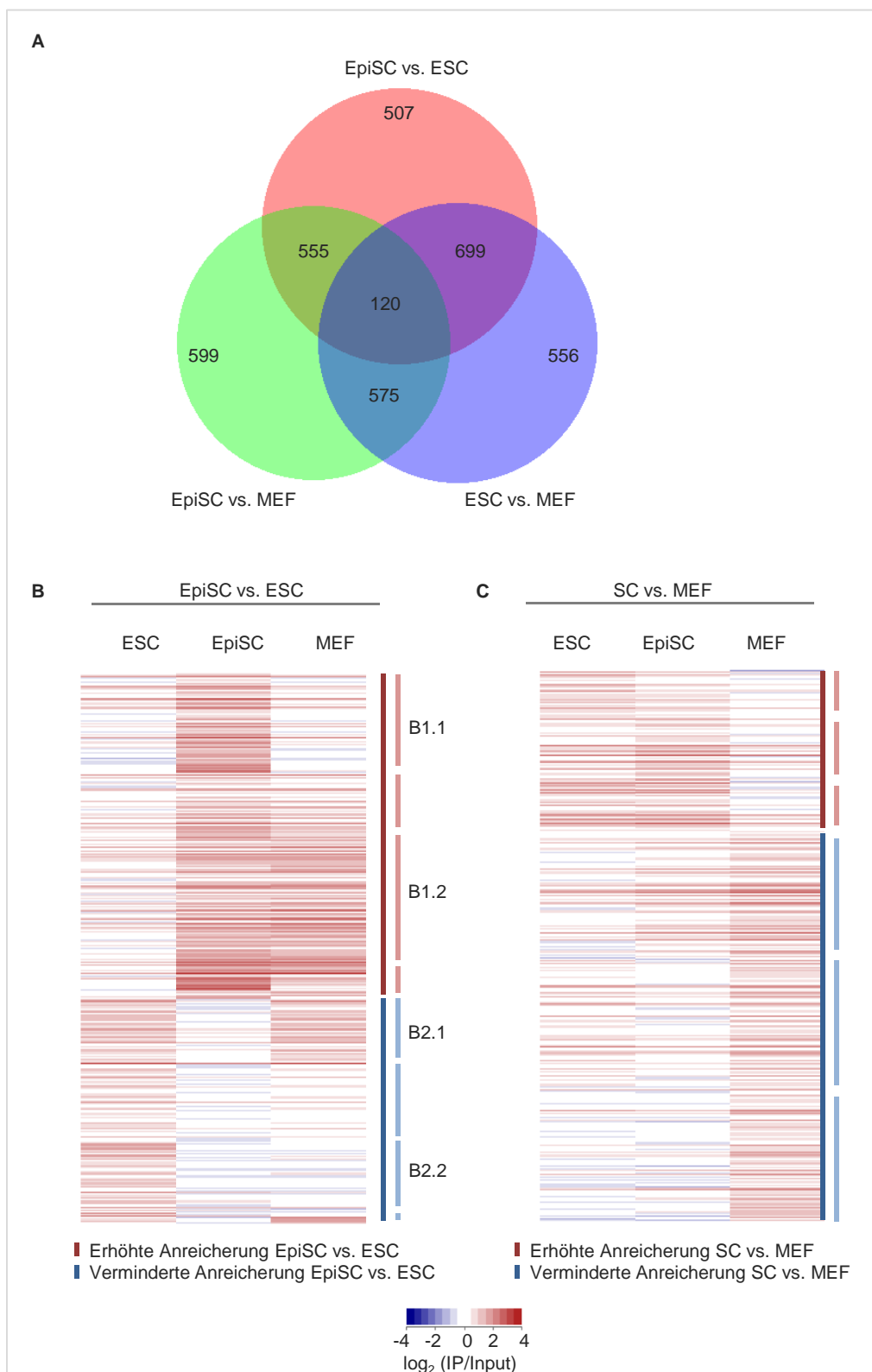


Abbildung 3.3-7: Differenziell methylierte Promotor-Regionen in EpiSC vs. ESC und MEF. (A) Venn-Diagramm der zwischen den drei Zelltypen differenziell angereicherten Promotor-Regionen. (B) und (C) Systematische Gruppierung der differenziell angereicherten Promotor-Regionen in Heatmaps zur Veranschaulichung zelltypspezifischer Muster. (B) Promotor-Regionen mit signifikant erhöhter oder verminderter Anreicherung der EpiSC, mit stammzell-, gewebe- oder EpiSC-spezifischen Mustern. (B) Dynamische Anreicherungsunterschiede zwischen ESC, EpiSC und MEF.

EpiSC- und SC-spezifische Hypomethylierung war dabei vermehrt mit Aufwärts-Promotoren, Hypermethylierung dagegen eher mit Kern-Promotoren assoziiert. Gewebespezifische Muster zeigten in diesem Zusammenhang keine klare Tendenz (Anhang C, Abbildung 6.3-2). Darüber hinaus wurden bei weiteren 1782 DMRs Muster erkannt, welche an- bzw. absteigende Anreicherungswerte zwischen den Zelltypen offenbarten und dynamische Methylierungsänderungen in Abhängigkeit des Entwicklungsstadiums andeuten könnten (Abbildung 3.3-7 (C)). Dabei war für die Zwischenstufe (Zelltyp mit mittlerem Anreicherungswert) meist kein signifikanter Unterschied im Vergleich zu den hypo- und hypermethylierten Zelltypen zu verzeichnen. Promotor-Regionen mit Anreicherungswerten der EpiSC zwischen ESC und MEF waren hauptsächlich in ESC hypomethyliert (ESC<EpiSC<MEF, 418/ 581 DMRs, Abbildung 3.3-7 (C2)). Umgekehrt konnte für Regionen mit ESC als „Zwischenstufe“ überwiegend eine Hypomethylierung der EpiSC detektiert werden (EpiSC<ESC<MEF, 430/632 DMRs, Abbildung 3.3-7 (C1)). DMRs mit mittleren Anreicherungswerten der MEF hingegen wiesen gleichermaßen in EpiSC und ESC hypomethylierte Regionen auf (EpiSC<MEF<ESC, 289/569 DMRs, Abbildung 3.3-7 (C3.1) und ESC<MEF<EpiSC, 280/569 DMRs, Abbildung 3.3-7 (C3.2)).

3.3.4.2 Funktionelle Analyse identifizierter Gen-Gruppen

Wie bereits zuvor beschrieben, wurde zur Identifizierung überrepräsentierter biologischer Prozesse der einzelnen Gen-Cluster eine funktionelle Analyse mittels Gen-Ontologie (*DAVID*-Software (Huang da et al. 2009)) vorgenommen. Die Oberbegriffe der jeweiligen Gruppen wurden in Abbildung 3.3-8 und Abbildung 3.3-9 übersichtlich zusammengefasst. Eine vollständige Auflistung der einzelnen Kategorien, sowie die Gene der identifizierten Gruppen, befinden sich im Anhang (Anhang C, Tabelle 6.3-6 bis Tabelle 6.3-11 und Tabelle 6.3-16 bis Tabelle 6.3-21).

Die hypermethylierten Gene des gewebespezifischen Clusters waren hinsichtlich reproduktiver Prozesse der Meiose angereichert (Abbildung 3.3-8 (A.2)). Während für Gene des EpiSC-spezifischen Clusters im Falle der Hypermethylierung lediglich ein Bezug zur Ionen Homeostase festgestellt werden konnte (Abbildung 3.3-8 (A.1)), zeigten die entsprechend hypomethylierten Gene (u.a. mehrere Hox-Gene, Anhang C, Tabelle 6.3-18) eine klare Zuordnung zu embryonaler Entwicklung (Regionalisierung, Organ-Entwicklung, Abbildung 3.3-8 (B.1)).

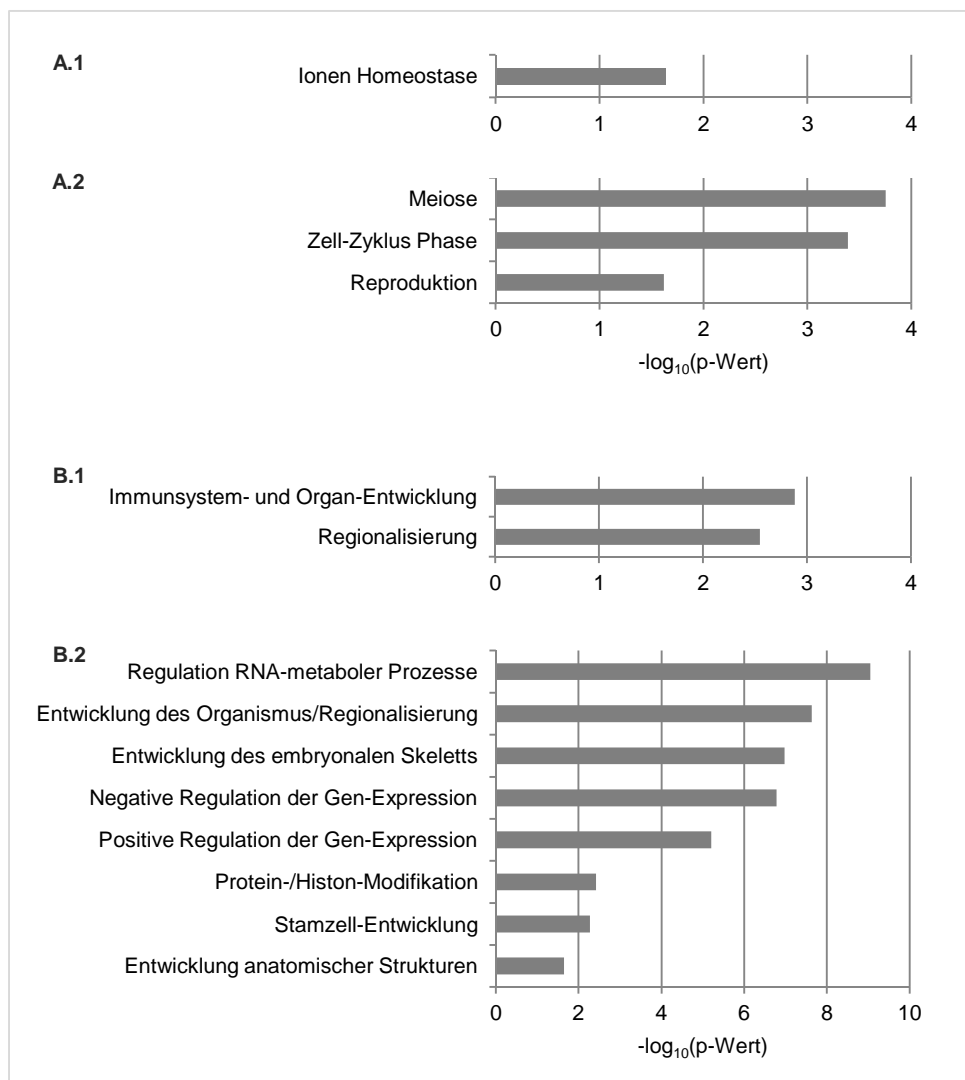


Abbildung 3.3-8: Überrepräsentierte biologische Prozesse für hypo- und hypermethylierte Gene der EpiSC. Gene mit statistisch signifikanter Hypo- und Hypermethylierung der EpiSC im Vergleich zu ESC/MEF mit stammzell-, gewebe- oder EpiSC-spezifischen Mustern wurden mittels der DAVID-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um angereicherte biologische Prozesse (GO-Ontologie) zu ermitteln, für (A.1) EpiSC-spezifische Hypermethylierung, (A.2) gewebespezifische Hypermethylierung, (B.1) EpiSC-spezifische Hypomethylierung, (B.2) stammzellspezifische Hypomethylierung. Dargestellt sind die Oberbegriffe der in Gruppen zusammengefassten Prozesse ähnlicher biologischer Bedeutung.

Ähnliche Kategorien (Entwicklung anatomischer Strukturen und des embryonalen Skeletts) waren ebenfalls für hypomethylierte Gene mit pluripotenzspezifischen Anreicherungsmustern vertreten (u.a. mehrere Hox- und Fox-Gene, Anhang C, Tabelle 6.3-19). Darüber hinaus waren Prozesse hinsichtlich Regulation der Genexpression, Stammzell-Entwicklung sowie Histon-Modifikationen zu verzeichnen (Abbildung 3.3-8 (B.2)).

Auch jene Gene, welche von EpiSC über ESC hin zu MEF eine zunehmende Methylierung aufwiesen, ließen neben Prozessen hinsichtlich der Regulation der Genexpression ebenfalls einen Bezug zur Organ- und Zellentwicklung erkennen (Abbildung 3.3-9 (A)). Im Gegensatz dazu konnte für das Cluster mit mittleren Anreicherungswerten der EpiSC im Vergleich zu hypomethylierten ESC und hypermethylierten MEF eine Zuordnung zu Phosphat-Metabolismus und Protein-Katabolismus herausgearbeitet werden (Abbildung 3.3-9 (B)).

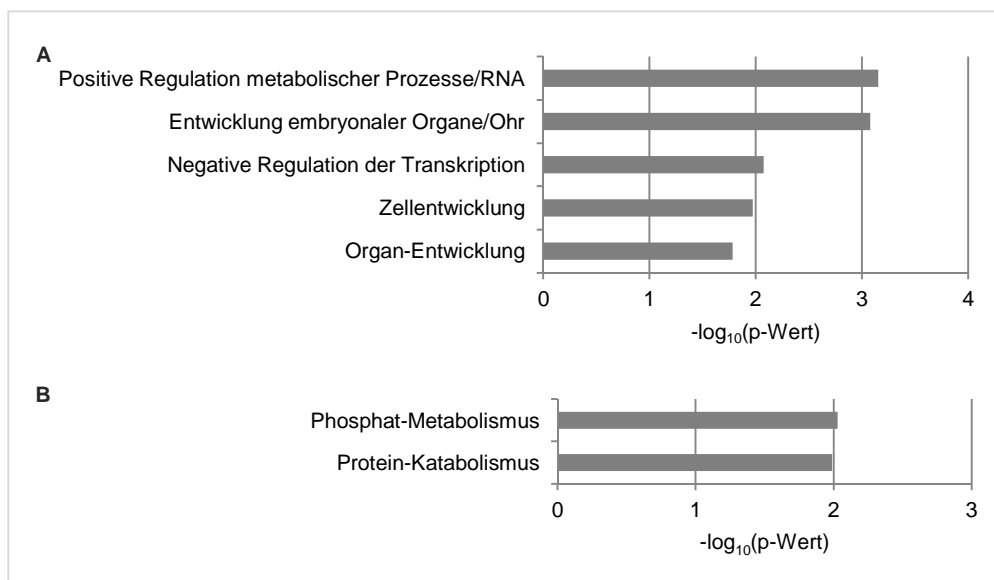


Abbildung 3.3-9: Überrepräsentierte biologische Prozesse für dynamische Anreicherungsunterschiede zwischen EpiSC, ESC und MEF. Dynamisch differentiell angereicherte Gene zwischen EpiSC/ESC/MEF wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um angereicherte biologische Prozesse (GO-Ontologie) zu ermitteln, für (A) EpiSC<ESC<MEF, (B) ESC<EpiSC<MEF. Dargestellt sind die Oberbegriffe der in Gruppen zusammengefassten Prozesse ähnlicher biologischer Bedeutung.

3.3.5 Validierung der MeDIP-Chip-Daten mittels Bisulfit-Sequenzierung

Zur Überprüfung der beschriebenen Resultate und Erfassung der quantitativen Methylierungsunterschiede zwischen den entsprechenden Proben wurden einzelne differentiell angereicherte Promotor-Regionen mittels Hochdurchsatz-Sequenzierung Bisulfit-behandelter DNA analysiert. Zu diesem Zweck wurden jeweils Kandidaten der verschiedenen ermittelten Gen-Cluster aus Abschnitten 3.3.5 und 3.3.6 ausgewählt.

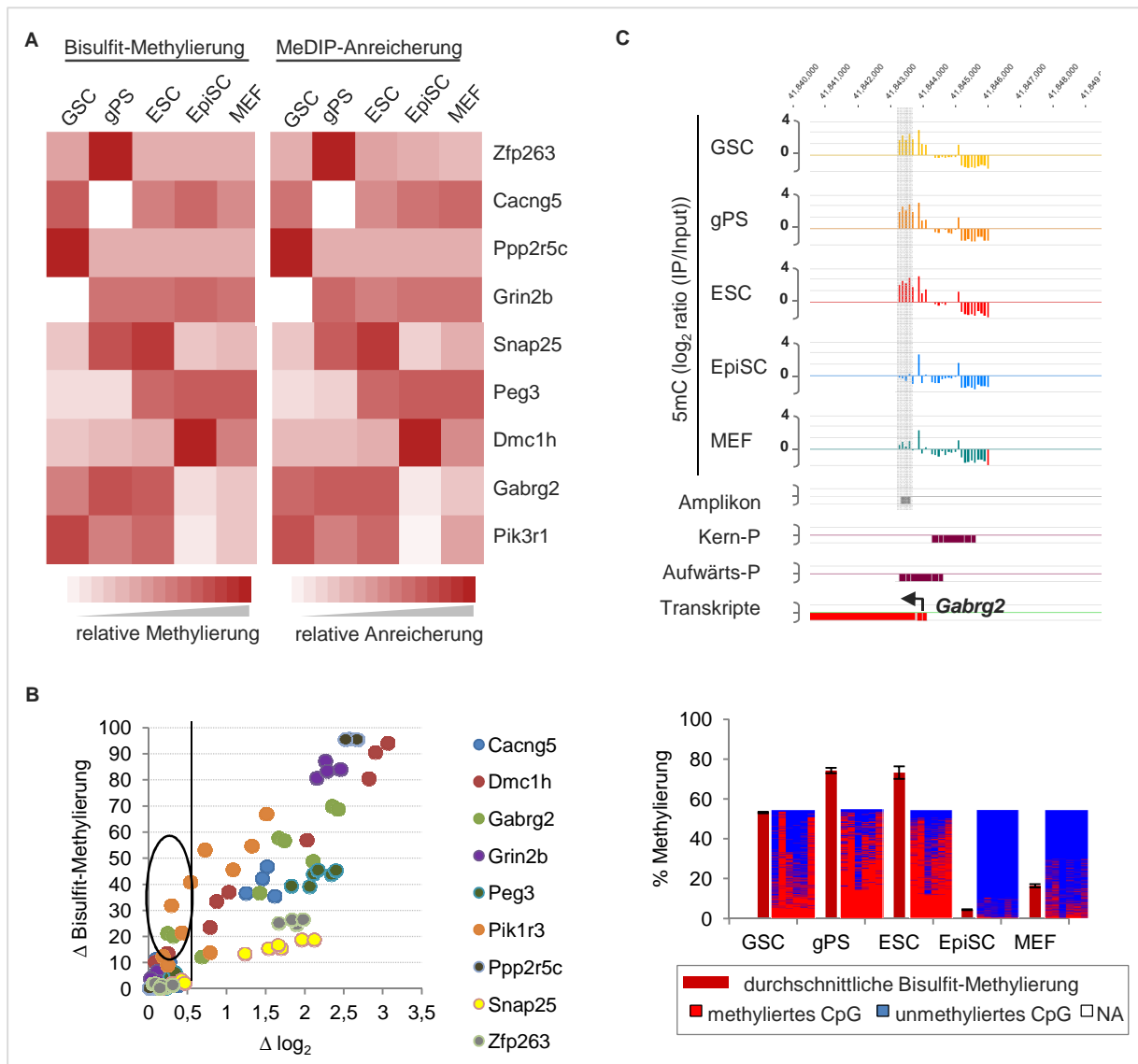


Abbildung 3.3-10: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung. (A) Heatmaps der mittleren Bisulfit-Methylierung bzw. Anreicherung der untersuchten Regionen, dargestellt als relativer Wert pro Amplikon (weiß: Zelltyp mit niedrigstem Wert, rot: Zelltyp mit höchstem Wert). (B) Gegenüberstellung der im paarweisen Vergleich der Zelltypen erhaltenen delta-Werte für Bisulfit-Methylierung und Anreicherung. Einige Regionen mit delta-Methylierung <30% wurden nicht als differentiell angereichert (>0,5 delta- \log_2) erfasst. (C) Anreicherungsprofil von *Gabrg2* (\log_2 -Ratios relativ zu genomischen Koordinaten sowie funktionellen Regionen) mit Resultaten der Bisulfit-Hochdurchsatz-Sequenzierung. Dargestellt ist die mittlere Bisulfit-Methylierung pro Zelltyp sowie die jeweiligen Methylierungsprofile (ein Replikat) mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen.

In Abbildung 3.3-10 (A) wurden Methylierungs- und Anreicherungsdaten der analysierten Amplikons jeweils in Form von Heatmaps gegenübergestellt, wobei zur Verdeutlichung der zelltypspezifischen Muster die Proben im Verhältnis zueinander dargestellt wurden. Dabei wird die hohe qualitative Übereinstimmung der Daten deutlich. In Abhängigkeit der Sequenzeigenschaften und Methylierungslevel (Anzahl der AK-Epitope) der jeweiligen

Regionen, welche einen Einfluss auf die Sensitivität ausüben, konnten Methylierungsunterschiede $>30\%$, z.T. sogar $>12\%$, (jeweils in beiden biologischen Replikaten) mit hoher Sensitivität und Spezifität mittels MeDIP-Chip ($\Delta\log_2 >0,5$) detektiert werden (Abbildung 3.3-10 (B)).

Bei detaillierter Betrachtung der einzelnen Methylierungsprofile konnten im Falle unvollständiger Methylierung meist unterschiedliche Sequenzpopulationen (mosaikal- bzw. hoch-methylierte sowie unmethylierte Sequenzen) erkannt werden, deren relatives Verhältnis zwischen den einzelnen Proben abwich (z.B. Zfp293, Cacng5, Anhang C, Abbildung 6.3-3). In einigen Fällen konnte zusätzlich auch ein deutlicher Positionseffekt verzeichnet werden, wobei eine Methylierungs-Präferenz bestimmter CpG-Dinukleotide innerhalb der Region ersichtlich war (z.B. Gabrg2, Abbildung 3.3-10 (C)). Die übrigen Sequenzprofile befinden sich im Anhang (Anhang C, Abbildung 6.3-3 bis Abbildung 6.3-6).

4 Diskussion

4.1 Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung

In den letzten Jahren wurden eine Reihe von Methoden zur Erstellung genomweiter Methylierungsprofile entwickelt (Laird 2010, Wu et al. 2012). Als Gold-Standard zur Erfassung quantitativer Methylierungsdaten auf Basenpaar-Ebene gilt noch immer die Bisulfit-Sequenzierung. Eine Übertragung/Anwendung auf genomweiter Ebene ist allerdings kostspielig, erfordert eine erhebliche Sequenzier-Leistung und erweist sich hinsichtlich der Analyse aufgrund der reduzierten Sequenzinformation als komplex. Alternative Ansätze basieren z.B. auf Immunpräzipitationen und erlauben, in Kombination mit Sequenzierung oder Microarrays, eine relativ kostengünstige Alternative.

Im Zuge dieser Entwicklung sollte zur Bearbeitung der in Abschnitten 3.2 und 3.3 dargestellten biologischen Fragestellungen die damals noch neue Methode MeDIP-Chip etabliert werden. Durch gründliche Überprüfung und Einstellung experimenteller Parameter konnte dabei die AK-vermittelte Anreicherung methylierter DNA erfolgreich etabliert werden. Des Weiteren wurde eine Abfolge bioinformatischer Analysen zur Qualitätskontrolle und Prä-Prozessierung der erhaltenen Microarray-Daten erarbeitet. Zur Identifizierung differentiell methylierter Regionen wurde, hinsichtlich der in Abschnitt 3.2 angestrebten Erfassung verschiedener funktioneller Bereiche, ein Ansatz zur Analyse chromosomenweiter Methylierungsprofile entwickelt. Dieser konnte anhand publizierter quantitativer Daten aus Bisulfit-Sequenzierung validiert werden. Durch Integration verschiedener Kontrollen sowohl auf experimenteller als auch analytischer Ebene konnten Defizite und Limitierungen der Methode erkannt und im weiteren Verlauf der Arbeit berücksichtigt werden. Die einzelnen Punkte werden im Folgenden näher diskutiert.

4.1.1 Wesentliche Aspekte hinsichtlich experimenteller Parameter

Basierend auf dem Originalprotokoll (Weber et al. 2005) konnte die Immunpräzipitation methylierter DNA mittels gegen 5-Methylcytosin gerichteter Antikörper erfolgreich etabliert werden. Verschiedene Qualitätskontrollen, insbesondere die qPCR, dienten einer

Bewertung einzelner Parameter bzw. Experimente, so dass hochwertige Präzipitate generiert bzw. identifiziert werden konnten.

Es wurde eine effiziente und spezifische Anreicherung methylierter Regionen im Vergleich zu minimalen Hintergrundsignalen erreicht. Eine der Grundvoraussetzungen ist dabei die **Qualität des Ausgangsmaterials** (hochmolekulare, RNA- und RNase-freie DNA). Es konnte beispielsweise gezeigt werden, dass die Anwesenheit von RNA und insbesondere RNase offenbar durch unspezifische Bindung des Antikörpers zu verstärkter Anreicherung unmethylierter DNA und somit zu erheblichen Hintergrundsignalen führte (Abbildung 6.1-1). Dieser Effekt trat allerdings erst im Verlaufe der Optimierungsversuche auf und schien abhängig von der AK-Charge zu sein. Um eine gleichbleibende Qualität des Präzipitats sicherzustellen, wurde im Folgenden (unter Verwendung ohnehin zeiteffektiverer kommerzieller Kits) auf die gründliche Entfernung des Enzyms geachtet.

Des Weiteren ist das eingesetzte **AK:DNA-Verhältnis** für die Güte der Präzipitation von Bedeutung. Wie in einer entsprechenden Versuchsreihe dargestellt werden konnte (Abschnitt 3.1.1.4), sind verbesserte Präzipitationsraten aufgrund höherer AK-Mengen mit einer Zunahme des Hintergrundsignals durch unspezifische AK-Reaktionen verbunden. Das hier gewählte Verhältnis von 2:1 stellt dabei einen Kompromiss zwischen Spezifität, Sensitivität (Erfassung CpG-armer Regionen), Ausbeute und Kosten dar. Weitere nicht näher ausgeführte Abwandlungen des Originalprotokolls konnten in dieser Hinsicht - insbesondere bzgl. der Anreicherungsrate - zusätzliche, wenn auch geringfügige, Effekte erzielen. Dazu zählen z.B. eine verlängerte Inkubation des primären Antikörper (ÜN statt 2h) sowie eine Denaturierung des Präzipitats vor dessen Aufreinigung, um die Effizienz des nachfolgenden Proteinase K-Verdau zur Degradation der AK-DNA-Komplexe zu erhöhen. Eine weitere Optimierung könnte durch Variation der Menge des 2° Antikörpers, ein in dieser Arbeit nicht näher überprüfter Faktor, beitragen. Dabei müsste durch systematische Überprüfung erneut ein Gleichgewicht zwischen Spezifität und Ausbeute ermittelt werden. Zu diskutieren ist auch einer der ersten Teilschritte des Protokolls - die **Fragmentierung genomischer DNA**. Die hierzu verwendeten Methoden hatten keinen Einfluss auf die Qualität der darauffolgenden Immunpräzipitation, unterschieden sich aber wesentlich in ihrer Effizienz. Die Nebulisierung lieferte hierbei, im Gegensatz zur Sonifizierung, reproduzierbar und verlässlich die zu erwartenden Fragmentgrößen. Bei der Druckvermittelten Zerstäubung der Flüssigkeit (Nebulisierung) lassen sich die Scherkräfte besser kontrollierten als im Falle der Ultraschall-basierten Methode (Wasserbad). Letztere wird möglicherweise durch weitere, bei der Standardisierung nicht berücksichtigte Parameter beeinflusst (z.B. Luftdruck oder Luftfeuchtigkeit). Auch könnte die Alterung des Geräts zu

einer ungleichmäßigen Leistung beigetragen haben. Neben der hier dargestellten Zweckmäßigkeit der Nebulisierung, sind auch folgende Nachteile zu verzeichnen: Zum Einen tritt ein erheblicher Materialverlust auf, da die zerstäubte Probe nicht komplett aus dem Nebulisierungsgefäß wiedergewonnen werden kann. Zum Anderen kann, insbesondere bei größeren Ausgangsmengen an DNA, aufgrund der notwendigen Volumina einzusetzenden Puffers die anschließende Aufreinigung unhandlich werden. Die Aufteilung der Probe in mehrere Fraktionen birgt dabei die Gefahr von Kontamination und Verwechslung. Eine enzymatisch vermittelte Fragmentierung könnte eine interessante Alternative zu den hier verwendeten mechanischen Verfahren darstellen. Neuere Produkte wie z.B. die Shearase von Zymo (Irvine, USA), welche mittels degenerierter Erkennungssequenzen ebenfalls zu einer zufälligen Fragmentierung führen, könnten eine noch gezieltere, vereinfachte und zeiteffektivere Bearbeitung ermöglichen. Des Weiteren könnte möglicherweise eine Verbesserung der Präzipitation bzw. der WGA erzielt werden, da im Gegensatz zur mechanischen Fragmentierung keine zusätzlichen Beschädigungen wie Einzelstrangbrüche auftreten sollten und zudem „blunt-ends“ generiert werden.

Unter Einbeziehung der gesammelten Erfahrung und verbesserten Kontrollen könnte eine erneute systematische und kritische Überprüfung der experimentellen Teilschritte zu einem effizienteren Ablauf bzw. einer Zeitersparnis führen.

4.1.2 Unterschiedliches Präzipitationsverhalten humaner und muriner DNA

Im Verlaufe der in dieser Arbeit durchgeführten Experimente konnten einige interessante Unterschiede bzgl. der Präzipitation humanen und murinen Materials festgestellt werden: eine bei Maus-DNA zu verzeichnende i) geringere Dynamik der Präzipitations-Effizienz in Abhängigkeit des eingesetzten AK:DNA-Verhältnisses; ii) niedrigere spezifische Anreicherung im Vergleich zum Hintergrund; iii) höhere Anreicherungsrate. Diese im Folgenden diskutierten Ergebnisse legen eine Etablierung bzw. Optimierung der Versuchsparameter für jede zu analysierende Spezies nahe.

Der erste Punkt könnte zumindest teilweise auf variierende Versuchs- bzw. PCR-Bedingungen zurückzuführen sein. Dabei wurden die Proben der humanen Versuchsreihe z.B. nicht, wie im Falle der murinen Tests, mit dem späteren standardisierten qPCR-Protokoll untersucht. Die bei murinen IPs festgestellte niedrigere Anreicherung im Verhältnis zum Hintergrund (bis zu 10-fach) konnte allerdings ebenfalls bei paralleler Versuchsdurchführung unter vergleichbaren, standardisierten Versuchsbedingungen (AK:DNA-Ratio 2:1) beobachtet werden. Diese resultierte aus einer geringeren spezifischen

Anreicherung methylierter Regionen im Verhältnis zu einer leicht erhöhten Anreicherung der Negativkontrolle. Technische Ursachen sind in diesem Zusammenhang nicht ersichtlich. Zum Einen wurden diese Resultate unabhängig der Herkunft bzw. Produktionscharge des Antikörpers erhalten (Daten nicht gezeigt). Des Weiteren konnte insgesamt nur ein geringer Hintergrund aufgrund unspezifischer Präzipitationen primärer und sekundärer Antikörper festgestellt werden. Zudem lagen die für murine IPs ermittelten Daten im Rahmen publizierter Werte der gleichen Regionen (Weber et al. 2005). Es ist zu berücksichtigen, dass in Mensch und Maus jeweils nur wenige und zudem unterschiedliche Regionen bzgl. ihrer Anreicherung analysiert wurden, welche in Abhängigkeit ihrer Sequenzeigenschaften und Bindestellen in unterschiedlicher Weise die Präzipitation beeinflussen könnten.

In Zusammenhang mit der beobachteten höheren, fast doppelten Anreicherungsrate muriner Proben ist zu vermuten, dass unterschiedliche Genom-Kompositionen bzw. Verteilungen verfügbarer CpG-Positionen zu abweichenden Anreicherungsmustern führen. Da repetitive Elemente einen Großteil des präzipitierten Materials darstellen sollten, wäre z.B. eine im Vergleich zum Menschen verstärkte Methylierung dieser Regionen denkbar. Dies würde nicht nur die insgesamt erhöhte Ausbeute erklären. In Verbindung mit Andeutungen der Literatur, dass murine Sequenzen generell weniger bzw. schwächere und kürzere CpG-Inseln aufweisen (Matsuo et al. 1993, Jiang et al. 2007, Aissani und Bernardi 1991, Cross et al. 1997), würde dies auch eine anteilig geringere Anreicherung uniquer Regionen bedingen. Zur Überprüfung dieser Hypothese wären vergleichende Analysen hochauflösender genomweiter Methylierungsdaten im Sequenzkontext nötig.

4.1.3 Qualitätskontrollen, Prä-Prozessierung und Vergleichbarkeit von Datensätzen

Zur Auswertung der mittels Microarray-Hybridisierung der MeDIP-Präzipitate erzeugten Daten wurde in der vorliegenden Arbeit eine Abfolge bioinformatischer Analysen erarbeitet. Ausgehend von gängigen Methoden zur Verarbeitung von Expressions- bzw. ChIP-Chip-Daten wurden funktionelle Kommandos zur Qualitätskontrolle und Prä-Prozessierung der vom Dienstleister erhaltenen Rohdaten zusammengestellt. Auf diese Weise war es möglich, problematische Datensets zu erkennen und gegebenenfalls von weiteren Analysen auszuschließen, um so spätere Fehlinterpretationen zu vermeiden.

Die Bedeutung einer gründlichen Überprüfung von Microarray-Daten zur Identifikation technischer Variabilitäten und Artefakte sowie deren Angleichung wurde bereits mehrfach

erläutert (Huber et al. 2005, Bolstad et al. 2003, Quackenbush 2002). Geeignete Normalisierungsverfahren müssen hierbei aufgrund der von üblichen Microarray-Analysen (Expressions-Arrays) abweichenden Datenstruktur des MeDIP-Chips mit Bedacht gewählt werden. In der vorliegenden Arbeit konnte anhand von Datenbeispielen gezeigt werden, dass zur Vermeidung künstlicher Signalverzerrungen globale Methoden wie Median-Zentrierung oder Quantil-Normalisierung vorteilhaft sind. Eine Inter-Array-Angleichung geringer technischer Variabilitäten wird dabei durch eine einfache Skalierung der Signalintensitäten erreicht. Damit scheinen diese Methoden einen vergleichsweise geringen Einfluss auf die Datenverteilung zu nehmen. Jedoch stellen die gewählten Ansätze keine optimale Lösung dar. Zum Einen könnte die globale Anpassung auch zu einem Informationsverlust führen, z.B. im Falle von Zelltypen mit unterschiedlichem Methylierungslevel (Stammzellen vs. differenzierte Zellen). Des Weiteren können stärkere Verzerrungen, welche auf aberranten Signalverteilungen der Kanäle im Verhältnis zueinander beruhen, nicht korrigiert werden. Die Anforderungen an eine entsprechende Intra-Array-Normalisierung sind allerdings aufgrund der Unterschiede des pro Farbstoff hybridisierten Materials (IP vs. Input) komplex und werden von keiner bisher verfügbaren Methode, wie z.B. LOWESS, erfüllt. In der Literatur finden sich bereits einige Ansätze hinsichtlich einer Entwicklung geeigneter Verfahren zur Analyse von MeDIP-Chip-Daten, welche z.T. eigens auf die verwendeten Microarrays zugeschnitten sind (Penterman et al. 2007, Dindot et al. 2009, Irizarry et al. 2009). Eine Bereitstellung und Standardisierung solcher Algorithmen könnte zu einer optimierten Datenverarbeitung beitragen.

Angesichts der erläuterten Problematik bzgl. der Korrektur von Datenverzerrungen ergibt sich eine Limitierung in der Vergleichbarkeit verschiedener Datensets. Anhand der Daten zweier unabhängiger Experimente konnte gezeigt werden, dass aberrante Signalverteilungen zu einer qualitativ geringeren Erfassung des Methylierungsmusters führen. Die Ursache für derartige Qualitätsunterschiede der MeDIP-Chip-Daten, kann nur vermutet werden. Da sowohl amplifizierte als auch nicht-amplifizierte Proben der entsprechenden Versuchsreihe betroffen waren, kann ein Einfluss der WGA ausgeschlossen werden. Neben sonstigen Abweichungen auf experimenteller Ebene (z.B. DNA-Qualität) sind differierende Bedingungen der Microarray-Prozessierung (Labeling, Hybridisierung, Scannen) wahrscheinlich. Tatsächlich konnten bereits leichtere Effekte auf die Signalverteilung bei erneuter Hybridisierung derselben Probe festgestellt werden (Anhang A, Abbildung 6.1-4). Durch rechtzeitige Identifikation und kritische Beurteilung solcher Artefakte mit Hilfe der dargestellten Analyse-Schritte können vergleichende Analysen entsprechender Datensets und damit spätere Fehlinterpretationen vermieden werden.

4.1.4 Verlust der Sensitivität durch genomrepräsentative Amplifikation

Im Falle limitierten Ausgangsmaterials war vor der Array-Prozessierung eine Amplifikation der Proben nötig. Da in der Literatur diesbezüglich Datenverzerrungen mit hohem CpG- bzw. GC-Gehalt in Verbindung gebracht wurden (Jia et al. 2010, Robinson et al. 2010, Johnson et al. 2008), sollte die Validität dieser Methode anhand entsprechender Daten kritisch beurteilt werden.

Tatsächlich konnte sowohl auf qPCR- als auch Array-Ebene ein solcher Einfluss erkannt werden, wobei CpG-reiche Regionen in +WGA-Proben vergleichsweise unterrepräsentiert waren. Die qPCR-Daten liefern dazu weiteren Aufschluss. Dabei war aufgrund der verringerten Amplifikation CpG-reicher Regionen in Präzipitaten eine im Vergleich zur Input-DNA stärkere Signalverschiebung zu verzeichnen. Dies ist vermutlich auf die unterschiedliche Zusammensetzung dieser Fraktionen zurückzuführen. Da sich die spezifische Anreicherung, sowohl bei Evaluierung der Array-Daten als auch der qPCRs, aus dem Verhältnis von IP zu Input ergibt, führt dies zu der beobachteten verringerten Anreicherung methylierter Regionen.

Bei Evaluierung der MeDIP-Chip-Daten im Vergleich zu quantitativen Methylierungswerten aus Bisulfit-Sequenzierung war für +WGA-Proben dann eine minderwertigere Erfassung methylierter Regionen zu verzeichnen. Ein ähnlicher Sensitivitätsverlust, aufgrund der komprimierten Signale bzw. Dynamik, hinsichtlich der Identifikation differentiell methylierter Regionen ist anzunehmen und wurde in der Literatur bereits angedeutet (Johnson et al. 2008, Robinson et al. 2010).

Dennoch wurden bei Verwendung der WGA sinnvolle Resultate erhalten, wobei biologische Replikate höchst reproduzierbare Daten lieferten. Unter Vermeidung experimenteller Schwankungen kann die WGA also im Falle reduzierten Probenmaterials eine annehmbare Alternative darstellen, wobei ein gewisser Informationsverlust akzeptiert werden muss.

Im Falle muriner MeDIP-Studien ist jedoch unbedingt die Verwendung nicht-amplifizierten Materials vorzuziehen. Aufgrund des bereits diskutierten abweichenden Präzipitationsverhaltens muriner DNA führte die WGA-abhängige Verzerrung zu nicht-tolerierbar niedrigen Anreicherungswerten. Jüngst wurde die erfolgreiche Präzipitation methylierter DNA von einzelnen/wenigen Zellen beschrieben (Borgel et al. 2010). Falls die anschließende Prozessierung (Labeling, Hybridisierung etc.) wie in der vorliegenden Arbeit nicht selbst vorgenommen wird, ergibt sich durch entsprechende Vorgaben des Dienstleisters also hauptsächlich eine Limitierung auf Ebene der Microarrays. Eine Anpassung der Protokolle zur Prozessierung geringerer Probenmengen von Seiten der Dienstleister wäre daher sinnvoll. Alternativ könnte eine Hochdurchsatz-Sequenzierung der

Präzipitate (MeDIP-Sequenzierung) in Erwägung gezogen werden, wobei es jedoch eher zu einer Verlagerung des Problems kommt. Hierbei werden CpG-arme Regionen aufgrund der geringen Anreicherung vergleichsweise schlecht abgedeckt (Robinson et al. 2010).

4.1.5 Identifikation von DMRs und Sensitivität der MeDIP-Chip-Pipeline

Unlängst gewannen Nicht-Promotor-Regionen bei der Analyse von DNA-Methylierungsmustern an Bedeutung (Suzuki 2008, Zilberman 2007, Ball 2009, Dindot 2009, Irizarry 2009). Gewebe- und krankheitsspezifische Änderungen der Methylierung wurden dabei nicht nur mit CpG-armen Ausläuferregionen von Promotor-CpG-Islands (CpG-I-Shores) sondern auch mit Genkörpern in Verbindung gebracht (Dindot et al. 2009, Ball et al. 2009, Irizarry et al. 2009, Doi et al. 2009, Suzuki und Bird 2008). Auch in der vorliegenden Arbeit war im Zusammenhang mit der in Abschnitt 3.2 dargestellten Kryopreservations-Studie eine Erfassung chromosomenweiter Methylierungsmuster von Interesse. In Ermangelung verfügbarer bioinformatischer Anwendungen wurde eine Strategie zur Analyse chromosomenweiter Anreicherungsprofile erarbeitet. Differentiell methylierte Regionen wurden dabei mit Hilfe eines auf Basis der $\Delta\log_2$ -Werte verwendeten Peak-Algorithmus in Kombination mit statistischen Analysen bzw. verschiedenen Filter-Kriterien identifiziert. In einer Evaluation solcher Regionen in Abhängigkeit quantitativer Daten für >260 Promotoren (Zhang et al. 2009) konnte die Validität dieses Ansatzes gezeigt werden. Dabei war für Methylierungsunterschiede >20% zunehmend eine verbesserte Analyse deutlich. Anhand der ermittelten moderaten AUC (AUC(20%): 0,74) ist jedoch ersichtlich, dass auch eine Reihe falsch-positiver und falsch-negativer Resultate erhalten werden. Diese Ergebnisse sind zwar z.T. auch von den zur ROC-Analyse verwendeten Kriterien abhängig (Überschneidungsbereich der Promotor-Region und DMR etc). Dennoch könnte eine gezielte Überprüfung der zur DMR-Identifikation (und -Evaluation) verwendeten Parameter (Größe des Fensters, Schwellenwerte für signifikante Anreicherung bzw. Anreicherungsunterschiede etc.) zu einer Verbesserung der Sensitivität und Spezifität führen. Des Weiteren könnten alternative Peak-Detektions-Algorithmen getestet werden. Qualitative Unterschiede konnten diesbezüglich in anderen Studien bei der Identifizierung angereicherter ChIP-Chip Regionen festgestellt werden, wobei sich z.B. das Programm TAMALg in Bezug auf Nimblegen Arrays als besonders geeignet erwies (Johnson et al. 2008). Ferner wurden im Zuge einiger MeDIP-Studien ähnliche Ansätze speziell für die Auswertung von Methylierungsdaten entwickelt (Dindot et al. 2009, Penterman et al. 2007,

Irizarry et al. 2009). Eine Bereitstellung und Standardisierung solcher Algorithmen könnte die Datenauswertung weiter verbessern.

4.1.6 Kritische Betrachtung der Einbeziehung von Sequenzinformationen in die Daten-Evaluation

Bei der im vorangehenden Abschnitt diskutierten Güte der Datenauswertung ist zu berücksichtigen, dass dabei lediglich qualitative Daten zugrunde liegen. Die Genauigkeit bzw. das Auflösungsvermögen derartiger bioinformatischer Analysen wird durch die Immunopräzipitation selbst bestimmt, welche von der Anzahl und Verteilung verfügbarer Bindestellen abhängig ist. Diese nicht-lineare Beziehung zwischen DNA-Methylierung CpG-Dichte und gemessenem Signal konnte in Abschnitt 3.1.1 verdeutlicht werden.

Verschiedene Studien befassen sich bereits mit einer mehr oder weniger komplexen Einbeziehung dieser Sequenzinformationen um verbesserte und im besten Falle quantitative Resultate zu erhalten (Mohn et al. 2008, Weber et al. 2007, Yamashita et al. 2009, Down et al. 2008, Pelizzola et al. 2008). Einfachere Ansätze beinhalten lediglich die Klassifizierung funktioneller Regionen wie Promotoren gemäß den Eigenschaften überlappender CpG-Inseln, um dann Schwellenwerte hinsichtlich der Anreicherung anzupassen (Mohn et al. 2008, Weber et al. 2007). Dem gegenüber stehen komplexere Algorithmen, welche durch Modellierung der CpG-abhängigen Präzipitation eine Quantifizierung der MeDIP-Daten anstreben (Yamashita et al. 2009, Pelizzola et al. 2008, Down et al. 2008). Solche Ansätze sind allerdings kritisch zu überprüfen. Die Beziehung zwischen DNA-Methylierung, CpG-Dichte und Anreicherung ist komplex und die zur Modellrechnung zugrunde liegenden Annahmen treffen evtl. nicht zwangsläufig auf das gesamte Datenset zu.

Die bekannteste Anwendung, Batman („*Bayesian tool for methylation analysis*“) (Down et al. 2008) wurde, mit Hilfe des Studenten *Pavlo Lutsik*, anhand verschiedener Datensets getestet und durch Korrelation der ermittelten Werte mit Methylierungsdaten aus Bisulfit-Sequenzierung (Zhang et al. 2009) evaluiert. Dabei konnte bei Ausgangsdaten hoher Qualität kein positiver Effekt, sondern eine verringerte Korrelation (sowohl mit als auch ohne vorherige Normalisierung) festgestellt werden. Lediglich Ausgangsdaten mit Signalverzerrungen konnten leicht höhere Korrelationen aufweisen. Interessanterweise galt dies insbesondere in Kombination mit der zuvor diskutierten LOWESS-Normalisierung, welche für sich allein verringerte Korrelationen bedingt. Insgesamt wurde der Eindruck gewonnen, dass die Verwendung von Batman, v.a. bei guter Datenqualität, eher zu einer

Verfälschung als Verbesserung der MeDIP-Daten führt. Dies könnte aufgrund der von Seiten des Programmierers unzureichend dokumentierten Parameter auf eine suboptimale Anwendung zurückzuführen sein. Zudem könnten einige zur Modellrechnung vorausgesetzte Kriterien (spezifische Fragmentverteilung, lineare Beziehung CpG-arter Regionen) nicht ausreichend zutreffen.

Interessant erscheinen hingegen Ansätze, welche experimentelle und bioinformatische Bausteine miteinander verbinden, z.B. MEDME („*Modeling experimental data with MeDIP enrichment*“) (Pelizzola et al. 2008). In diesem Fall wird der Einfluss der CpG-Dichte auf Präzipitation und Hybridisierungsverhalten anhand von Referenz-Arrays mit komplett methylierter DNA erfasst, modelliert und zur Quantifizierung verwendet. Da diese Referenz theoretisch für jede Probe durchzuführen ist, stellt dies im Falle verschiedener zu analysierender Zelltypen allerdings eine aufwändige und kostspielige Methode dar.

Der Nutzen derartiger Anwendungen sollte kritisch abgewägt werden. Sind tatsächlich quantitative Daten erwünscht, sollten eher geradewegs andere Methoden herangezogen werden.

4.1.7 Fazit

Wie in der vorliegenden Arbeit gezeigt werden konnte, eignet MeDIP-Chip sich zur Erfassung relativer Methylierungsprofile und dient damit einer ersten Charakterisierung der (behandelten) Zelltypen von Interesse. Die CpG-abhängige Anreicherung (und Amplifikation) des Materials limitiert dabei die Sensitivität bzgl. der Identifikation (differenziell) methylierter Regionen. Alternative bioinformatische Ansätze könnten in diesem Zusammenhang zwar eine erhöhte Genauigkeit nicht jedoch eine tatsächliche Quantifizierbarkeit erreichen. In Abhängigkeit der Fragestellung könnten andere Ansätze vorteilhaft sein. Ähnliche Methoden unter Verwendung Methyl-bindender Proteine (Gebhard et al. 2006a, Gebhard et al. 2006b, Jorgensen et al. 2006, Cross et al. 1994) erfassen vergleichsweise weniger CpG-arme, dafür vermehrt CpG-reiche Regionen (Nair et al. 2011) und erlauben durch eine Fraktionierung des Präzipitats in Abhängigkeit der CpG-Dichte/Methylierung eine gezieltere globale Analyse (Gebhard et al. 2006a, Gebhard et al. 2006b).

4.2 Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryokonservierung

4.2.1 Geringfügige Abweichungen der MeDIP-Anreicherung nach Kryokonservierung reflektieren technische und biologische Schwankungen

Schwerpunkt der vorliegenden Arbeit war die Charakterisierung chromosomenweiter DNA-Methylierungsmuster kryokonservierter HepG2-Zellen zur Evaluation einer potentiellen Stress-vermittelten Epimutagenese.

Neben den erfahrungsgemäß mittels Kryopreservation verursachten Beeinträchtigungen der Zell-Vitalität durch physikalischen und biochemischen Stress sind auch sublethale Effekte auf Ebene epigenetischer Modifikationen denkbar. Vermehrt gebildete Sauerstoffradikale (ROS) (Ahn et al. 2002, Rauen et al. 1999, Niu et al. 2010, Thomson et al. 2009, Fuller et al. 1988) könnten dabei durch Schädigung der DNA bzw. veränderte Signalwege direkten oder indirekten Einfluss auf die Homöostase der DNA-Methylierung zellulärer Systeme nehmen. Ausgehend von Krebsstudien in Verbindung mit oxidativem Stress wäre z.B. eine Hypomethylierung der DNA durch chemische Modifikation der CpG-Dinukleotide bzw. behinderte Interaktion mit Faktoren wie DNMTs zu vermuten (Valinluck und Sowers 2007, Turk et al. 1995, Weitzman et al. 1994, Donkena et al. 2010).

Aufgabe der vorliegenden Arbeit war deshalb eine umfassende Analyse chromosomenweiter DNA-Methylierungsprofile in Zusammenhang mit Kryokonservierung, um mögliche Veränderungen des epigenetischen Gedächtnisses aufzudecken. DMSO-behandelte bzw. kryokonservierte Zellen der hepatischen Karzinomlinie HepG2 wurden nach 48- und 120-stündiger Kultivierung (Erholungsphase), ebenso wie unbehandelte Kontrollen, mittels der zuvor etablierten Methode MeDIP-Chip analysiert. Unter Verwendung eines Array-Designs, welches kontinuierliche Bereiche der Chromosomen 20, 21 und 22 abdeckt, konnten verschiedene funktionelle intra- und intergenische Regionen berücksichtigt werden. Dabei konnten keine signifikanten globalen Veränderungen der DNA-Methylierung verzeichnet werden. Die lediglich geringfügigen Abweichungen der Anreicherung zwischen Kontrollen und DMSO-behandelten bzw. kryokonservierten Proben bewegten sich im Bereich erwarteter zufälliger Fluktuationen und konnten nicht mittels Bisulfit-Sequenzierung validiert werden. Mehrere Anzeichen sprechen dafür, dass es sich hierbei hauptsächlich um biologische und technische Variabilität CpG-arter Regionen handelt, was im Folgenden näher diskutiert werden soll.

Bereits bei oberflächlicher Betrachtung der Daten unter Einbeziehung aller Oligonukleotide wird die erhebliche Ähnlichkeit der HepG2-Proben deutlich. Neben den im paarweisen Vergleich berechneten äußerst hohen Korrelationswerten zwischen Replikaten und Proben (0,94-0,98), besteht der Umstand, dass die hierarchische Clusteranalyse zu keiner systematischen Trennung derselben führt. Die in Abschnitt 3.3 beschriebenen Daten muriner Stammzellen hingegen erlaubten trotz ähnlicher Korrelationen eine Diskriminierung der Zelltypen in Abhängigkeit der biologischen Replikate, welche größere Schwankungen als die hier verwendeten technischen Replikate vermuten lassen. Dies weist bereits darauf hin, dass mögliche Stress-induzierte Effekte sich offensichtlich im Rahmen experimenteller Variabilität bewegen.

In einer detaillierteren Analyse wurden DMSO-behandelte bzw. kryokonservierte Proben jeweils auf Basis der $\Delta\log_2$ -Werte mit unbehandelten HepG2-Zellen verglichen. Die mittels Peak-Detektions-Algorithmus bestimmten Regionen differentieller Anreicherung waren nur ohne vorherige Korrektur der Daten hinsichtlich der zu erwartenden Falsch-Positiv-Rate signifikant. Die Anreicherungsunterschiede waren dabei, insbesondere im Gegensatz zu den im gewebespezifischen Vergleich identifizierten Regionen, von geringer Anzahl, Ausprägung (unter 2-fach) und Größe (durchschnittlich 200 nt). In Anbetracht der mit Hilfe der Positivkontrolle evaluierten Sensitivität des experimentellen und analytischen Ansatzes, könnten potentielle Methylierungsunterschiede demzufolge an der Auflösungsgrenze des Ansatzes (<20%) liegen. Es ist zu bedenken, dass die zu erwartende Dimension der möglichen Stress-induzierten Effekte nicht bekannt ist und im Vergleich zu gewebespezifischen Mustern nur wenige ausgewählte CpG-Dinukleotide betreffen könnte. Interessanterweise lassen die DPs der behandelten Proben tendenziell eine Hypermethylierung in Zusammenhang mit DMSO bzw. Kryokonservierung erkennen. Dies entspricht nicht den vermuteten Auswirkungen oxidativen Stresses. Wie bereits erwähnt, würde eine ROS-vermittelte chemische Modifikation bzw. Schädigung der CpG-Dinukleotide zu einer direkten Demethylierung bzw. Abwesenheit des AK-Substrats führen (z.B. Oxidation zu 5hmC). Möglicherweise sind andere, ROS-unabhängige Mechanismen für diese potentiellen Methylierungsänderungen verantwortlich. Aufgrund der geringen Anzahl betrachteter Regionen könnte diese Tendenz jedoch auch als Zufälligkeit verstanden werden, zumal die Anreicherungsunterschiede der behandelten Proben sich hinsichtlich Anzahl, Größe und Ausmaß im Rahmen der zwischen den Kontrollen *Ktr* und *NKtr* selbst beobachteten Variationen bewegen. Dies deutet eher auf technische und biologische Schwankungen (möglicherweise auch Kultivierungseffekte) statt auf einen tatsächlichen Einfluss abiotischer Stressoren hin.

Diese Annahme wird durch den Umstand unterstützt, dass es sich bei den beobachteten Schwankungen der Anreicherung offensichtlich um stochastische Ereignisse handelt, da in den analysierten Proben unterschiedliche Regionen und Gene betroffen sind. Zudem konnten die DPs in einer Analyse gemeinsamer Sequenzeigenschaften tendenziell CpG-armen Regionen zugeordnet werden. Tatsächlich verursachen solche genomischen Abschnitte bei Immunpräzipitations-basierten Ansätzen bekanntermaßen ein gewisses Rauschen. Dies ist in der Abhängigkeit der Anreicherung von Anzahl und Verteilung verfügbarer Bindestelle begründet.

Zur Überprüfung der beschriebenen Daten wurde eine exemplarische Analyse drei identifizierter DPs mittels quantitativer Bisulfit-Sequenzierung vorgenommen. Zwischen behandelten Proben und Kontrollen wurden dabei Methylierungsunterschiede <10% festgestellt, womit sie im Rahmen der zwischen den Kontrollen *NKtr* und *Ktr* selbst zu verzeichnenden Variabilität lagen. Die im MeDIP-Chip erkennbaren Tendenzen bzgl. Hypo- bzw. Hypermethylierung entsprechender Proben ließen sich nicht verifizieren. Abgesehen von der geringen Anzahl analysierter Sequenzen ($n \leq 16$) lässt sich dies mit der Tatsache begründen, dass die Methylierungsschwankungen tatsächlich unter der Auflösungsgrenze des MeDIP-Chips liegen und somit ungenaue Messwerte liefern. Obgleich die Analyse aufgrund der geringen Anzahl überprüfter Regionen nicht repräsentativ scheint und einige DPs tatsächlich Stress-induzierte Methylierungsänderungen aufweisen könnten, unterstützen die Resultate generell die Hypothese, dass es sich bei den identifizierten DPs hauptsächlich um technische und biologische Fluktuationen unabhängig thermalen oder chemischen Stresses handelt. Diese Folgerungen wurden durch weitere Bisulfit-Sequenzierungen repetitiver Elemente (LINE1, HERVK) bekräftigt. Diese genomweite Methylierungsmuster repräsentierende Regionen wiesen stabile Signaturen mit individuellen Variationen bis 6% auf. Eine (schwache) Demethylierung der kryokonservierten Proben (*f48h*, *f120h*) in HERVK-Elementen könnte mit Hilfe weiterer Replikate/Experimente und Kultivierungszeitpunkte überprüft werden, um so einen potentiellen Stress-induzierten Effekt von technischer und biologischer Variabilität abzugrenzen.

Unter Verwendung solch sensitiver Methoden könnten potentielle Stress-induzierte Methylierungsänderungen mit einer höheren Auflösung genomweit untersucht werden. Zusätzliche DPs könnten validiert und möglicherweise weitere Regionen, welche aufgrund der Auflösungsgrenze des MeDIP-Chips bisher nicht erfasst wurden, identifiziert werden. Allerdings ist die biologische Signifikanz solch geringer Methylierungsunterschiede, welche

zudem eher intergenische Regionen statt regulatorische Bereiche oder Gene betreffen, fraglich. Durch Erfassung von Transkriptomdaten der HepG2-Proben konnte der Frage eines möglichen Einflusses auf die Genexpression nachgegangen werden.

4.2.2 Änderungen der mRNA-Expression nach Kryokonservierung zeigen keinen Zusammenhang mit DNA-Methylierung

Basierend auf bereits veröffentlichten Daten hinsichtlich Expressionsänderungen in Zusammenhang mit Hypothermie bzw. Kryopreservation (Liu et al. 2000, Sonna et al. 2006, Sonna et al. 2010) und zur weiteren Erhärtung unserer Hypothese, wurde im Zuge der hier beschriebenen Studie zusätzlich die mRNA-Expression untersucht. Statistisch signifikante, z.T. zeitabhängige, Änderungen der Expression konnten sowohl in Zusammenhang mit DMSO-Behandlung als auch Kryopreservation festgestellt werden. Dabei war allerdings keine Überschneidung mit differentiell angereicherten Genen des MeDIP-Chips zu verzeichnen. Ausgehend von der exemplarischen Analyse der Gene der Chromosomen 20-22 scheint somit kein Stress-induzierter Effekt der DNA-Methylierung auf die mRNA-Expression, oder umgekehrt, zu bestehen. Dies unterstützt die oben erläuterte Annahme, dass die im MeDIP-Chip detektierten Anreicherungsunterschiede lediglich technische und biologische Fluktuationen mit geringer Signifikanz repräsentieren.

Dennoch wäre ein Effekt auf einzelne Gene, welche in der vorliegenden Methylierungsstudie nicht berücksichtigt wurden, denkbar. Eine zusätzliche exemplarische Analyse vier differentiell exprimierter Gene mittels Bisulfit-Sequenzierung (CEBPA, FOS, IGFBP1, SPRY4) welche nur individuelle Variationen unter 6% aufzeigte, scheint dies jedoch nicht zu indizieren. Ferner könnte die aufgrund abiotischen Stresses veränderte Expression direkt oder indirekt die DNA-Methylierung ausgewählter Gene beeinflussen. Im Zuge einer Krebsstudie an hepatozellulären Karzinomzellen (HCC) wurde beispielsweise eine ROS-induzierte Überexpression von *Snail* mit einer Rekrutierung von Histon-Deacetylase 1 und DNMT1 in Verbindung gebracht, was zu einer Hypermethylierung des *E-Cadherin*-Promotors führte (Lim et al. 2008).

Dessen ungeachtet scheinen die beobachteten Expressionsänderungen nach DMSO-Behandlung bzw. Kryopreservation eher mit anderen regulatorischen Prozessen als DNA-Methylierung assoziiert zu sein. Bei systematischer Analyse der betroffenen Gene hinsichtlich biologischer Prozesse konnte eine Anreicherung konventioneller Signalwege zur Regulation der Zellregeneration und der Zellerhaltung festgestellt werden. Dies legt nahe, dass die Änderungen der Expression Teil der innewohnenden Stress-Antwort einer Zelle

sind, was angesichts der physikalischen und biochemischen Beeinträchtigung einer solchen Behandlung auch zu erwarten ist. Eine Deregulation verschiedener Stress-Gene wurde bereits mehrfach im Zusammenhang mit moderater Hypothermie und Kryopreservation berichtet. Die Änderungen traten dabei überwiegend augenblicklich nach thermalem Stress bzw. während des Wieder-Erwärmens auf, wobei basale Expressionslevel nach einer gewissen Kultivierungsperiode erreicht wurden (Liu et al. 2000, Sonna et al. 2010, Underhill und Smales 2007). Während diese Studien jedoch eine Über-Expression typischer Faktoren wie HSPs, CIRP, RBM3 etc. berichten, wiesen diese Gene in der vorliegenden Arbeit keine signifikanten Expressionsänderungen auf. Stattdessen war eine verminderte Expression anderer „Stress-reaktiver“ Gene zu verzeichnen. Um diese Diskrepanz zu erklären, wären Daten zu früheren Zeitpunkten der Erholungsphase nach dem Auftauen nötig - obgleich ein detailliertes Verständnis der Natur und Kinetik betroffener Prozesse und Signalwege nicht Ziel dieser Arbeit ist. Möglicherweise spiegeln die Zellen ein eher spätes Stadium der Regeneration wider, in welchem Expressionslevel prominenter Kandidaten wie HSPs bereits wieder angepasst werden und basale Level erreichen. Gene, welche scheinbar eine transient verminderte Expression aufweisen bevor basale Level bei 120h erreicht werden (z.B. Stress-Gene des MAPK-Signalwegs: JUN, FOS, DUSP), könnten nach anfänglicher Stress-induzierter Aktivierung übermäßig reguliert worden sein. Dies würde, angesichts des Fehlens früherer „*post-thawing*“ Zeitpunkte, fälschlicherweise eine transiente Herunter-Regulierung andeuten. Ähnlich unverhältnismäßige De-Regulationen wurden z.B. bzgl. der zellulären Konzentration verschiedener Stress-Proteine nach Kryopreservation von Fibroblasten beobachtet (Liu et al. 2000).

Nicht ganz eindeutig ist die Natur der erst gegen Ende der Erholungsphase (120h) detektierbaren, erheblichen Änderungen der Expression. Ein Zusammenhang mit abiotischem Stress erscheint aufgrund der Anreicherung von Haushalts-Prozessen wie Metabolismus und Translation sowie des stark verzögerten Auftretens unwahrscheinlich (sowohl direkte Einflüsse auf die Signalkaskaden als auch indirekte Effekte über veränderte Chromatin-Strukturen würden eher zu einer unmittelbaren Antwort führen). Es scheint sich hierbei eher um Kultivierungseffekte zu handeln. Möglicherweise befinden sich die Zellen nach Behandlung und Erholungsphase in einem mehr oder weniger synchronisierten Zustand bzw. einem anderen Abschnitt des Zellzyklus als die mit ihnen verglichenen Kontrollzellen.

4.2.3 Fazit

Die vorliegende Arbeit gibt einen Einblick in die Auswirkungen der Kryopreservation an einem Zellkultur-Modell. Die dargestellten Resultate indizieren keine gravierenden Langzeit-Effekte in Zusammenhang mit DNA-Methylierung und Expression nachfolgend Kryokonservierung. Weder der Prozess des Einfrierens und Auftauens noch das toxische Kryoprotektivum DMSO scheinen DNA-Methylierungsmuster signifikant zu beeinflussen. Dies scheint angesichts früher publizierter Daten zur Wechselbeziehung von ROS und DNA-Methylierung, zu Kälteschock-induzierter Demethylierung in Pflanzen sowie zu Effekten in Folge ausgedehnter DMSO-Behandlungen erstaunlich (Valinluck und Sowers 2007, Valinluck et al. 2004, Turk et al. 1995, Weitzman et al. 1994, Kaity et al. 2008, Peredo et al. 2008, Steward et al. 2002, Steward et al. 2000, Iwatani et al. 2006). Weitere Analysen unter Berücksichtigung der folgenden Gesichtspunkte könnten zum Aufschluss dieser unerwarteten Resultate beitragen. Zum Einen wäre eine Evaluierung epigenetischer Muster hinsichtlich anderer zeitlicher Messpunkte sowie längerer Preservation-Perioden von Interesse. Entsprechende Effekte der hauptsächlich zur Langzeitlagerung lebenden biologischen Materials entwickelten Methode zeichnen sich möglicherweise erst im späteren Verlauf der Konservierung ab. Es sei auch zu bedenken, dass andere Gewebe- oder Zelltypen (z.B. Stammzellen oder frühe Embryonen) eine höhere epigenetische Sensitivität auf abiotische Stressoren aufweisen könnten. Murine Keimzellen und Embryonen zeigten, wie bereits erwähnt, in einigen Studien Hinweise auf eine veränderte Methylierung (Milroy et al. 2011, Wang et al. 2010). Dies sollte mittels chromosomenweiter Ansätze systematisch analysiert werden. Des Weiteren könnten DMSO-Behandlung bzw. Kryopreservation eine andere epigenetische Ebene wie Histon-Modifikationen beeinflussen. Diese wurden in Zusammenhang mit verschiedenen Stressoren wie Hypoxie, DMSO und Hypothermie bereits als reaktiv erachtet (Suo et al. 2010, Yan et al. 2010). Eine detaillierte Analyse ausgewählter, die Genexpression beeinflussender Histonmodifikationen wie H3K4, H3K9, H3K27 z.B. mittels ChIP-Chip könnte eine mögliche Epimutagenese in HepG2-Zellen klären und zur weiteren Interpretation der dargestellten Transkriptomdaten beitragen.

Letztendlich wäre es denkbar, dass in der hier getesteten Zelllinie potentielle Stress- bzw. ROS-vermittelte Schäden vor einer Manifestation bereits repariert wurden. Es gibt Hinweise auf eine innerhalb weniger Stunden nach Kryokonservierung erfolgende Reparatur von Doppelstrangbrüchen in Oozyten (Peng et al. 2008, Trapphoff et al. 2010). Auch vermutlich entstehende oxidierte Basen wie 5-Hydroxymethylcytosin könnten, falls keine weitere Modifikationen wie Deaminierung und Carboxylierung auftreten, mittels z.B.

Basen-Excisions-Reparatur vor dem ersten untersuchten Zeitpunkt repariert worden sein. Dies könnte auch eine Art epigenetischer Stress-Antwort unter Beeinflussung des Zellzyklus involvieren. Milutinovic und Mitarbeiter (Milutinovic et al. 2003) berichten im Zuge einer Krebsstudie von einem intra-S-Phasen-Arrest bei DNMT1-Knockout, wodurch eine Manifestation fehlerhafter DNA-Methylierung verhindert wird.

Insgesamt legt die vorliegende Arbeit den Schluss nahe, dass die Kryopreservation biologischen Materials zu keiner gravierenden Beeinflussung der DNA-Methylierung führt. Potentielle geringfügige Abweichungen scheinen, da kein Einfluss auf die Genexpression festgestellt werden konnte, nicht zu Langzeiteffekten zu führen.

4.3 Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen

In den vergangenen Jahren konnten Stammzelllinien unterschiedlichen ontogenetischen Alters und Differenzierungspotentials, z.T. auch durch Reprogrammierung und Transdifferenzierung, generiert werden. Dies bietet nicht nur Ansatzmöglichkeiten im Bereich der medizinischen Forschung sondern trägt insbesondere zum Grundverständnis molekularer Mechanismen bei, welche an Pluripotenz, früher Differenzierung und Reprogrammierung beteiligt sind. In diesem Zusammenhang sind auch dynamische Änderungen epigenetischer Signaturen von Bedeutung.

Zur näheren Charakterisierung ausgewählter Stammzelltypen sollten in der vorliegenden Arbeit deren genomweite DNA-Methylierungsprofile analysiert werden. Dabei waren zum Einen durch Reprogrammierung von Keimbahnstammzellen generierte ESC-ähnliche Zellen (Ko et al. 2009), zum Anderen die Stadien der vorbereiteten und naiven Pluripotenz (Nichols und Smith 2009) von Interesse. Die vergleichenden Analysen der Promotor-Methylierung mittels MeDIP-Chip konnten dabei zum Verständnis epigenetischer Signaturen der zugrundeliegenden Entwicklungsstadien beitragen. Zudem werden in der folgenden Diskussion durch Einbeziehung zur Verfügung gestellter mRNA-Expressionsdaten interessante Fragen aufgeworfen, welche Ansatzpunkte für tiefere Untersuchungen bieten.

4.3.1 Die MeDIP-Chip-Daten erlauben eine sensitive Erfassung von Methylierungsunterschieden im Promotor-Bereich

Zur Analyse genomweiter Methylierungsmuster wurde die zuvor beschriebene Methode MeDIP in Kombination mit Promotor-Arrays herangezogen. Nach Qualitätskontrolle und Prä-Prozessierung der Daten wurden die mittleren Anreicherungsdaten der Kern- und Aufwärtspromotoren (± 500 bp um TSS bzw. -500 bis -1000 bp) berechnet und zur Ermittlung differentiell methylierter Regionen verwendet.

Durch Hochdurchsatz-Bisulfit-Sequenzierung neun ausgewählter Promotor-Regionen konnte die Validität der erhobenen Daten näher überprüft werden. Dabei wurde eine qualitative Übereinstimmung der Methylierungs- und Anreicherungsdaten hinsichtlich zelltypspezifischer Muster verzeichnet. Regionen differentieller Methylierung wurden zudem mit hoher Sensitivität erfasst. Die MeDIP-Chip-Daten scheinen somit Methylierungsprofile ausreichend zu repräsentieren und bieten daher eine solide und verlässliche Grundlage für vergleichende Analysen und Interpretationen.

Durch die hier erreichte Sensitivität konnten zum Teil auch Methylierungsunterschiede $<30\%$ identifiziert werden. Solch geringe Differenzen haben unter Umständen keine biologische Signifikanz und spiegeln z.B. lediglich biologische Variabilität wieder. Dies führt möglicherweise zu späteren Fehlinterpretationen. Aufgrund der bereits erläuterten Abhängigkeit der AK-basierten Präzisierung von der CpG-Dichte, welche keine Quantifizierbarkeit der Anreicherungsdaten erlaubt, stellt das Ansetzen strengerer Schwellenwerte eine gewisse Herausforderung dar. In dieser Hinsicht kann eine Einbeziehung der Sequenzinformation hilfreich sein, wie z.B. der Ansatz von Weber und Mohn (Mohn et al. 2008, Weber et al. 2007). Dabei werden Promotoren mit niedrigem, mittlerem und hohem CpG-Gehalt unterschieden (LCP, ICP, HCP) und Schwellenwerte einer Hypermethylierung/*de novo* Methylierung entsprechend festgelegt. Dies erfordert allerdings einen erhöhten bioinformatischen Aufwand sowie eine ausreichende Validierung (Bisulfit-Sequenzierung) zur Anpassung der Parameter in Abhängigkeit der Prä-Prozessierung/Datendynamik. Eine entsprechende Umsetzung könnte jedoch zu einer verfeinerten Analyse des in dieser Arbeit vorgestellten Datensets beitragen.

4.3.2 Distinkte Methylierungsmuster der Aufwärts- und Kern-Promotoren

Unabhängig der im Folgenden zu betrachtenden Fragestellungen wurden die Methylierungsdaten aller analysierten Zelltypen verglichen. Aufwärts- und Kern-

Promotoren wiesen dabei unabhängig des Zelltyps oder Differenzierungspotentials distinkte Methylierungsmuster auf. Diese waren weder in pluripotenten noch terminal differenzierten Zellen durch erhöhte Hypo- oder Hypermethylierung einer der Promotorklassen gekennzeichnet. Die spezifische Gruppierung der Zelltypen untereinander, welche deren generelle Beziehung andeutet, konnte dabei in beiden Promotor-Klassen erkannt werden (siehe Abbildung 3.3-3). Offenbar sind Aufwärts- und Kern-Promotoren für die zelltypspezifischen Methylierungsprofile verantwortlich, wobei jeweils unterschiedliche Gengruppen markiert werden. Dementsprechend ist in terminal differenzierten MEFs Hypermethylierung im Vergleich zu pluripotenten Stammzellen (ESC und EpisC) sowohl in Kern- als auch Aufwärts-Promotoren mit Genen der Embryonalentwicklung (Regionalisierung, Morphogenese) assoziiert. Dennoch betrifft diese Hypermethylierung tendenziell eher Aufwärtspromotoren (Abbildung 6.3-2) und reflektiert somit die bereits bekannte Tatsache, dass spezifische Methylierungsereignisse während der Differenzierung offenbar vornehmlich in distalen Regionen, >1 kb aufwärts des Promotors, auftreten (Meissner et al. 2008).

4.3.3 Genomweite Methylierungs- und Expressionsmuster der pluripotenten gPS heben sich deutlich von den keimbahnspezifischen Profilen der GSC

Die allein durch entsprechende Kultivierung spontan zur Pluripotenz zurückgekehrten gPS stellen ein interessantes Modell zum Verständnis der Neusetzung epigenetischer Signaturen während dieser Reprogrammierung dar und könnten des Weiteren von therapeutischem Nutzen sein. Daher war eine Charakterisierung dieser Zellen auf Ebene der DNA-Methylierung im Vergleich zu ihren keimbahnspezifischen „Ausgangszellen“ und pluripotenten ESC von besonderem Interesse.

Die Resultate der MeDIP-Chip-Analyse wiesen auf eine genomweite Reprogrammierung der Promotor-Methylierung in gPS zu ESC-ähnlichen Mustern hin. Im Gegensatz dazu wiesen sie deutliche Unterschiede zu unipotenten GSC auf. Keimbahnspezifische Gene (Spermatogenese) waren demzufolge, wie zu erwarten, in gPS und ESC hypermethyliert. Trotz der genomweit ähnlichen Anreicherungsmuster der gPS und ESC waren diese nicht identisch. 648 DMRs ließen dabei hauptsächlich eine Hypomethylierung aufwärts-gelegener Promotor-Regionen in gPS erkennen. Diese zeigten sowohl keimbahnspezifische Muster (GSC=gPS), u.a. in imprinteten Genen, als auch gPS-spezifische Muster, welche mit Phosphorylierung und Gewebe-Entwicklung assoziiert waren.

Die Resultate lassen daher eine zum Teil fehlerhafte, nämlich unvollständige bzw. aberrante, Reprogrammierung vermuten, welche auch bereits in anderen Reprogrammierungssystemen erkannt wurde. Die aus somatischen Zellen generierten iPSC (Mensch und Maus) sind diesbezüglich schon recht gut charakterisiert und neben genetischen Aberrationen wurde mehrfach über epigenetische Änderungen berichtet. Verschiedene Studien konnten dabei unterschiedlich ausgeprägte Abweichungen des Methylierungsmusters in iPSC gegenüber ESC aufzeigen (Kim et al. 2010, Mikkelsen et al. 2008, Nishino et al. 2011, Ohi et al. 2011, Doi et al. 2009, Lister et al. 2011). Obwohl diese zum Teil konsistent in mehreren iPSC-Linien aufzutreten scheinen, wird mitunter der stochastische Charakter solcher Änderungen betont (Lister et al. 2011, Nishino et al. 2011). Wie auch in der vorliegenden Arbeit, im Falle der gPS, können bei iPSC sowohl unvollständig reprogrammierte Methylierungsmuster der „Mutterzellen“ als auch aberrante Muster spezifisch für die reprogrammierte Zelle unterschieden werden (epigenetisches Gedächtnis vs. epigenetische Mutation) (Lister et al. 2011).

Die Ursachen solch potentiell fehlerhafter Reprogrammierung auf Ebene der DNA-Methylierung sind unklar. Die Mechanismen der direkten Reprogrammierung, insbesondere ohne Einsatz von Faktoren wie im Falle der gPS, ist noch nicht gut verstanden, so dass an diesem Punkt lediglich spekuliert werden kann. Möglicherweise führt eine gestörte Balance von Methylierungs- und Demethylierungs-Ereignissen zu einem fehlerhaften Muster. i) Da hauptsächlich in gPS hypomethylierte Regionen vorkommen, wäre zum Einen eine beeinträchtigte oder verzögerte Aktivität der Methylierungsmaschinerie denkbar (Abbildung 4.3-1 (B)). In diesem Zusammenhang wurde bereits eine ineffiziente Rekrutierung entsprechender Faktoren zu „abgelegenen“ Genen vermutet, welche in iPSC z.T. von unvollständiger Reprogrammierung betroffen sind (Ohi et al. 2011). Auch könnte eine fehlerhafte Bindung von unmethylierten Promotoren durch in gPS erneut exprimiertes Tet1 eine Rolle spielen. Studien deuten darauf hin, dass Tet1 in ESC die Hypomethylierung von Promotor-CGIs durch Blockierung der Methylierungsmaschinerie erhalten (Xu et al. 2011, Wu et al. 2011b, Wu et al. 2011a, Williams et al. 2011). ii) Auf der anderen Seite könnte, z.B. im Falle der gPS-spezifischen Hypomethylierung, eine Überaktivität oder fehlende Spezifität der Demethylierungsereignisse vermutet werden (Abbildung 4.3-1 (C)). Falls die in gPS auftretende Reprogrammierung aktive Demethylierungs-Mechanismen involviert, wäre eine unspezifische Modifikation (Hydroxylierung, Carboxylierung) bzw. Demethylierung einzelner Regionen mittels Tet-Enzymen möglich. iii) Angesichts des komplexen epigenetischen Codes mit gegenseitiger Abhängigkeit verschiedener Komponenten scheint auch ein indirekter Effekt durch fehlerhafte Histonmodifikationen

oder Histon-modifizierende Enzyme denkbar (Abbildung 4.3-1 (B) und (C)). Dadurch könnte eine *de novo* Methylierung in gPS behindert bzw. der Demethylierungsmaschinerie Zugang gewährt werden.

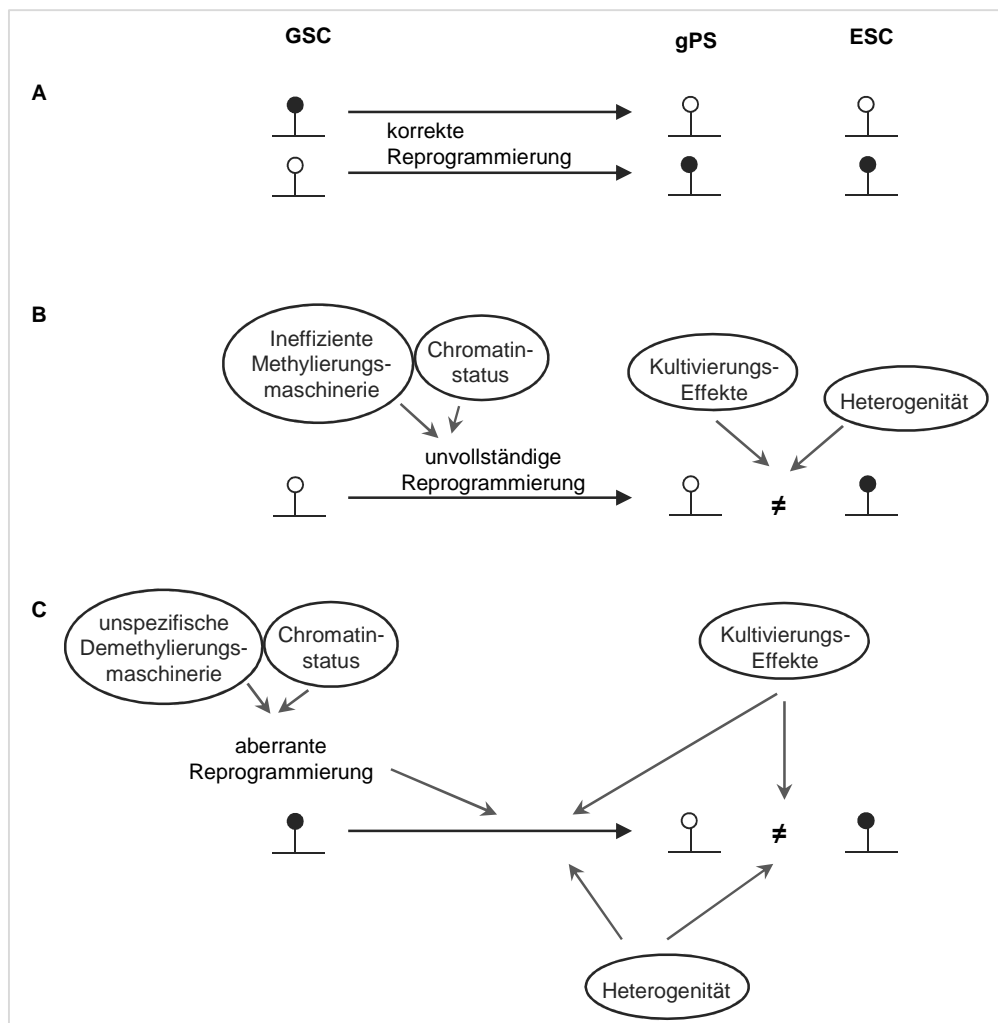


Abbildung 4.3-1: Modell zu möglichen Ursachen der in gPS beobachteten Hypomethylierung. (A) schematische Darstellung einer korrekten Reprogrammierung von GSC zu ESC-ähnlichen DNA-Methylierungsmustern in gPS. (B) und (C) schematische Darstellung der Hypomethylierung in gPS als Folge unvollständiger bzw. aberranter Reprogrammierung oder als experimentelles Artefakt aufgrund genetischer und Kultivierungs-bedingter Einflüsse. ● und ○: hyper- bzw. hypomethylierte Region

Abgesehen von einem tatsächlichen Zusammenhang mit Reprogrammierungs-Ereignissen könnten die beobachteten Methylierungsdifferenzen z.T. auch auf bereits bekannte Einflüsse wie Kultivierungseffekte zurückzuführen sein (Borgel et al. 2010, Pannetier und Feil 2007) (Abbildung 4.3-1 (B) und (C)). Eine Akkumulation von Promotor-Methylierung, wie sie diesbezüglich in ESC festgestellt wurde (Borgel et al. 2010), könnte in der vorliegenden Arbeit eine vermeintlich aberrante Hypomethylierung der gPS im Vergleich zu

den als Referenz erachteten ESC aufzeigen. Des Weiteren wurden in ESC und EpiSC Subpopulationen von Zellen identifiziert, welche sich z.T. in Expression, Chromatin-Modifikationen und DNA-Methylierung unterscheiden (Chambers et al. 2007, Hayashi et al. 2008, Toyooka et al. 2008, Han et al. 2010). Diese heterogene Natur von Stammzelllinien selbst - oder eine Heterogenität bzgl. der Reprogrammierungseffizienz - wobei nur ein gewisser Anteil der gPS-Zellen fehlerhafte Methylierungsmuster aufweisen würde - sollte berücksichtigt werden. Die in der Bisulfit-Analyse erkannten Subpopulationen von Sequenzen hinsichtlich ihrer Methylierungsmuster könnten ein Hinweis darauf sein (z.B. *Cacng5*, *Zfp263*, Anhang C, Abbildung 6.3-3). Interessant wäre auch eine Analyse unterschiedlicher gPS-Passagen, da im Falle der iPSC über z.T. transiente Methylierungsunterschiede berichtet wurde, welche mit längerer Kultivierung verschwinden (Nishino et al. 2011).

Um mögliche funktionelle Auswirkungen der DNA-Methylierungsunterschiede zwischen gPS und ESC abzuschätzen, wurden die Resultate mit zur Verfügung gestellten Expressionsdaten der jeweiligen Zelltypen abgeglichen (*Mouse Genome 430 2.0 Array*, Affymetrix, MPI Münster). Zum untersuchten Zeitpunkt konnte zwar bei 96 Genen eine differentielle Expression zwischen gPS und ESC (>2-fach) festgestellt werden, jedoch war dies nur im Falle von vier Genen mit einer Methylierungsänderung verbunden (basierend auf überlappenden Gensets der verwendeten Arrays) (Abbildung 4.3-2 (A)).

Die in gPS deregulierte Expression ist offenbar eher auf alternative Faktoren zurückzuführen. Angesichts der bekannten Regulationsmechanismen pluripotenter Zellen könnte es sich hierbei um Störungen der TF-Netzwerke und Chromatinstruktur handeln. Die hier betroffenen entwicklungspezifischen Gene (Morphogenese) werden in der Regel mittels bivalenter Domänen reguliert, deren gestörte Ausbildung durch fehlende Etablierung von H3K27me3 in iPSC zu unvollständiger Reprogrammierung führt (Mikkelsen et al. 2008). Auch könnte DNA-Methylierung außerhalb der analysierten Promotorregionen (z.B. in Genkörpern oder LMRs) sowie hmC und Tet1 eine Rolle spielen.

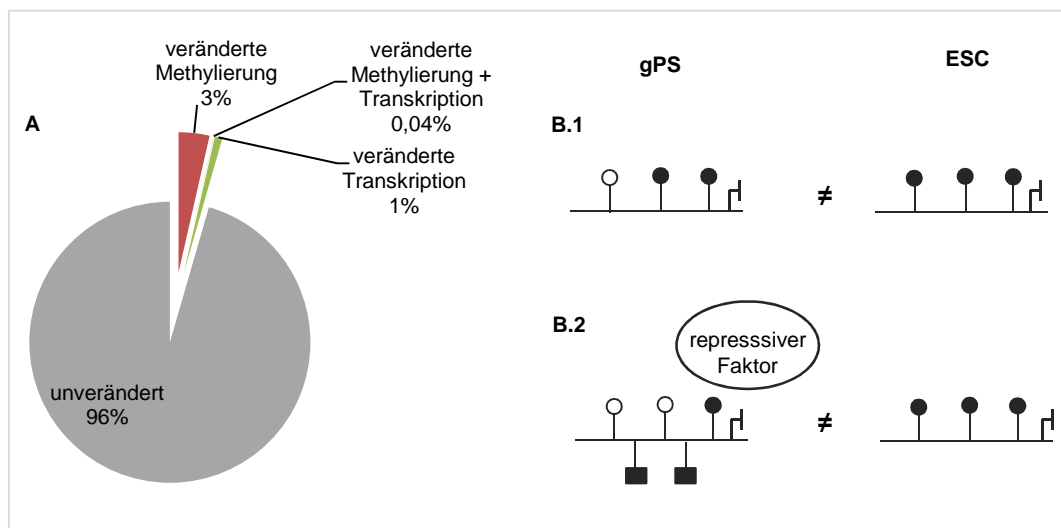


Abbildung 4.3-2: DNA-Methylierung vs. Gen-Expression in gPS und ESC. (A) Änderungen der DNA-Methylierung und mRNA-Expression in gPS treten offenbar unabhängig voneinander auf. (B) Modelle zur Interpretation von Methylierungsunterschieden zwischen gPS und ESC bei gleichbleibender Transkription (schematische Darstellung). Daten basierend auf überlappendem Genset der Methylierungs- und Expressionsarrays (14670 individuelle Gene) (B.1) Geringe Abweichungen der DNA-Methylierung (hoch-sensitive Erfassung) bewirken keine Aktivierung der Transkription. (B.2) Stilllegung hypomethylierter Gene mittels repressiver Chromatinmodifikationen bzw. repressiver Faktoren. ● und ○: hyper- bzw. hypomethyliertes CpG; ■ repressive Chromatinmodifikation

Interessant hingegen ist die Beobachtung, dass umgekehrt die identifizierten Methylierungsänderungen offenbar keinen direkten Effekt auf die Expression ausüben. Möglicherweise sind diese Methylierungsdifferenzen z.T. sehr gering (diskutierte Sensitivität, siehe Abschnitt 4.3.1) bzw. betreffen nur einen gewissen Anteil der Zellen (diskutierte Heterogenität, siehe oben), so dass dies keinen signifikant erfassbaren Einfluss auf die Expression darstellt (Abbildung 4.3-2 (B.1)). Die dargestellten MeDIP-Chip-Daten weisen durch Betrachtung von Kern- und Aufwärtspromotoren sowie alternative TSS zudem eine gewisse Redundanz auf. Interessanterweise waren die Methylierungsänderungen vornehmlich mit Aufwärts-Promotoren assoziiert (Abbildung 6.3-1). Möglicherweise könnten einige der DMRs also vernachlässigt werden, da der (unveränderte) Methylierungsstatus der Kern-Promotoren der entsprechenden Gene, oder aufwärtsgelegene regulatorische Regionen (z.B. LMRs (Stadler et al. 2011)), ausschlaggebend für die Expression ist. Ähnliche Spekulationen wurden bereits in anderen Studien erhoben (Eckhardt et al. 2006). Des Weiteren könnte die Hypomethylierung der gPS teilweise durch z.B. fehlende TFs, repressive Chromatinmodifikationen oder andere Faktoren kompensiert werden (Abbildung 4.3-2 (B.2)). Dabei stellt sich die Frage, wie die Demethylierungsmaschinerie Zugang zu hochkondensierten Targets erlangen kann. Denkbar wäre eine fehlerhafte Ausbildung bivalenter Domänen, wobei 5hmC und Tet1 involviert

sein könnten. Letzteres ist laut aktueller Studien auch durch Interaktion mit dem Co-Repressor Sin3A an einer Histon-Deacetylierung und folgenden Stilllegung von Genen beteiligt (Williams et al. 2011).

Die hier dargestellten Ergebnisse bzgl. der Integration von Methylierungs- und Expressions-Änderungen beruhen lediglich auf einem groben Vergleich. Es ist zu berücksichtigen, dass eine detailliertere Analyse zu einer Verschiebung der Resultate führen könnte. Dabei könnten die Parameter bzw. Schwellenwerte bzgl. der Identifikation differentiell exprimierter Gene kritischer festgelegt werden. Des Weiteren könnte die (hier nicht näher überprüfte) Datenqualität und Prä-Prozessierung eine erhebliche Rolle spielen. Erst kürzlich wurde im Falle der iPSC gezeigt, dass in Abhängigkeit der Datenanalyse stark unterschiedliche Resultate erhalten werden (Chin et al. 2010, Guenther et al. 2010). Nicht zuletzt sollte die Validität der Expressionsdaten entweder mittels Einzelgenanalysen oder unter Verwendung einer anderen Plattform überprüft werden.

Fazit:

In der vorliegenden Arbeit konnte zur Charakterisierung der gPS auf der Ebene epigenetischer Signaturen beigetragen werden. Dabei wurde gezeigt, dass deren DNA-Methylierungsmuster genomweit hin zu ESC-ähnlichen Mustern reprogrammiert wurden. Diese mittels einer Promotor-basierten Analyse erhaltenen Erkenntnisse wurden durch Bisulfit-Hochdurchsatz-Sequenzierung repetitiver Elemente bestätigt (Daten der Doktorandin Julia Arand, Anhang C, Abbildung 6.3-7). Dabei spiegelten gPS, im Gegensatz zu GSC, die typischen pluripotenzspezifischen Muster, auch hinsichtlich der nicht-CpG-Methylierung (CpA in mSat), wider. Eine Reihe differentiell methylierter Regionen zwischen gPS und ESC könnten - neben Kultivierungseffekten und Heterogenität der Zelllinien - auf eine ebenfalls in iPSC auftretende (stochastische) unvollständige und aberrante Reprogrammierung hindeuten. Die Ursachen und funktionellen Auswirkungen dieser fehlerhaften Methylierung sind nicht klar. Offenbar scheint aber kein direkter Zusammenhang mit der Genexpression zu bestehen. Ähnlich wie in iPSC scheinen fehlerhafte Methylierungssignaturen und deregulierte Expression größtenteils durch alternative Mechanismen kompensiert zu werden, so dass Eigenschaften wie Pluripotenz und Differenzierungspotential nicht wesentlich beeinträchtigt werden. Lediglich die Fähigkeit der tetraploiden Komplementation ist nicht gegeben (Ko et al. 2009), wobei der genaue Bezug zu Expression bzw. Methylierung nicht klar ist. Allerdings stellt diese Einschränkung das mögliche therapeutische Potential der Zellen in Frage.

Die Ergebnisse verdeutlichen, dass weitere Analysen zur Charakterisierung der gPS im Vergleich zu ESC (und GSC) nötig sind, um die tatsächliche Zellkompetenz sowie das Ausmaß der epigenetischen Reprogrammierung abschätzen zu können und funktionelle Konsequenzen in Zusammenhang mit der Genexpression aufzudecken. Neben einer kritischen Re-Analyse und Validierung der Expressionsdaten wären Methylierungsdaten zusätzlicher Zelllinien und Kultivierungszeitpunkte, insbesondere aber die Erfassung von Histonmodifikationen (wie H3K4me1, H3K4me3, H3K27me3) nötig.

4.3.4 Die Methylierungsprofile der EpiSC spiegeln das Stadium der eingeschränkten Pluripotenz wider

Stammzellen der Post-Implantations-Entwicklung, EpiSC, stellen ein interessantes Modell zur Analyse der frühen Embryonalentwicklung dar. Zum besseren Verständnis des durch sie repräsentierten Stadiums der „vorbereiteten“ Pluripotenz insbesondere hinsichtlich epigenetischer Signaturen sollten diese Zellen im Vergleich zu naiv-pluripotenten ESC und terminal differenzierten MEF charakterisiert werden.

Die Resultate der MeDIP-Chip-Analyse wiesen genomweit distinkte Methylierungsmuster der EpiSC auf, welche sich bereits deutlich von denen der naiv-pluripotenten ESC unterscheiden und eine ausgeprägte Ähnlichkeit mit differenzierten MEFs erkennen ließen. Unter Berücksichtigung der Aufwärts- und Kernpromotoren gruppierten EpiSC in der hierarchischen Clusteranalyse mit MEF „außerhalb“ der ESC sowie der übrigen analysierten Stammzelllinien. Bei detaillierterer Analyse identifizierter DMRs zwischen den drei Zelltypen konnten in EpiSC Promotor-Regionen mit gewebespezifischen Methylierungsmustern (hauptsächlich hypermethyliert), stammzellspezifischen Methylierungsmustern (hauptsächlich hypomethyliert) und EpiSC-spezifischen Methylierungsmustern (hypo- und hypermethyliert) erkannt werden. Dabei gingen insbesondere Gene der Stammzellerhaltung bzw. frühen und späten Embryonalentwicklung mit Hypomethylierung einher, während Gene, die mit sehr frühen Entwicklungsereignissen in Zusammenhang stehen (Meiose, Dynamik der Ionenkanäle) (Moody 1991, Toshi 2010 für Ionenkanäle), durch stabile reprimierende Markierung der Methylierung gekennzeichnet waren.

Durch Einbeziehung zur Verfügung gestellter Expressionsdaten (*Mouse Genome 430 2.0 Array*, Affymetrix, MPI Münster) werden die dynamischen und komplexen Vorgänge während der frühen Entwicklung deutlich, welche zur Vorbereitung der Zellen und

schließlich zur Differenzierung führen. Basierend auf der Gruppierung der Zelltypen zueinander (Cluster-Analyse) bzw. der Anzahl differentiell exprimierter Gene zwischen EpiSC, ESC und MEF ist die im Gegensatz zur epigenetischen Signatur erhöhte Ähnlichkeit der Stammzellen hinsichtlich ihrer Expressionsmuster auffällig (Abbildung 4.3-3 (A)).

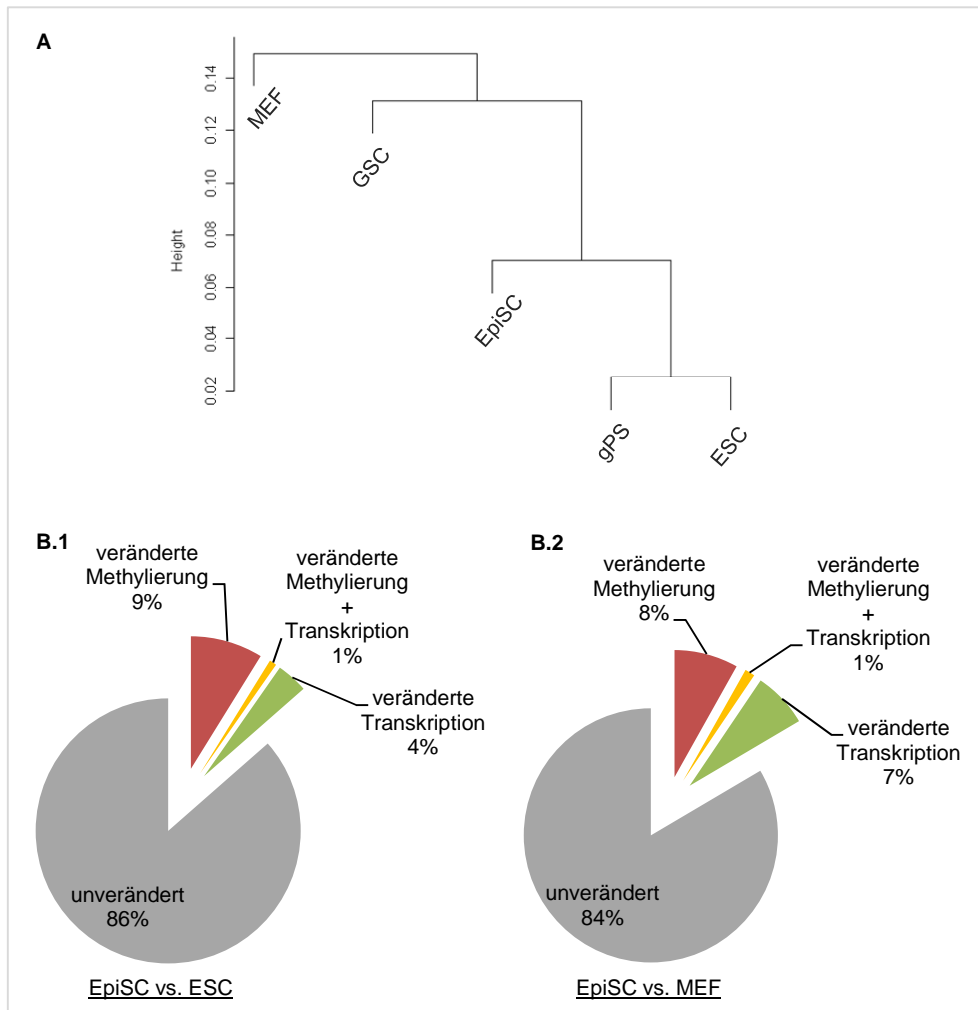


Abbildung 4.3-3: Gen-Expression und DNA-Methylierung in EpiSC, ESC und MEF. (A) Hierarchische Clusteranalyse der Transkriptom-Daten weist hohe Ähnlichkeit der EpiSC mit (naiv) pluripotenten ESC und gPS auf. (B.1, B.2) Änderungen der DNA-Methylierung und Transkription zwischen EpiSC und ESC (B.1) bzw. EpiSC und MEF (B.2) betreffen hauptsächlich unterschiedliche Gene (basierend auf überlappendem Genset der Methylierungs- und Expressionsarrays, 14521 individuelle Gene).

Interessanterweise scheint nur bei einem geringen Anteil differentiell exprimierter Gene ein Zusammenhang mit Änderungen der DNA-Methylierung zu bestehen (ausgehend von überlappenden Gensets der verwendeten Arrays, Abbildung 4.3-3 (B.1) und (B.2)). Hierbei handelt es sich u.a. um Gene der Embryonalentwicklung.

Der Hauptteil der an in EpiSC zu verzeichnenden Expressionsänderungen scheint jedoch mittels DNA-unabhängiger Mechanismen reguliert zu werden. Stamzellenspezifisch unterexprimierte Gene sind u.a. mit der Entwicklung (verschiedene Organe) assoziiert, welche in pluripotenten Zellen bekanntermaßen hauptsächlich über Chromatin-Modifikationen und Chromatin-modifizierende Faktoren reguliert werden (Bernstein et al. 2006, Mikkelsen et al. 2007). Diese Gene könnten also z.T. in bivalenten Domänen liegen. Des Weiteren wurde berichtet dass mit der Entwicklung/Differenzierung einhergehende Änderungen der DNA-Methylierung meist in distalen Regionen (>1kb von TSS) vorkommen (Meissner 2008), welche in der vorliegenden Arbeit durch die Aufwärtspromotoren (-1,5 bis -0,5kb) nur marginal abgedeckt werden. Tatsächlich konnte auch in der vorliegenden Arbeit festgestellt werden, dass entwicklungspezifische Gene tendenziell Methylierungsänderungen in Aufwärtspromotoren zeigten (Abbildung 6.3-2). Des Weiteren könnten diese Gene durch DNA-Methylierungs-unabhängige Mechanismen wie die An-/Abwesenheit spezifischer Transkriptionsfaktoren, die Verwendung alternativer Promotoren, alternatives Spleissen sowie ncRNAs reguliert werden.

Interessant, wenn auch unerwartet, ist die erhebliche Anzahl differentiell methylierter Regionen zwischen den drei Zelltypen, welche nicht mit Expressionsänderungen in Zusammenhang zu stehen scheinen. Hierbei könnte es sich um eine Widerspiegelung dynamischer Regulationen der Embryonalentwicklung durch komplexe Interaktion unterschiedlicher epigenetischer Ebenen (DNA-Methylierung und -Hydroxymethylierung, Chromatin-Modifikationen) mit regulativen Faktoren (TFs, PRCs, etc.) handeln. In Ermanglung entsprechender Daten für EpiSC kann an diesem Punkt - basierend auf bekannten Mechanismen zur Regulation der hier z.T. mit differentieller Methylierung assoziierten entwicklungspezifischen Gene (Bernstein 2006, Mikkelsen 2008, Meissner 2010) - lediglich spekuliert werden. Es wäre denkbar, dass ein Teil dieser DMRs, insbesondere im Falle transkriptionell inaktiver Gene, mit bivalenten Regionen und deren dynamischer Änderungen zu unterschiedlichen Zeitpunkten der Entwicklung assoziiert sind (Abbildung 4.3-4).

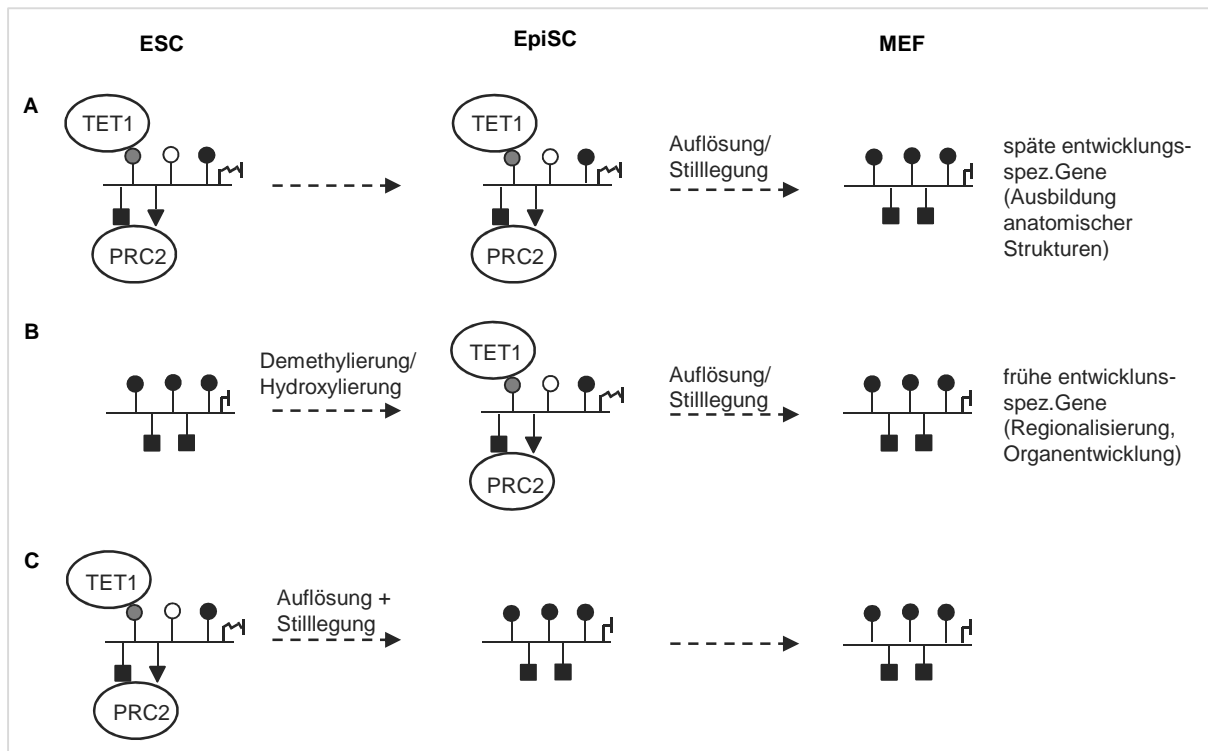


Abbildung 4.3-4: Modelle zur Interpretation der Methylierungsdaten im Kontext der Genexpression mit potentieller Assoziation transkriptionell inaktiver Gene mit bivalenten Domänen (schematische Darstellung). (A) Stammzell-spezifische Hypomethylierung der ESC und EpiSC im Vergleich zu MEF lässt mögliche Auflösung und spätere stabile Stilllegung bivalenter Domänen im fortgeschrittenen Stadium der Entwicklung vermuten. (B) EpiSC-spezifische Hypomethylierung lässt mögliche Ausbildung neuer bivalenter Domänen in EpiSC mittels Demethylierung bzw. Hydroxylierung sowie spätere stabile Stilllegung dieser frühen entwicklungs-spezifischen Gene vermuten. (C) Gewebespezifische Hypermethylierung der EpiSC und MEF im Vergleich zu ESC lässt frühe Auflösung und Stilllegung bivalenter Domänen vermuten. ● und ○: hyper- bzw. hypomethyliertes CpG; ● hydroxyliertes CpG, ■ und ▼: repressive bzw. aktive Chromatinmodifikation

Neben der Auflösung in univalente Regionen und späterer Stilllegung mittels Methylierung (Abbildung 4.3-4 (A) und (C)), wäre - in Zusammenhang mit EpiSC-spezifischer Hypomethylierung - auch eine Bildung neuer Domänen möglich (Abbildung 4.3-4 (B)). Dabei könnte die enzymatische Funktion von Tet1 (Demethylierung/Hydroxylierung), welches gemäß der hier analysierten Daten in EpiSC noch transkribiert wird, eine Rolle spielen. Auch könnte die Expression dieser Gene eher durch aufwärtsgelegene Regionen, wie z.B. distale LMRs (Stadler et al. 2011), oder Genkörper-Methylierung bestimmt werden. Zur Klärung dieser Annahmen wären weitere Analysen erforderlich wie die Erfassung von Methylierungsmustern außerhalb von Promotor-Regionen (z.B. MeDIP-Sequenzierung), die Generierung entsprechender 5hmC-Profile (z.B. hMeDIP, Immunopräzipitation mittels hmC-AK in Kombination mit Microarrays oder Sequenzierung) sowie die Einbeziehung von Histonmodifikationen (ChIP-Chip oder ChIP-Sequenzierung).

Wie bereits in Abschnitt 4.2 diskutiert, können mögliche experimentelle und analytische Artefakte nicht ausgeschlossen werden, welche zu einer Verfälschung der präsentierten Ergebnisse führen. Neben Abweichungen des genetischen Hintergrunds zwischen den analysierten Zelltypen könnten auch heterogene Subpopulationen, sowie Kultivierungsbedingungen vergleichende Analysen beeinflussen. Auch könnte eine detailliertere Analyse der Expressionsdaten selbst - wie bereits in 4.2 erläutert - zu abweichenden Resultaten führen.

Besonders wichtig ist es zu betonen, dass die hier verglichenen Zelltypen zwar aufeinanderfolgende Stadien der Entwicklung repräsentieren, jedoch keine direkten Differenzierungsprodukte darstellen. Dadurch werden Interpretationen nicht nur erschwert sondern möglicherweise auch verfälscht.

Fazit:

EpiSC weisen distinkte Methylierungsmuster der Promotoren auf, welche (abgesehen von EpiSC-spezifischen Profilen) Ähnlichkeiten mit pluripotenten ESC als auch terminal differenzierten MEF aufweisen. In Verbindung mit genomweiten Transkriptionsdaten weisen diese Resultate auf einen dynamischen Wechsel der epigenetischen Regulation, von flexiblen Histonmodifikationen und bivalenten Domänen zu einer sekundären, stabilen Langzeitmarkierung, während der Entwicklung hin. Eine globale Ausprägung somatischer Signaturen wird zudem durch die im Vergleich zu ESC auftretende Hypermethylierung repetitiver Elemente sowie das Fehlen der in undifferenzierten Zellen typischen CpA-Methylierung (mSat) bestätigt (Analysen der Doktorandin Julia Arand, Anhang C, Abbildung 6.3-7).

Die Resultate verdeutlichen, dass DNA-Methylierung nur ein Puzzleteil in der komplexen Regulation der frühen Entwicklung darstellt und eine Charakterisierung der vorbereiteten Pluripotenz, abgesehen von einer kritischen Re-Analyse und Validierung der Expressionsdaten sowie der Berücksichtigung des genetischen Hintergrunds, nur in Kombination mit weiteren Daten wie Histonmodifikationen und 5hmC erreicht werden kann.

Zusammenfassung

Thema der vorliegenden Arbeit war die Generierung genomweiter DNA Methylierungsprofile verschiedener Zelltypen mit Hilfe der Immunopräzipitation methylierter DNA (MeDIP) und anschließender Microarray-Analyse. Dies erforderte eine laborinterne Etablierung und Optimierung experimenteller Parameter sowie die Erarbeitung bioinformatischer Analyseschritte. Zum Einen wurden kryokonservierte HepG2-Zellen im Vergleich zu unbehandelten Kontrollen analysiert um eine potentielle Epimutagenese infolge abiotischen Stresses zu untersuchen. Die Resultate führten zu dem Schluss, dass die DMSO-basierte Kryopreservation der analysierten Zellen keine vererbaren Langzeiteffekte die DNA-Methylierung betreffend verursachen.

Zum Anderen wurden verschiedene murine Stammzelltypen epigenetisch charakterisiert, um so zu einem besseren Verständnis entwicklungspezifischer Vorgänge beizutragen. Die Promotor-Regionen reprogrammierter pluripotenter Keimbahnstammzellen (gPS) wiesen dabei zwar ESC-ähnliche Muster auf, deuteten aber z.T. auf eine unvollständige bzw. aberrante Reprogrammierung hin. Die Resultate ließen eine Kompensation fehlerhafter Methylierungsmuster vermuten, da kein Zusammenhang mit Expressionsmustern oder eine Beeinträchtigung des Zellpotentials vorlag. Methylierungsmuster der dem Epiblast entstammenden Zellen (EpiSC) scheinen dagegen ein Stadium der vorbereiteten Pluripotenz wider zu spiegeln, wie es sich während der frühen Embryonalentwicklung findet.

Summary

Main topic of this thesis was the analysis of genomewide DNA-methylation patterns in different celltypes using the immunoprecipitation of methylated DNA (MeDIP) with subsequent microarray-hybridization. To this end, wet-lab procedures and bioinformatic workflows have been established and optimized.

Firstly, chromosome-wide methylation patterns of cryopreserved HepG2 cells were compared to untreated controls to explore a potential epimutagenesis after abiotic stress induction. As a result, no global changes in DNA methylation patterns were observed between two groups by bisulfite sequencing and MeDIP-chip, respectively. Furthermore, there was no direct link between expression patterns and MeDIP-enrichments. These results indicate that DMSO-treatment and freezing/thawing of HepG2 cells do not cause severe heritable long-term effects in terms of DNA methylation.

Secondly, different murine stem cell types were characterized by MeDIP-chip, thus contributing to the understanding of developmental events. Promoter-regions of pluripotent germline stem cells (gPS) possessed ESC-like patterns but indicated incomplete and aberrant reprogramming. A compensation of altered methylation patterns by unknown mechanisms was assumed as no direct relationship to expression patterns or an impairment of cell potency was obvious. Methylation signatures of epiblast stem cells (EpiSC) represented a naive pluripotent state found similar during early embryonic development.

5 Referenzen

- Ahn HJ, Sohn IP, Kwon HC, Jo DH, Park YD, Min CK (2002) Characteristics of the cell membrane fluidity, actin fibers, and mitochondrial dysfunctions of frozen-thawed two-cell mouse embryos. *Mol Reprod Dev* **61**(4): 466-476.
- Aissani B, Bernardi G (1991) CpG islands: features and distribution in the genomes of vertebrates. *Gene* **106**(2): 173-183.
- Al-Khtib M, Perret A, Khoueiry R, Ibala-Romdhane S, Blachere T, Greze C, Lornage J, Lefevre A (2011) Vitriification at the germinal vesicle stage does not affect the methylation profile of H19 and KCNQ1OT1 imprinting centers in human oocytes subsequently matured in vitro. *Fertil Steril* **95**(6): 1955-1960.
- Albert M, Peters AH (2009) Genetic and epigenetic control of early mouse development. *Curr Opin Genet Dev* **19**(2): 113-121.
- Arand J, Spieler D, Karius T, Meilinger D, Meissner A, Jenuwein T, Xu G, Leonhardt H, Wolf V, Walter J, in Revision. In vivo control of CpG and non-CpG methylation by DNA methyltransferases. In *PLoS Genet*.
- Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* **27**(4): 361-368.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129**(4): 823-837.
- Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR et al. (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**(2): 169-181.
- Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. *Cell* **128**(4): 669-681.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**(2): 315-326.
- Bestor TH (2000) The DNA methyltransferases of mammals. *Hum Mol Genet* **9**(16): 2395-2402.
- Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* **16**(1): 6-21.
- Bird AP, Wolffe AP (1999) Methylation-induced repression--belts, braces, and chromatin. *Cell* **99**(5): 451-454.
- Bock C, Halachev K, Buch J, Lengauer T (2009) EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi)genomic data. *Genome Biol* **10**(2): R14.
- Bock C, Reither S, Mikeska T, Paulsen M, Walter J, Lengauer T (2005) BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21**(21): 4067-4068.
- Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, Jager N, Gnirke A, Stunnenberg HG, Meissner A (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol* **28**(10): 1106-1114.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**(2): 185-193.
- Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, Forne T, Weber M (2010) Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* **42**(12): 1093-1100.
- Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**(5845): 1760-1764.
- Bourc'his D, Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**(7004): 96-99.
- Branco MR, Ficiz G, Reik W (2012) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet* **13**(1): 7-13.
- Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA et al. (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**(7150): 191-195.
- Cairns BR (2007) Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol* **14**(11): 989-996.

- Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* **450**(7173): 1230-1234.
- Chavez L, Jozefczuk J, Grimm C, Dietrich J, Timmermann B, Lehrach H, Herwig R, Adjaye J (2010) Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. *Genome Res* **20**(10): 1441-1450.
- Chen H, Yan Y, Davidson TL, Shinkai Y, Costa M (2006) Hypoxic stress induces dimethylated histone H3 lysine 9 through histone methyltransferase G9a in mammalian cells. *Cancer Res* **66**(18): 9009-9016.
- Chin MH, Pellegrini M, Plath K, Lowry WE (2010) Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. *Cell Stem Cell* **7**(2): 263-269.
- Chinnusamy V, Zhu JK (2009) Epigenetic regulation of stress responses in plants. *Curr Opin Plant Biol* **12**(2): 133-139.
- Choi EK, Uyeno S, Nishida N, Okumoto T, Fujimura S, Aoki Y, Nata M, Sagisaka K, Fukuda Y, Nakao K et al. (1996) Alterations of c-fos gene methylation in the processes of aging and tumorigenesis in human liver. *Mutat Res* **354**(1): 123-128.
- Chou YF, Chen HH, Eijpe M, Yabuuchi A, Chenoweth JG, Tesar P, Lu J, McKay RD, Geijsen N (2008) The growth factor environment defines distinct pluripotent ground states in novel blastocyst-derived stem cells. *Cell* **135**(3): 449-461.
- Clark SJ, Harrison J, Paul CL, Frommer M (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res* **22**(15): 2990-2997.
- Clark SJ, Statham A, Stirzaker C, Molloy PL, Frommer M (2006) DNA methylation: bisulphite modification and analysis. *Nat Protoc* **1**(5): 2353-2364.
- Cleveland WS, Devlin SJ (1988) Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* **83** (403): 596-610.
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**(7184): 215-219.
- Colman A, Dreesen O (2009) Induced pluripotent stem cells and the stability of the differentiated state. *EMBO Rep* **10**(7): 714-721.
- Cortese R, Lewin J, Backdahl L, Krispin M, Wasserkort R, Eckhardt F, Beck S (2011) Genome-wide screen for differential DNA methylation associated with neural cell differentiation in mouse. *PLoS One* **6**(10): e26002.
- Cross SH, Charlton JA, Nan X, Bird AP (1994) Purification of CpG islands using a methylated DNA binding column. *Nat Genet* **6**(3): 236-244.
- Cross SH, Lee M, Clark VH, Craig JM, Bird AP, Bickmore WA (1997) The chromosomal distribution of CpG islands in the mouse: evidence for genome scrambling in the rodent lineage. *Genomics* **40**(3): 454-461.
- Dann CT, Alvarado AL, Molyneux LA, Denard BS, Garbers DL, Porteus MH (2008) Spermatogonial stem cell self-renewal requires OCT4, a factor downregulated during retinoic acid-induced differentiation. *Stem Cells* **26**(11): 2928-2937.
- de Las Heras I, Sundaram YMA, Schatrowski S, Stancheva I. Prometheus: a software for the integrated analysis of promoter and expression arrays (Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh, UK, in Vorbereitung).
- Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A* **99**(8): 5261-5266.
- Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. *Genes Dev* **25**(10): 1010-1022.
- Dindot SV, Person R, Strivens M, Garcia R, Beaudet AL (2009) Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions. *Genome Res* **19**(8): 1374-1383.
- Dodge JE, Ramsahoye BH, Wo ZG, Okano M, Li E (2002) De novo methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* **289**(1-2): 41-48.
- Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S et al. (2009) Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* **41**(12): 1350-1353.

- Donkena KV, Young CY, Tindall DJ (2010) Oxidative stress and DNA methylation in prostate cancer. *Obstet Gynecol Int* **2010**: 302051.
- Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM et al. (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat Biotechnol* **26**(7): 779-785.
- Dudziec E, Gogol-Doring A, Cookson V, Chen W, Catto J (2012) Integrated epigenome profiling of repressive histone modifications, DNA methylation and gene expression in normal and malignant urothelial cells. *PLoS One* **7**(3): e32750.
- Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA et al. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* **38**(12): 1378-1385.
- Esteller M, Herman JG (2002) Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J Pathol* **196**(1): 1-7.
- Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**(5819): 154-156.
- Farthing CR, Ficz G, Ng RK, Chan CF, Andrews S, Dean W, Hemberger M, Reik W (2008) Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet* **4**(6): e1000116.
- Feil R (2006) Environmental and nutritional effects on the epigenetic regulation of genes. *Mutat Res* **600**(1-2): 46-57.
- Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, Reik W (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**(7347): 398-402.
- Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of ageing. *Nature* **408**(6809): 239-247.
- Fisher CL, Fisher AG (2011) Chromatin states in pluripotent, differentiated, and reprogrammed cells. *Curr Opin Genet Dev* **21**(2): 140-146.
- Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, Van Neste L, Jaenisch R, Fan G (2008) Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2**(2): 160-169.
- Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* **89**(5): 1827-1831.
- Fuller BJ, Gower JD, Green CJ (1988) Free radical damage and organ preservation: fact or fiction? A review of the interrelationship between oxidative stress and physiological ion disbalance. *Cryobiology* **25**(5): 377-393.
- Gao D, Critser JK (2000) Mechanisms of cryoinjury in living cells. *ILAR J* **41**(4): 187-196.
- Gardiner-Garden M, Frommer M (1987) CpG islands in vertebrate genomes. *J Mol Biol* **196**(2): 261-282.
- Gebhard C, Schwarzfischer L, Pham TH, Andreesen R, Mackensen A, Rehli M (2006a) Rapid and sensitive detection of CpG-methylation using methyl-binding (MB)-PCR. *Nucleic Acids Res* **34**(11): e82.
- Gebhard C, Schwarzfischer L, Pham TH, Schilling E, Klug M, Andreesen R, Rehli M (2006b) Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. *Cancer Res* **66**(12): 6118-6128.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**(10): R80.
- Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74**: 481-514.
- Greber B, Wu G, Bernemann C, Joo JY, Han DW, Ko K, Tapia N, Sabour D, Sternecker J, Tesar P et al. (2010) Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* **6**(3): 215-226.
- Guan K, Nayernia K, Maier LS, Wagner S, Dressel R, Lee JH, Nolte J, Wolf F, Li M, Engel W et al. (2006) Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* **440**(7088): 1199-1203.
- Guenther MG, Frampton GM, Soldner F, Hockemeyer D, Mitalipova M, Jaenisch R, Young RA (2010) Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell* **7**(2): 249-257.

- Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**(1): 77-88.
- Guibert S, Forne T, Weber M (2012) Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res* **22**(4): 633-641.
- Guo G, Yang J, Nichols J, Hall JS, Eyres I, Mansfield W, Smith A (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**(7): 1063-1069.
- Haines TR, Rodenhiser DI, Ainsworth PJ (2001) Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. *Dev Biol* **240**(2): 585-598.
- Han DW, Greber B, Wu G, Tapia N, Arauzo-Bravo MJ, Ko K, Bernemann C, Stehling M, Scholer HR (2011) Direct reprogramming of fibroblasts into epiblast stem cells. *Nat Cell Biol* **13**(1): 66-71.
- Han DW, Tapia N, Joo JY, Greber B, Arauzo-Bravo MJ, Bernemann C, Ko K, Wu G, Stehling M, Do JT et al. (2010) Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. *Cell* **143**(4): 617-627.
- Hanna JH, Saha K, Jaenisch R (2010) Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**(4): 508-525.
- Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y et al. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol* **28**(10): 1097-1105.
- Hayashi K, Lopes SM, Tang F, Surani MA (2008) Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**(4): 391-401.
- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L et al. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**(6047): 1303-1307.
- Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA et al. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**(3): 311-318.
- Hitchler MJ, Domann FE (2009) Metabolic defects provide a spark for the epigenetic switch in cancer. *Free Radic Biol Med* **47**(2): 115-127.
- Hochedlinger K, Plath K (2009) Epigenetic reprogramming and induced pluripotency. *Development* **136**(4): 509-523.
- Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, Uren PJ, Park J, Butler J, Rafii S, McCombie WR et al. (2011) Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol Cell* **44**(1): 17-28.
- Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L et al. (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res* **19**(9): 1593-1605.
- Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**(1): 44-57.
- Huang J, Renault V, Sengenès J, Touleimat N, Michel S, Lathrop M, Tost J (2012) MeQA: a pipeline for MeDIP-seq data quality assessment and analysis. *Bioinformatics* **28**(4): 587-588.
- Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5**(1): e8888.
- Huber W, Irizarry R, Gentleman R (2005) Preprocessing Overview. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, (ed. R Gentleman, V Carey, W Huber, R Irizarry, S Dudoit), p. 473. Springer, New York, NY.
- Hulsen T, de Vlieg J, Alkema W (2008) BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**: 488.
- Ibanez de Caceres I, Dulaimi E, Hoffman AM, Al-Saleem T, Uzzo RG, Cairns P (2006) Identification of novel target genes by an epigenetic reactivation screen of renal cancer. *Cancer Res* **66**(10): 5021-5028.
- Ihaka R, Gentleman R (1996) R: A Language for Data Analysis and Graphics. *J Comput Graph Stat* **5**(3): 299-314

- Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M et al. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* **41**(2): 178-186.
- Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**(7310): 1129-1133.
- Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**(6047): 1300-1303.
- Itskovitz-Eldor J, Schuldiner M, Karsenti D, Eden A, Yanuka O, Amit M, Soreq H, Benvenisty N (2000) Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol Med* **6**(2): 88-95.
- Iwatani M, Ikegami K, Kremenska Y, Hattori N, Tanaka S, Yagi S, Shiota K (2006) Dimethyl sulfoxide has an impact on epigenetic profile in mouse embryoid body. *Stem Cells* **24**(11): 2549-2556.
- Izadyar F, Pau F, Marh J, Slepko N, Wang T, Gonzalez R, Ramos T, Howerton K, Sayre C, Silva F (2008) Generation of multipotent cell lines from a distinct population of male germ line stem cells. *Reproduction* **135**(6): 771-784.
- Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33** Suppl: 245-254.
- Jaenisch R, Young R (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**(4): 567-582.
- Jenuwein T, Allis CD (2001) Translating the histone code. *Science* **293**(5532): 1074-1080.
- Jia J, Pekowska A, Jaeger S, Benoukraf T, Ferrier P, Spicuglia S (2010) Assessing the efficiency and significance of Methylated DNA Immunoprecipitation (MeDIP) assays in using in vitro methylated genomic DNA. *BMC Res Notes* **3**: 240.
- Jiang C, Han L, Su B, Li WH, Zhao Z (2007) Features and trend of loss of promoter-associated CpG islands in the human and mouse genomes. *Mol Biol Evol* **24**(9): 1991-2000.
- Jin SG, Kadam S, Pfeifer GP (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res* **38**(11): e125.
- Johnson DS, Li W, Gordon DB, Bhattacharjee A, Curry B, Ghosh J, Brizuela L, Carroll JS, Brown M, Flicek P et al. (2008) Systematic evaluation of variability in ChIP-chip experiments using predefined DNA targets. *Genome Res* **18**(3): 393-403.
- Jorgensen HF, Adie K, Chaubert P, Bird AP (2006) Engineering a high-affinity methyl-CpG-binding protein. *Nucleic Acids Res* **34**(13): e96.
- Kaikkonen MU, Lam MT, Glass CK (2011) Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc Res* **90**(3): 430-440.
- Kaity A, Ashmore SE, Drew RA, Dulloo ME (2008) Assessment of genetic and epigenetic changes following cryopreservation in papaya. *Plant Cell Rep* **27**(9): 1529-1539.
- Kanatsu-Shinohara M, Lee J, Inoue K, Ogonuki N, Miki H, Toyokuni S, Ikawa M, Nakamura T, Ogura A, Shinohara T (2008) Pluripotency of a single spermatogonial stem cell in mice. *Biol Reprod* **78**(4): 681-687.
- Kanatsu-Shinohara M, Ogonuki N, Inoue K, Miki H, Ogura A, Toyokuni S, Shinohara T (2003) Long-term proliferation in culture and germline transmission of mouse male germline stem cells. *Biol Reprod* **69**(2): 612-616.
- Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**(6994): 900-903.
- Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H et al. (2006) Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat Genet* **38**(2): 149-153.
- Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI et al. (2010) Epigenetic memory in induced pluripotent stem cells. *Nature* **467**(7313): 285-290.
- Ko K, Tapia N, Wu G, Kim JB, Bravo MJ, Sasse P, Glaser T, Ruau D, Han DW, Greber B et al. (2009) Induction of pluripotency in adult unipotent germline stem cells. *Cell Stem Cell* **5**(1): 87-96.
- Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G et al. (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**(2): 200-213.

- Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**(4): 693-705.
- Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**(5929): 929-930.
- Kubota H, Avarbock MR, Brinster RL (2004) Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. *Proc Natl Acad Sci U S A* **101**(47): 16489-16494.
- Laird PW (2010) Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* **11**(3): 191-203.
- Laurent L, Wong E, Li G, Huynh T, Tsiganos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J et al. (2010) Dynamic changes in the human methylome during differentiation. *Genome Res* **20**(3): 320-331.
- Lim SO, Gu JM, Kim MS, Kim HS, Park YN, Park CK, Cho JW, Park YM, Jung G (2008) Epigenetic changes induced by reactive oxygen species in hepatocellular carcinoma: methylation of the E-cadherin promoter. *Gastroenterology* **135**(6): 2128-2140, 2140 e2121-2128.
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**(3): 523-536.
- Lister R, Pelizzola M, Downen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**(7271): 315-322.
- Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S et al. (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**(7336): 68-73.
- Liu K, Yang Y, Mansbridge J (2000) Comparison of the stress response to cryopreservation in monolayer and three-dimensional human fibroblast cultures: stress proteins, MAP kinases, and growth factor gene expression. *Tissue Eng* **6**(5): 539-554.
- Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R (2006) Interchromosomal interactions and olfactory receptor choice. *Cell* **126**(2): 403-413.
- Lu GD, Leung CH, Yan B, Tan CM, Low SY, Aung MO, Salto-Tellez M, Lim SG, Hooi SC (2010) C/EBPalpha is up-regulated in a subset of hepatocellular carcinomas and plays a role in cell growth and proliferation. *Gastroenterology* **139**(2): 632-643, 643 e631-634.
- Lutsik P, Feuerbach L, Arand J, Lengauer T, Walter J, Bock C (2011) BiQ Analyzer HT: locus-specific analysis of DNA methylation by high-throughput bisulfite sequencing. *Nucleic Acids Res* **39**(Web Server issue): W551-556.
- Matsuo K, Clay O, Takahashi T, Silke J, Schaffner W (1993) Evidence for erosion of mouse CpG islands during mammalian evolution. *Somat Cell Mol Genet* **19**(6): 543-555.
- Mazur P (1963) Kinetics of Water Loss from Cells at Subzero Temperatures and the Likelihood of Intracellular Freezing. *J Gen Physiol* **47**: 347-369.
- Mazur P (1984) Freezing of living cells: mechanisms and implications. *Am J Physiol* **247**(3 Pt 1): C125-142.
- Mazur P, Leibo SP, Chu EH (1972) A two-factor hypothesis of freezing injury. Evidence from Chinese hamster tissue-culture cells. *Exp Cell Res* **71**(2): 345-355.
- Mazur P, Schneider U (1986) Osmotic responses of preimplantation mouse and bovine embryos and their cryobiological implications. *Cell Biophys* **8**(4): 259-285.
- Meissner A (2010) Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* **28**(10): 1079-1088.
- Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* **33**(18): 5868-5877.
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB et al. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**(7205): 766-770.
- Meryman HT (2007) Cryopreservation of living cells: principles and practice. *Transfusion* **47**(5): 935-945.
- Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**(7200): 49-55.

- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**(7153): 553-560.
- Milroy C, Liu L, Hammoud S, Hammoud A, Peterson CM, Carrell DT (2011) Differential methylation of pluripotency gene promoters in in vitro matured and vitrified, in vivo-matured mouse oocytes. *Fertil Steril* **95**(6): 2094-2099.
- Milutinovic S, Zhuang Q, Niveleau A, Szyf M (2003) Epigenomic stress response. Knockdown of DNA methyltransferase 1 triggers an intra-S-phase arrest of DNA replication and induction of stress response genes. *J Biol Chem* **278**(17): 14985-14995.
- Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* **30**(6): 755-766.
- Nagy A, Rossant J, Nagy R, Abramow-Newerly W, Roder JC (1993) Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc Natl Acad Sci U S A* **90**(18): 8424-8428.
- Nair SS, Coolen MW, Stirzaker C, Song JZ, Statham AL, Strbenac D, Robinson MD, Clark SJ (2011) Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias. *Epigenetics* **6**(1): 34-44.
- Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* **23**(1): 58-61.
- Nichols J, Smith A (2009) Naive and primed pluripotent states. *Cell Stem Cell* **4**(6): 487-492.
- Nishino K, Toyoda M, Yamazaki-Inoue M, Fukawatase Y, Chikazawa E, Sakaguchi H, Akutsu H, Umezawa A (2011) DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genet* **7**(5): e1002085.
- Niu X, Arthur PG, Jeffrey GP (2010) Iron and oxidative stress in cold-initiated necrotic death of rat hepatocyte. *Transplant Proc* **42**(5): 1563-1568.
- O'Geen H, Nicolet CM, Blahnik K, Green R, Farnham PJ (2006) Comparison of sample preparation methods for ChIP-chip assays. *Biotechniques* **41**(5): 577-580.
- Ohi Y, Qin H, Hong C, Blouin L, Polo JM, Guo T, Qi Z, Downey SL, Manos PD, Rossi DJ et al. (2011) Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat Cell Biol* **13**(5): 541-549.
- Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**(3): 247-257.
- Okano M, Xie S, Li E (1998) Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* **26**(11): 2536-2540.
- Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* **448**(7151): 313-317.
- Palmke N, Santacruz D, Walter J (2011) Comprehensive analysis of DNA-methylation in mammalian tissues using MeDIP-chip. *Methods* **53**(2): 175-184.
- Pannetier M, Feil R (2007) Epigenetic stability of embryonic stem cells and developmental potential. *Trends Biotechnol* **25**(12): 556-562.
- Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P et al. (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**(7347): 394-397.
- Pegg DE (2002) The history and principles of cryopreservation. *Semin Reprod Med* **20**(1): 5-13.
- Pelizzola M, Koga Y, Urban AE, Krauthammer M, Weissman S, Halaban R, Molinaro AM (2008) MEDME: an experimental and analytical methodology for the estimation of DNA methylation levels based on microarray derived MeDIP-enrichment. *Genome Res* **18**(10): 1652-1659.
- Peng L, Wang S, Yin S, Li C, Li Z, Liu Q (2008) Autophosphorylation of H2AX in a cell-specific frozen dependent way. *Cryobiology* **57**(2): 175-177.
- Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL (2007) DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci U S A* **104**(16): 6752-6757.
- Peredo EL, Arroyo-Garcia R, Reed BM, Revilla MA (2008) Genetic and epigenetic stability of cryopreserved and cold-stored hops (*Humulus lupulus* L.). *Cryobiology* **57**(3): 234-241.
- Quackenbush J (2002) Microarray data normalization and transformation. *Nat Genet* **32** Suppl: 496-501.

- R.D.C.Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* **97**(10): 5237-5242.
- Rauen U, Polzar B, Stephan H, Mannherz HG, de Groot H (1999) Cold-induced apoptosis in cultured hepatocytes and liver endothelial cells: mediation by reactive oxygen species. *FASEB J* **13**(1): 155-168.
- Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. *Science* **293**(5532): 1089-1093.
- Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **38**: 413-443.
- Robertson KD (2005) DNA methylation and human disease. *Nat Rev Genet* **6**(8): 597-610.
- Robertson KD, Wolffe AP (2000) DNA methylation in health and disease. *Nat Rev Genet* **1**(1): 11-19.
- Robinson MD, Storzaker C, Statham AL, Coolen MW, Song JZ, Nair SS, Strbenac D, Speed TP, Clark SJ (2010) Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation. *Genome Res* **20**(12): 1719-1729.
- Rohde C, Zhang Y, Jurkowski TP, Stamerjohanns H, Reinhardt R, Jeltsch A (2008) Bisulfite sequencing Data Presentation and Compilation (BDPC) web server--a useful tool for DNA methylation analysis. *Nucleic Acids Res* **36**(5): e34.
- Scacheri PC, Crawford GE, Davis S (2006) Statistics for ChIP-chip and DNase hypersensitivity experiments on NimbleGen arrays. *Methods Enzymol* **411**: 270-282.
- Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. *Cell* **128**(4): 735-745.
- Seandel M, James D, Shmelkov SV, Falciatori I, Kim J, Chavala S, Scherr DS, Zhang F, Torres R, Gale NW et al. (2007) Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature* **449**(7160): 346-350.
- Siest G, Visvikis S, Herbeth B, Gueguen R, Vincent-Viry M, Sass C, Beaud B, Lecomte E, Steinmetz J, Locuty J et al. (1998) Objectives, design and recruitment of a familial and longitudinal cohort for studying gene-environment interactions in the field of cardiovascular risk: the Stanislas cohort. *Clin Chem Lab Med* **36**(1): 35-42.
- Silva J, Smith A (2008) Capturing pluripotency. *Cell* **132**(4): 532-536.
- Smith KP, Luong MX, Stein GS (2009) Pluripotency: toward a gold standard for human ES and iPS cells. *J Cell Physiol* **220**(1): 21-29.
- Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**: Article3.
- Smyth GK (2005) Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, (ed. R Gentleman, V Carey, S Dudoit, R Irizarry, W Huber), pp. 397-420. Springer New York.
- Smyth GK, Speed T (2003) Normalization of cDNA microarray data. *Methods* **31**(4): 265-273.
- Sonna LA, Cullivan ML, Sheldon HK, Pratt RE, Lilly CM (2003) Effect of hypoxia on gene expression by human hepatocytes (HepG2). *Physiol Genomics* **12**(3): 195-207.
- Sonna LA, Kuhlmeier MM, Carter HC, Hasday JD, Lilly CM, Fairchild KD (2006) Effect of moderate hypothermia on gene expression by THP-1 cells: a DNA microarray study. *Physiol Genomics* **26**(1): 91-98.
- Sonna LA, Kuhlmeier MM, Khatri P, Chen D, Lilly CM (2010) A microarray analysis of the effects of moderate hypothermia and rewarming on gene expression by human hepatocytes (HepG2). *Cell Stress Chaperones* **15**(5): 687-702.
- Sorensen AL, Collas P (2009) Immunoprecipitation of methylated DNA. *Methods Mol Biol* **567**: 249-262.
- Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK et al. (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**(7378): 490-495.
- Steponkus PL, Stout DG, Wolfe J, Lovelace RVE (1985) Possible role of transient electric fields in freezing-induced membrane destabilization. *Journal of Membrane Biology* **85**(3): 191-198.
- Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H (2002) Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J Biol Chem* **277**(40): 37741-37746.

- Steward N, Kusano T, Sano H (2000) Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells. *Nucleic Acids Res* **28**(17): 3250-3259.
- Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol* **12**(6): R54.
- Suo L, Meng Q, Pei Y, Fu X, Wang Y, Bunch TD, Zhu S (2010) Effect of cryopreservation on acetylation patterns of lysine 12 of histone H4 (acH4K12) in mouse oocytes and zygotes. *J Assist Reprod Genet* **27**(12): 735-741.
- Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**(6): 465-476.
- Szulwach KE, Li X, Li Y, Song CX, Han JW, Kim S, Namburi S, Hermetz K, Kim JJ, Rudd MK et al. (2011) Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet* **7**(6): e1002154.
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**(5929): 930-935.
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**(4): 663-676.
- Tan MP (2010) Analysis of DNA methylation of maize in response to osmotic and salt stress based on methylation-sensitive amplified polymorphism. *Plant Physiol Biochem* **48**(1): 21-26.
- Tegelenbosch RA, de Rooij DG (1993) A quantitative study of spermatogonial multiplication and stem cell renewal in the C3H/101 F1 hybrid mouse. *Mutat Res* **290**(2): 193-200.
- Telenius H, Carter NP, Bebb CE, Nordenskjold M, Ponder BA, Tunnacliffe A (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* **13**(3): 718-725.
- Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**(7150): 196-199.
- Thaler R, Spitzer S, Karlic H, Klaushofer K, Varga F (2012) DMSO is a strong inducer of DNA hydroxymethylation in pre-osteoblastic MC3T3-E1 cells. *Epigenetics* **7**(6).
- Thomson LK, Fleming SD, Aitken RJ, De Iulii GN, Zieschang JA, Clark AM (2009) Cryopreservation-induced human sperm DNA damage is predominantly mediated by oxidative stress rather than apoptosis. *Hum Reprod* **24**(9): 2061-2070.
- Toedling J, Skylar O, Krueger T, Fischer JJ, Sperling S, Huber W (2007) Ringo--an R/Bioconductor package for analyzing ChIP-chip readouts. *BMC Bioinformatics* **8**: 221.
- Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H (2008) Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**(5): 909-918.
- Trapphoff T, El Hajj N, Zechner U, Haaf T, Eichenlaub-Ritter U (2010) DNA integrity, growth pattern, spindle formation, chromosomal constitution and imprinting patterns of mouse oocytes from vitrified pre-antral follicles. *Hum Reprod* **25**(12): 3025-3042.
- Turk PW, Laayoun A, Smith SS, Weitzman SA (1995) DNA adduct 8-hydroxyl-2'-deoxyguanosine (8-hydroxyguanine) affects function of human DNA methyltransferase. *Carcinogenesis* **16**(5): 1253-1255.
- Underhill MF, Smales CM (2007) The cold-shock response in mammalian cells: investigating the HeLa cell cold-shock proteome. *Cytotechnology* **53**(1-3): 47-53.
- Valinluck V, Sowers LC (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* **67**(3): 946-950.
- Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC (2004) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res* **32**(14): 4100-4108.
- van Rossum G (1995) Python tutorial. *Technical Report CS-R9526*. Centrum voor Wiskunde en Informatica (CWI), Amsterdam
- Visvikis-Siest S, Siest G (2008) The STANISLAS Cohort: a 10-year follow-up of supposed healthy families. Gene-environment interactions, reference values and evaluation of biomarkers in prevention of cardiovascular diseases. *Clin Chem Lab Med* **46**(6): 733-747.

- Visvikis S, Sass C, Pallaud C, Grow MA, Zannad F, Siest G, Erlich HA, Cheng S (2000) Familial studies on the genetics of cardiovascular diseases: the Stanislas cohort. *Clin Chem Lab Med* **38**(9): 827-832.
- Wang J, Thompson B, Ren C, Ittmann M, Kwabi-Addo B (2006) Sprouty4, a suppressor of tumor cell motility, is down regulated by DNA methylation in human prostate cancer. *Prostate* **66**(6): 613-624.
- Wang Z, Xu L, He F (2010) Embryo vitrification affects the methylation of the H19/Igf2 differentially methylated domain and the expression of H19 and Igf2. *Fertil Steril* **93**(8): 2729-2733.
- Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* **37**(8): 853-862.
- Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**(4): 457-466.
- Weitzman SA, Turk PW, Milkowski DH, Kozlowski K (1994) Free radical adducts induce alterations in DNA cytosine methylation. *Proc Natl Acad Sci U S A* **91**(4): 1261-1264.
- Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP (2009) Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* **41**(2): 246-250.
- Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappaport J, Helin K (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**(7347): 343-348.
- Wobus AM, Holzhausen H, Jakel P, Schoneich J (1984) Characterization of a pluripotent stem cell line derived from a mouse embryo. *Exp Cell Res* **152**(1): 212-219.
- Wongtawan T, Taylor JE, Lawson KA, Wilmut I, Pennings S (2011) Histone H4K20me3 and HP1alpha are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J Cell Sci* **124**(Pt 11): 1878-1890.
- Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y (2011a) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* **25**(7): 679-684.
- Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y (2011b) Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**(7347): 389-393.
- Wu H, Tao J, Sun YE (2012) Regulation and function of mammalian DNA methylation patterns: a genomic perspective. *Brief Funct Genomics*.
- Wu H, Zhang Y (2011) Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev* **25**(23): 2436-2452.
- Wu SC, Zhang Y (2010) Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* **11**(9): 607-620.
- Wusteman MC, Pegg DE, Robinson MP, Wang LH, Fitch P (2002) Vitrification media: toxicity, permeability, and dielectric properties. *Cryobiology* **44**(1): 24-37.
- Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S et al. (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* **42**(4): 451-464.
- Yamashita S, Hosoya K, Gyobu K, Takeshima H, Ushijima T (2009) Development of a novel output value for quantitative assessment in methylated DNA immunoprecipitation-CpG island microarray analysis. *DNA Res* **16**(5): 275-286.
- Yan LY, Yan J, Qiao J, Zhao PL, Liu P (2010) Effects of oocyte vitrification on histone modifications. *Reprod Fertil Dev* **22**(6): 920-925.
- Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**(4): e15.
- Yu J, Thomson JA (2008) Pluripotent stem cell lines. *Genes Dev* **22**(15): 1987-1997.
- Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N (1992) Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* **89**(13): 5847-5851.
- Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J et al. (2009) DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet* **5**(3): e1000438.

- Zhong L, Xu YH, Wang JB (2009) DNA-methylation changes induced by salt stress in wheat *Triticum aestivum*. *African Journal of Biotechnology* **8**: 6201-6207.
- Zhu JQ, Liu JH, Liang XW, Xu BZ, Hou Y, Zhao XX, Sun QY (2008) Heat stress causes aberrant DNA methylation of H19 and Igf-2r in mouse blastocysts. *Mol Cells* **25**(2): 211-215.
- Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* **39**(1): 61-69.
- Ziller MJ, Muller F, Liao J, Zhang Y, Gu H, Bock C, Boyle P, Epstein CB, Bernstein BE, Lengauer T et al. (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet* **7**(12): e1002389.
- Zribi N, Feki Chakroun N, El Euch H, Gargouri J, Bahloul A, Ammar Keskes L (2010) Effects of cryopreservation on human sperm deoxyribonucleic acid integrity. *Fertil Steril* **93**(1): 159-166.

6 Anhang

6.1 Anhang A: Etablierung und Optimierung von MeDIP-Chip zur genomweiten Analyse der DNA-Methylierung

6.1.1 Resultate zur Etablierung experimenteller Parameter

Tabelle 6.1-1: Kontrollregionen zur Überprüfung der MeDIP-Qualität mittels qPCR - *Homo sapiens*. Sequenzinformationen wie CpG- und GC-Gehalt des analysierten Bereichs sind ebenso angegeben wie entsprechender Methylierungsstatus der überlappenden Region (gemäß Zhang et al. 2009).

Amplikon	Methylierung (Bisulfit Sequenzierung)				analysierter Bereich in MeDIP			
	Symbol	Status	Leuko* [%]	HEK293 [%]	HepG2 [%]	# CpG	GC-Gehalt	CpG_o/e**
TRPM2	TRPM2	mosaikal	35,5	83,0	70,0	18	0,66	0,78
MRPS6	MRPS6	mosaikal	52,3	99,8	92,0	12	0,56	0,62
KRTAP10-5	KRTAP10-5	mosaikal	60,4	80,8	84,8	9	0,59	0,40
SH3BGR	SH3BGR	methyliert	98,1	72,1	79,1	9	0,56	0,38
H2BFS	H2BFS	methyliert	95,5	87,9	92,4	20	0,60	0,82
C21orf93	C21orf93	methyliert	100,0	95,5	95,8	16	0,64	0,78
METS	METS	unmethyliert	0,0***	0,0***	0,0***	6	0,11	0,61
CpG-frei	CpG-frei	CpG-frei	0,0	0,0	0,0	0	0,35	0,00

*Leukozyten; **Verhältnis beobachteter/erwarteter CpG-Gehalt („CpG observed/expected“); ***erwarteter Methylierungsstatus gemäß Analysen an verschiedenen humanen Zelllinien (gemäß Gebhard et al. 2006a)

Tabelle 6.1-2: Kontrollregionen zur Überprüfung der MeDIP-Qualität mittels qPCR - *Mus musculus*. Sequenzinformationen wie CpG- und GC-Gehalt des analysierten Bereichs sind ebenso angegeben wie entsprechender Methylierungsstatus der überlappenden Region (gemäß Ko et al. 2009). Dabei sind die Methylierungslevel der elterlichen Allele getrennt angegeben (maternal/paternal).

Amplikon	Methylierung (Bisulfit Sequenzierung)			analysierter Bereich in MeDIP			
	Symbol	Status	somatische Zellen [%]	Keimbahn [%]	# CpG	GC-Gehalt	CpG_o/e*
H19	H19	imprinted	100/0	100/0	7	0,53	0,56
Igf2r	Igf2r	imprinted	100/0	0/0	9	0,54	0,41
Xist	Xist	hemi-methyliert	50**	0**	8	0,53	0,79
CpG-frei	CpG-frei	CpG frei	0	0	0	0,42	0,00

*Verhältnis beobachteter/erwarteter CpG-Gehalt („CpG observed/expected“), **erwarteter Methylierungsstatus aufgrund vollständiger Methylierung des Xist auf dem inaktiven X-Chromosom und fehlender Methylierung auf dem aktiven X-Chromosom

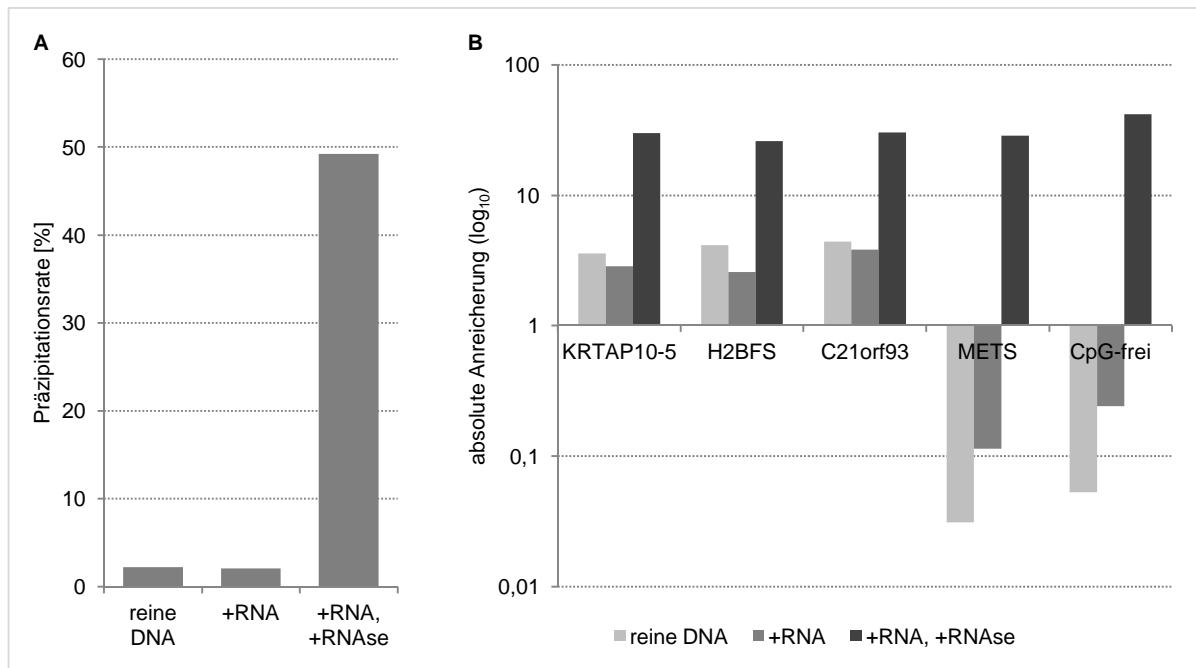


Abbildung 6.1-1: Unspezifische Präzipitation in Zusammenhang mit RNase. MeDIP-Experimente wurden mit reiner DNA und unter Zusatz von 2 µg RNA bzw. 2 µg RNA und 40 µg RNaseA (Invitrogen) durchgeführt (jeweils 2 µg DNA aus Leukozyten (676)). (A) Die Quantifizierung der Präzipitate (Fluorometer) ergab eine ungewöhnlich hohe Präzipitationsrate von 49% bei Vorhandensein von RNase. (B) Die Überprüfung der Präzipitate mittels qPCR (dargestellt ist die absolute Anreicherung in logarithmischer Skalierung) weist eine verminderte Anreicherung methylierter Regionen und eine weniger starke Abreicherung der Negativkontrollen bei Vorhandensein von RNA auf. Zusätzliche RNase resultiert in ungewöhnlicher hoher Anreicherung methylierter Regionen und einer ebenso ausgeprägten unspezifischen Präzipitation der Negativkontrollen. Vergleichbare Experimente unter Verwendung von RNase verschiedener Hersteller führten zu ähnlichen Resultaten. Die Immunopräzipitationen sowie qPCRs wurden von der Diplomandin *Sabrina Pichler* durchgeführt.

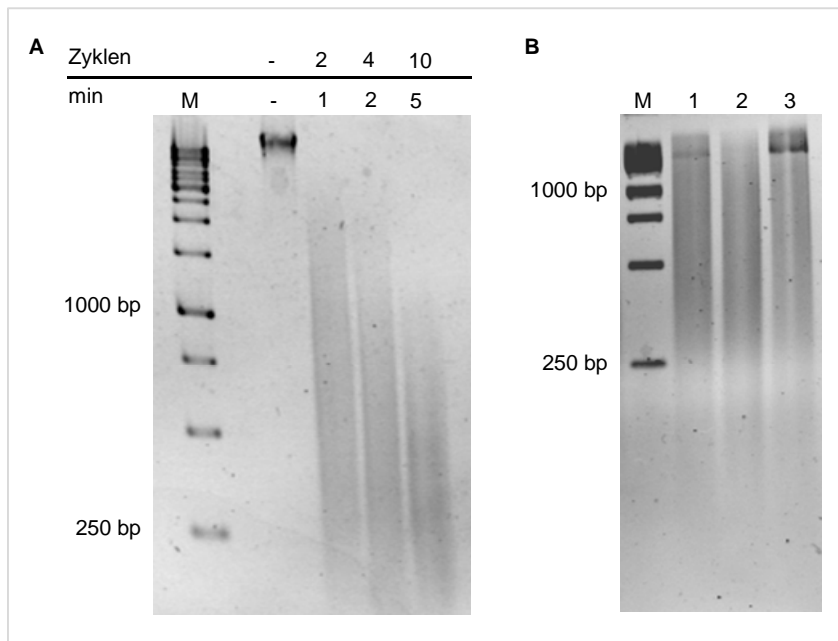


Abbildung 6.1-2: Fragmentierung genomischer DNA mittels Sonifizierung. (A) Fragmentverteilung in Abhängigkeit der Scherzyklen. Genomische DNA wurde 2, 4, und 10 An/Aus-Zyklen von je 15 s sonifiziert. Gezeigt ist die Größenverteilung der Fragmente auf einem 1,5% Agarosegel im Vergleich zu unbehandelter genomischer DNA (Spur 2). Spuren 3 und 4 weisen unvollständig sonifizierte DNA auf, Spur 5 dagegen fragmentierte DNA im verlangten Bereich zwischen 200 und 1000 nt (siehe 1kb-Leiter, Spur 1). M: Größenstandard 1kb-Leiter. Diese Sonifizierung wurde von der Diplomandin *Sabrina Pichler* vorgenommen. (B) Drei Aliquots einer genomischen DNA (HepG2) wurden in 5 An/Aus-Zyklen von je 15 s sonifiziert. Die Triplikate 1-3 weisen trotz gleicher Bedingungen ein unterschiedliches Maß der Fragmentierung auf. M: Größenstandard 1kb-Leiter.

6.1.2 Resultate zur bioinformatischen Analyse von MeDIP-Chip-Daten

Tabelle 6.1-3: Evaluierung Batman-transformierter MeDIP-Chip-Daten. MeDIP-Chip-Daten unterschiedlicher Experimente (A: HEK293T (WGA), Replikat 2; Leukozyten (480) (WGA), Replikat 1; HepG2 (*Ktr*) (WGA), Replikat 1) wurden mit Hilfe des *Batman*-Algorithmus (Down et al. 2008) analysiert und die erhaltenen Werte mit quantitativen Methylierungsdaten (Bisulfit-Sequenzierung, Zhang 2009) abgeglichen. Angegeben sind die Pearson-Korrelationswerte der mittleren Batman-Werte vs. Bisulfit-Methylierung aller Regionen. Eine erhöhte Quantifizierbarkeit der MeDIP-Werte und damit eine erhöhte Korrelation mit Bisulfit-Methylierungs-Werten ist lediglich im Falle der Experimente A und B gegeben. Die bioinformatischen Analysen wurden vom Master-Studenten *Pavlo Lutsik* vorgenommen.

Experiment	Rohdaten	Loess-norm.	Quantil-norm.	Rohdaten, Batman-trans.	Loess-norm., Batman trans.	Quantil-norm., Batman trans.
A	0,36	0,28	0,38	0,54	0,70	0,59
B	0,51	0,35	0,34	0,73	0,76	/
C	0,65	0,58	0,65	0,26	0,55	/

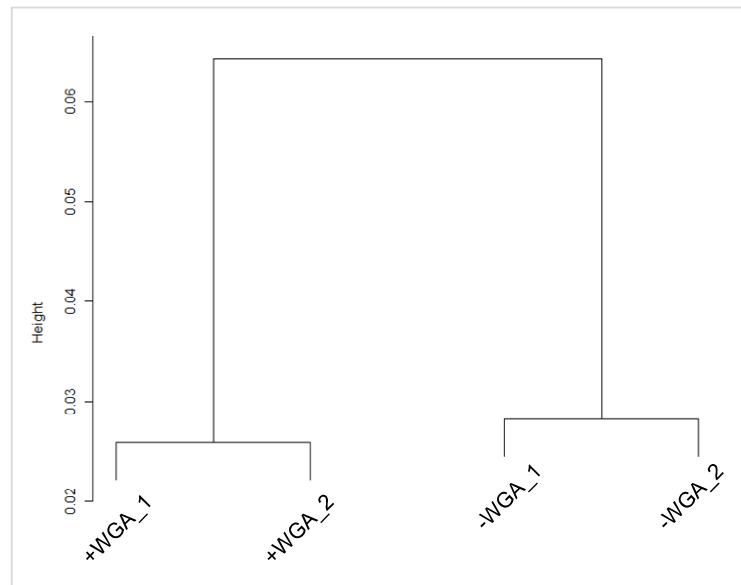


Abbildung 6.1-3: Hierarchische Clusteranalyse von MeDIP-Chip-Proben mit und ohne genomrepräsentative Amplifikation (HEK293T). Die Proben gruppieren nicht gemäß der biologischen Replikate (1 und 2) sondern der Behandlung (+WGA und -WGA).

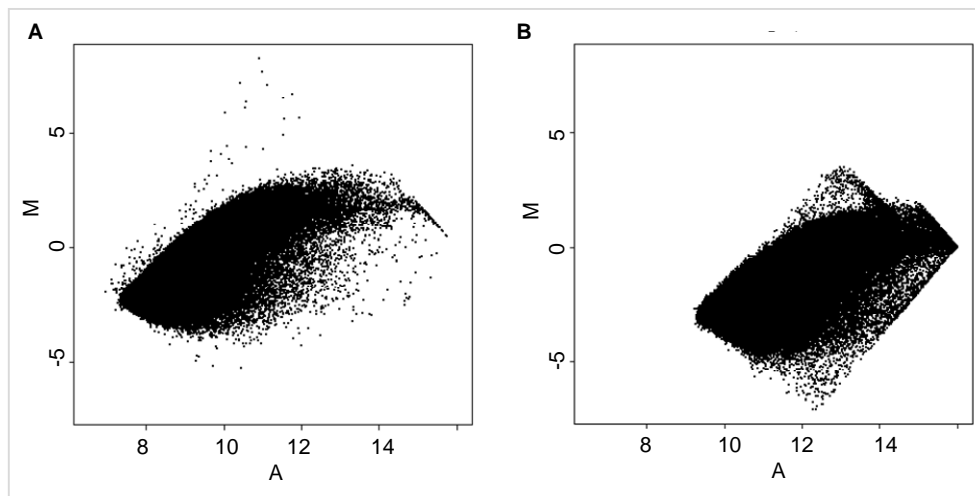


Abbildung 6.1-4: Unterschiedliche Signalverteilung in Abhängigkeit der Hybridisierung. MA-plots der MeDIP-Probe HepG2 *NKtr* (Replikate 1) nach zeitlich auseinanderliegenden Hybridisierungen 1 (A) und 2 (B).

6.2 Anhang B: Analyse chromosomenweiter DNA-Methylierung in Verbindung mit Kryopreservation

6.2.1 Protokoll zur Kryopreservation von HepG2 Zellen

Die HepG2-Zellen wurden vom IBMT gemäß Standardprotokollen kultiviert und z.T. kryokonserviert. Dazu wurden die Zellen zunächst auf eine Dichte von $1 \times 10^6 \text{ ml}^{-1}$ eingestellt, mit 10% DMSO (v/v) versetzt, aliquotiert (1 ml) und 30 min bei 4 °C vorbehandelt. Anschließend wurden die Zellen in einem programmierbaren Gerät schrittweise eingefroren (1 °C/min bis -7 °C, 10 min bei -7 °C, 1 °C/min bis -80 °C), in einen Stickstofftank überführt und mindestens über Nacht in der Gasphase belassen. Zum Auftauen wurden die Gefäße direkt in ein Wasserbad (37 °C) gebracht und sanft geschüttelt, bis das Eis geschmolzen war. Die Zellen wurden dann in vorgewärmtes Kulturmedium gebracht und anschließend kultiviert. Zellpellets folgender Proben wurden zur Verfügung gestellt:

- (1) unbehandelte Kontrolle, nicht-konfluent (*NKtr*)
- (2) unbehandelte Kontrolle, konfluent (*Ktr*)
- (3) mit DMSO vorbehandelte Zellen; folgende Kultivierung für 48 h (*DMSO*)
- (4) DMSO-behandelte, kryokonservierte Zellen; folgende Kultivierung für 48 h (*f48h*)
- (5) DMSO-behandelte, kryokonservierte Zellen; folgende Kultivierung für 120 h (*f120h*)

6.2.2 Resultate zur Immunpräzipitation von HepG2-Zellen

Tabelle 6.2-1: Ausbeute der Immunpräzipitation methylierter DNA an behandelten und unbehandelten HepG2 Zellen. Angegeben sind die absoluten Mengen angereicherter DNA (ng) sowie die errechnete Präzipitationsrate (%) der jeweiligen technischen Replikate pro Probe.

Probe		Präzipitationsrate	
		[ng]	[%]
<i>NKtr</i>	Replikat 1	93,6	2,3
	Replikat 2	135,2	3,4
<i>Ktr</i>	Replikat 1	92,8	2,3
	Replikat 2	114,4	2,9
<i>DMSO</i>	Replikat 1	104,8	2,6
	Replikat 2	103,6	2,6
<i>f48h</i>	Replikat 1	96,4	2,4
	Replikat 2	122,4	3,1
<i>f120h</i>	Replikat 1	109,2	2,7
	Replikat 2	115,2	2,9
<i>Durchschnitt</i>		108,8	2,7
<i>Standardabweichung</i>		12,8	0,3

6.2.3 Resultate zur Evaluierung differentiell angereicherter bzw. exprimierter Gene in Verbindung mit Kryopreservation

Tabelle 6.2-2: Gene/Promotoren +/-1 kb um identifizierte DPs in DMSO (basierend auf Hg18, Emsembl 2009)

Gen-Symbol	Beschreibung
ADORA2A	adenosine A2a receptor
ADRBK2	adrenergic, beta, receptor kinase 2
AL121918.26	(protein-coding gene)
AP000351.13	(unprocessed pseudogene)
C21orf34	chromosome 21 open reading frame 34
C21orf9	non-protein coding RNA 110
CRYZL1	crystallin, zeta (quinone reductase)-like 1
EDN3	endothelin 3
EIF3D	eukaryotic translation initiation factor 3, subunit D
ERG	v-ets erythroblastosis virus E26 oncogene homolog (avian)
FOXRED2	FAD-dependent oxidoreductase domain containing 2
GRIK1	glutamate receptor, ionotropic, kainate 1
IFNAR1	interferon (alpha, beta and omega) receptor 1
IGLV1-20	immunoglobulin lambda variable (I)-20 (pseudogene)
IGLV3-21	immunoglobulin lambda variable 2-11;...
KRTAP11-1	keratin associated protein 11-1
LARGE	like-glycosyltransferase
MKL1	megakaryoblastic leukemia (translocation) 1
MRPS6	mitochondrial ribosomal protein S6
NCAM2	neural cell adhesion molecule 2
PCBP3	poly(rC) binding protein 3
RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)
RP1-130H16.15	coiled-coil domain containing 157
TAF4	TAF4 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 135kDa
TBC1D10A	TBC1 domain family, member 10A
TSHZ2	teashirt zinc finger homeobox 2
TXN2	thioredoxin 2
UFD1L	ubiquitin fusion degradation 1 like (yeast)
ZNF294	ring finger protein 160

Tabelle 6.2-3: Gene/Promotoren +/-1 kb um identifizierte DPs in f48h (basierend auf Hg18, Emsembl 2009)

Gen-Symbol	Beschreibung
C21orf29	chromosome 21 open reading frame 29
C21orf91	chromosome 21 open reading frame 91
CABIN1	calcineurin binding protein 1
CACNG2	calcium channel, voltage-dependent, gamma subunit 2
CDH4	cadherin 4, type 1, R-cadherin (retinal)
CLDN14	claudin 14
CXADR	coxsackie virus and adenovirus receptor pseudogene 2; coxsackie virus and adenovirus receptor
DSCAM	Down syndrome cell adhesion molecule
GGT1A1	gamma-glutamyltransferase 5
GMEB2	glucocorticoid modulatory element binding protein 2
KRTAP10-12	keratin associated protein 10-12
KRTAP21-1	keratin associated protein 21-1
PCBP3	poly(rC) binding protein 3
PHACTR3	phosphatase and actin regulator 3
TTC28	chromosome 6 open reading frame 35; hCG1820764; tetratricopeptide repeat domain 28
USP25	ubiquitin specific peptidase 25
ZGPAT	zinc finger, CCCH-type with G patch domain

Tabelle 6.2-4: Gene/Promotoren +/-1 kb um identifizierte DPs in 120h (basierend auf Hg18, Ensembl 2009)

Gen-Symbol	Beschreibung
AP000473.2	Putative uncharacterized protein C21orf34
AP1B1	adaptor-related protein complex 1, beta 1 subunit
BCL2L13	BCL2-like 13 (apoptosis facilitator)
C21orf34	chromosome 21 open reading frame 34
C21orf42	non-protein coding RNA 158
CDH4	cadherin 4, type 1, R-cadherin (retinal)
DEPDC5	DEP domain containing 5
DSCR4	Down syndrome critical region gene 4
ENTHD1	ENTH domain-containing protein 1
HLC5	holocarboxylase synthetase (biotin-(proprionyl-Coenzyme A-carboxylase (ATP-hydrolysing)) ligase)
HSCB	HscB iron-sulfur cluster co-chaperone homolog (E. coli)
IL17RA	interleukin 17 receptor A
KCNJ15	potassium inwardly-rectifying channel, subfamily J, member 15
PATZ1	POZ (BTB) and AT hook containing zinc finger 1
PVALB	parvalbumin
SLC25A18	solute carrier family 25 (mitochondrial carrier), member 18
SYN3	synapsin III
TIAM1	T-cell lymphoma invasion and metastasis 1
TSHZ2	teashirt zinc finger homeobox 2
USP25	ubiquitin specific peptidase 25

Tabelle 6.2-5: Gene mit erhöhter Expression - DMSO-Effekt. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Gen-Symbol	Gen-Symbol	Gen-Symbol
AGPAT9	HS.156542	MT1G
AKR1B10	ID3	MT2A
APOM	INPP1	MTE
C1ORF64	KIAA1199	PLAC8
C1ORF97	LCN15	PTGR1
C20ORF127	LOC124220	RBP1
CIDEC	LOC441019	SLC26A3
CTHRC1	LOC649025	SLC2A1
ERV3	MIXL1	SLC43A2
EXOG	MT1A	SLC7A9
GCNT3	MT1E	TIMP1

Tabelle 6.2-6: Gene mit verminderter Expression - DMSO-Effekt. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
A2M	CD24	HPX	MSLN	SERPINF1
ANKRD38	CEBPA	HS.223856	NEDD4L	SERPINF2
APOA5	CES1	HS.282795	NFE2	SLC29A4
APOC3	CFB	HS.444329	NR1H4	SLC6A12
ASS1	COL18A1	IGF2	PALM	SP5
B4GALNT4	COL4A5	INHBE	PAQR8	SPP1
BTG2	ENO3	IRX3	PCOLCE	SULT2A1
C16ORF79	F5	LAMB2	PGC	TCEA3
C2	FGB	LAMC1	PLA2G2A	TFR2
C2ORF82	FGL1	LEAP2	PROX1	TMEM45B
C3	FLNB	LIME1	RARRES2	TMEM86B
C4BPA	FZD4	LIPC	RELN	TOX3
C5	GJC2	LOC401720	RPS23	TTYH1
C5ORF13	GLUD1	LOC402509	RXRA	VAV3
C6	GPAM	LOC653879	SERPINA7	VTN
C8B	HAMP	MMP11	SERPIND1	

Tabelle 6.2-7: Gene mit erhöhter Expression - Kryopreservations-Effekt

Gen-Symbol	Bezeichnung
CYP24A1	cytochrome P450, family 24, subfamily A, polypeptide 1
DUSP5	dual specificity phosphatase 5
HS.386232	human transcribed locus
HS.543887	human transcribed locus
IGFBP1	insulin-like growth factor binding protein 1
ITGA2	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)
LOC653506	meteorin, glial cell differentiation regulator-like; similar to meteorin, glial cell differentiation regulator-like
S100P	S100 calcium binding protein P
SERPINE1	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
SPRY4	sprouty homolog 4 (Drosophila)
STS-1	ubiquitin associated and SH3 domain containing, B
VGF	VGF nerve growth factor inducible

Tabelle 6.2-8: Gene mit verminderter Expression - Kryopreservations-Effekt.

Gen-Symbol	Bezeichnung
AFP	alpha-fetoprotein
DUSP1	dual specificity phosphatase 1
ERP27	endoplasmic reticulum protein 27
FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog
FOSB	FBJ murine osteosarcoma viral oncogene homolog B
HP	haptoglobin-related protein; haptoglobin
HPR	haptoglobin-related protein; haptoglobin
JUN	jun oncogene
LOC284422	similar to HSPC323
SNORD13	small nucleolar RNA, C/D box 13
TFF3	trefoil factor 3 (intestinal)

Tabelle 6.2-9: Gene mit erhöhter Expression - später Effekt. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
ADM2	CD55	FAM43A	IL11	PDGFRB	SLC7A5
AKNA	CEBPB	FILIP1L	IL1RN	PIM1	SMOX
ARG2	CEBPG	FLJ35767	ISG20	PPARGC1A	SOX8
ARHGGEF2	CHCHD10	FST	ITPR3	PROSAPIP1	SRGAP1
ASNS	CPLX1	GARS	ITPRIP	PSAT1	STC2
ATF4	CTH	GATAD2A	JAG1	RASD1	TARS
AXIN1	DBNDD2	GDF15	LOC55908	RBM24	TAT
AXL	DDIT3	GK	LOC650215	RRP12	TIGA1
BAIAP2L2	DDIT4	GMDS	LOC730525	SARS	TMC6
BCAT1	DKFZP686O24166	GOT1	LONP1	SH2B3	TNFRSF12A
BEX2	DLC1	GRPEL2	MARS	SHMT2	TNFSF4
BMP4	DLK1	HMGA1	MBNL2	SLC16A6	TRIB1
BPIL1	DNAJC12	HOXD1	MOCOS	SLC17A2	TRIB3
C10ORF10	DSCR6	HS.28170	MTHFD1L	SLC1A5	UTP15
C13ORF15	EEF1A2	HS.566710	MTHFD2	SLC20A1	WARS
C5AR1	EIF1	HS.82028	NAV2	SLC22A15	WNK4
CACNA1H	EIF2C2	HSD17B2	NKD2	SLC38A1	YARS
CARS	ENPP1	HSPA9	NPAS1	SLC38A2	
CBS	ERN1	IER3	OSCAR	SLC4A7	
CCL20	ETV4	IFRD1	OSTBETA	SLC7A1	
CCNB1IP1	FAM100B	IGFBP3	PALLD	SLC7A2	

Tabelle 6.2-10: Gene mit verminderter Expression - später Effekt. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
ACAT2	C1ORF53	EPDR1	IDI1	PCNA	TK1
ADH6	CALM3	FDFT1	LIPA	PLA2G12B	TM7SF2
ALOX5AP	CAT	FGG	LOC148915	PPP1R3C	TMEM14A
AQP11	CETN2	FN3KRP	LOC158160	PRDX2	TMEM37
ARL6IP5	CREG1	GCHFR	LOC653874	RBP2	TSC22D1
ASB9	CRIP1	GINS2	LOC730534	REEP6	TSPAN13
ATP6V0E2	DBI	GSTA4	LSM3	RPA3	TYMS
BRP44	DDC	GSTM1	MAGED1	SC4MOL	UBD
C11ORF67	DGAT1	GSTM2	METTTL7B	SERPINB1	UHRF1
C16ORF33	DIO1	HIST1H2BD	MGST3	SH3BGRL	
C1ORF106	EBP	HMG2	MT1F	SLC16A10	
C1ORF41	EHHADH	IDH1	NSDHL	SPTLC3	

Tabelle 6.2-11: GO-Annotationen überrepräsentierter biologischer Prozesse für herunterregulierte Gene mit DMSO-Effekt. Gene mit statistisch signifikanter deregulierter Expression wurden mittels der DAVID-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem HT12 bead-Chip abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Antwort des Immunsystems		
GO:0006956-complement activation	2,8E-08	7
GO:0002541-activation of plasma proteins involved in acute inflammatory response	3,3E-08	7
GO:0048584-positive regulation of response to stimulus	1,1E-07	11
GO:0006958-complement activation, classical pathway	2,1E-07	6
GO:0050778-positive regulation of immune response	2,8E-07	9
GO:0002455-humoral immune response mediated by circulating immunoglobulin	3,0E-07	6
GO:0006959-humoral immune response	1,5E-06	7
GO:0051605-protein maturation by peptide bond cleavage	2,5E-06	7
GO:0002253-activation of immune response	4,0E-06	7
GO:0016064-immunoglobulin mediated immune response	5,1E-06	6
GO:0019724-B cell mediated immunity	6,1E-06	6
GO:0016485-protein processing	1,1E-05	7
GO:0002684-positive regulation of immune system process	1,2E-05	9
GO:0002449-lymphocyte mediated immunity	1,8E-05	6
GO:0051604-protein maturation	1,9E-05	7
GO:0002250-adaptive immune response	2,7E-05	6
GO:0002460-adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	2,7E-05	6
GO:0002252-immune effector process	3,4E-05	7
GO:0045087-innate immune response	3,9E-05	7
GO:0002443-leukocyte mediated immunity	5,0E-05	6
Makromoleküle: Umbau und Metabolismus		
GO:0034367-macromolecular complex remodeling	1,1E-04	4
GO:0034368-protein-lipid complex remodeling	1,1E-04	4
GO:0034369-plasma lipoprotein particle remodeling	1,1E-04	4
GO:0070328-triglyceride homeostasis	5,9E-04	3
GO:0019433-triglyceride catabolic process	5,9E-04	3
GO:0046464-acylglycerol catabolic process	7,6E-04	3
GO:0046461-neutral lipid catabolic process	7,6E-04	3
GO:0044269-glycerol ether catabolic process	7,6E-04	3
GO:0006641-triglyceride metabolic process	1,1E-03	4
GO:0046503-glycerolipid catabolic process	1,4E-03	3
GO:0006639-acylglycerol metabolic process	1,5E-03	4
GO:0006638-neutral lipid metabolic process	1,6E-03	4
GO:0006662-glycerol ether metabolic process	1,7E-03	4
GO:0018904-organic ether metabolic process	1,9E-03	4
GO:0046486-glycerolipid metabolic process	6,0E-03	5
GO:0016042-lipid catabolic process	7,6E-03	5
GO:0044242-cellular lipid catabolic process	4,6E-02	3

GO:0006869-lipid transport	1,4E-01	3
GO:0010876-lipid localization	1,6E-01	3
Regulation metabolischer Prozesse (Lipide)		
GO:0046890-regulation of lipid biosynthetic process	1,2E-03	4
GO:0019216-regulation of lipid metabolic process	1,7E-03	5
GO:0010565-regulation of cellular ketone metabolic process	2,2E-03	4
GO:0050994-regulation of lipid catabolic process	6,5E-03	3
GO:0045833-negative regulation of lipid metabolic process	7,6E-03	3
GO:0009894-regulation of catabolic process	1,0E-02	4
GO:0019217-regulation of fatty acid metabolic process	2,0E-02	3
GO:0031329-regulation of cellular catabolic process	3,3E-02	3
Endozytose		
GO:0030100-regulation of endocytosis	2,9E-03	4
GO:0045807-positive regulation of endocytosis	8,1E-03	3
GO:0060627-regulation of vesicle-mediated transport	1,0E-02	4
Regulation der Immunantwort		
GO:0002824-positive regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	8,1E-03	3
GO:0002821-positive regulation of adaptive immune response	8,6E-03	3
GO:0002705-positive regulation of leukocyte mediated immunity	1,1E-02	3
GO:0002708-positive regulation of lymphocyte mediated immunity	1,1E-02	3
GO:0002699-positive regulation of immune effector process	1,8E-02	3
GO:0002706-regulation of lymphocyte mediated immunity	2,6E-02	3
GO:0002822-regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	2,6E-02	3
GO:0002819-regulation of adaptive immune response	2,7E-02	3
GO:0002703-regulation of leukocyte mediated immunity	3,3E-02	3
Zell-Morphogenese		
GO:0000904-cell morphogenesis involved in differentiation	5,4E-03	6
GO:0000902-cell morphogenesis	2,4E-02	6
GO:0032989-cellular component morphogenesis	3,6E-02	6
Antwort auf externen Stimulus / Stress		
GO:0050729-positive regulation of inflammatory response	8,1E-03	3
GO:0032103-positive regulation of response to external stimulus	3,5E-02	3
GO:0031349-positive regulation of defense response	4,4E-02	3
Regeneration und Wachstum		
GO:0031099-regeneration	4,0E-03	4
GO:0040007-growth	5,3E-02	4
GO:0048589-developmental growth	6,3E-02	3

Tabelle 6.2-12: GO-Annotationen überrepräsentierter biologischer Prozesse für herunterregulierte Gene mit Kryopreservations-Effekt. Gene mit statistisch signifikanter deregulierter Expression wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem HT12 bead-Chip abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Antwort auf externen Stimulus / Stress		
GO:0051591-response to cAMP	1,6E-04	3
GO:0000302-response to reactive oxygen species	4,9E-04	3
GO:0006979-response to oxidative stress	2,3E-03	3
GO:0010035-response to inorganic substance	3,5E-03	3
GO:0009991-response to extracellular stimulus	4,1E-03	3
GO:0009628-response to abiotic stimulus	1,1E-02	3

Tabelle 6.2-13: GO-Annotationen überrepräsentierter biologischer Prozesse für hochregulierte Gene mit spätem Effekt. Gene mit statistisch signifikanter deregulierter Expression wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem HT12 bead-Chip abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Aminosäure-Aktivierung		
GO:0043038-amino acid activation	8,0E-07	7
GO:0043039-tRNA aminoacylation	8,0E-07	7
GO:0006418-tRNA aminoacylation for protein translation	8,0E-07	7
GO:0006399-tRNA metabolic process	1,8E-04	7
GO:0034660-ncRNA metabolic process	1,0E-03	8
Phosphat-Metabolismus		
GO:0051174-regulation of phosphorus metabolic process	9,2E-06	15
GO:0019220-regulation of phosphate metabolic process	9,2E-06	15
GO:0042325-regulation of phosphorylation	2,9E-05	14
Transport (Aminosäuren)		
GO:0006865-amino acid transport	5,1E-04	6
GO:0046942-carboxylic acid transport	7,2E-04	7
GO:0015849-organic acid transport	7,5E-04	7
GO:0015837-amine transport	1,8E-03	6
Biosynthese (Aminosäuren)		
GO:0008652-cellular amino acid biosynthetic process	3,0E-05	6
GO:0009309-amine biosynthetic process	3,0E-04	6
GO:0016053-organic acid biosynthetic process	4,9E-03	6
GO:0046394-carboxylic acid biosynthetic process	4,9E-03	6
GO:0044271-nitrogen compound biosynthetic process	9,2E-03	8
Regulation der Transkription		
GO:0043433-negative regulation of transcription factor activity	3,4E-04	5
GO:0043392-negative regulation of DNA binding	5,5E-04	5
GO:0051100-negative regulation of binding	9,5E-04	5
GO:0051090-regulation of transcription factor activity	7,2E-03	5
GO:0051101-regulation of DNA binding	1,3E-02	5
GO:0051098-regulation of binding	2,6E-02	5
Katabolismus (Aminosäuren)		
GO:0009310-amine catabolic process	2,2E-03	5
GO:0009063-cellular amino acid catabolic process	1,2E-02	4
GO:0016054-organic acid catabolic process	4,4E-02	4
GO:0046395-carboxylic acid catabolic process	4,4E-02	4

Tabelle 6.2-14: GO-Annotationen überrepräsentierter biologischer Prozesse für herunterregulierte Gene mit spätem Effekt. Gene mit statistisch signifikanter deregulierter Expression wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der auf dem HT12 bead-Chip abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Lipid-Biosynthese		
GO:0016126-sterol biosynthetic process	1,4E-07	6
GO:0006695-cholesterol biosynthetic process	2,1E-06	5
GO:0006694-steroid biosynthetic process	9,4E-06	6

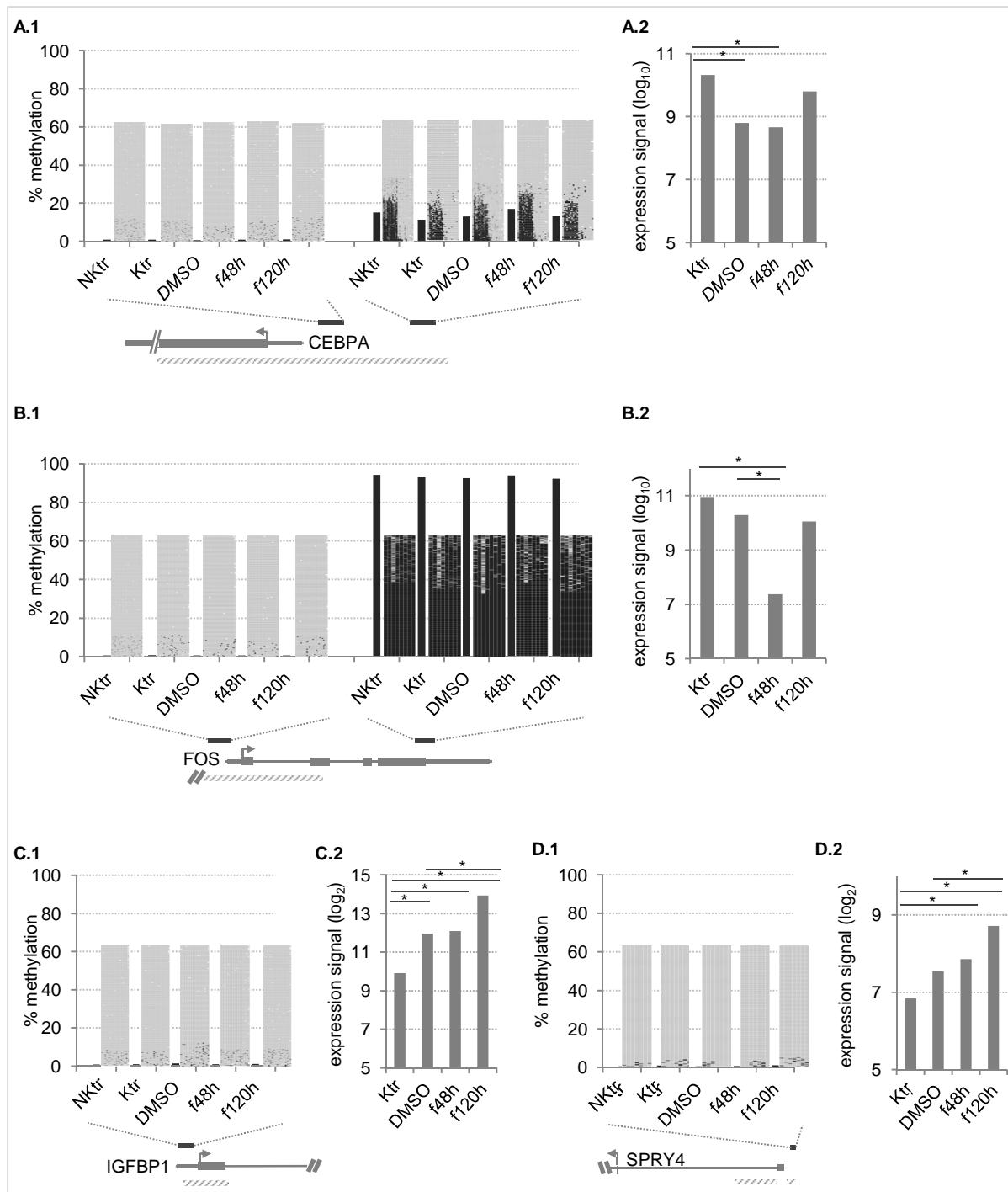


Abbildung 6.2-1: Bisulfite-Sequenzierung vier differentiell exprimierter Gene. Pro Gen sind jeweils mittlerer Methylierungsstatus sowie entsprechende Methylierungsmuster (Vertikale: individuelle Sequenzen, Horizontale: einzelne CpG-Positionen) dargestellt. Positionen der analysierten Amplikons (schwarz) sind darunter relativ zu CpG-Inseln (gestrichelt) und TSS (Pfeil) angegeben. Zum Vergleich wurden mittlere Expressionslevel (illumina) aufgetragen. *signifikanter Expressionsunterschied (ANOVA). (A) CEBPA (zwei Amplikons), (B) FOS (zwei Amplikons), (C) IGFBP1, (D) SPRY4.

6.3 Anhang C: Vergleichende Analyse genomweiter DNA-Methylierungsmuster verschiedener muriner Stammzellen

6.3.1 Resultate zur Evaluierung differentiell angereicherter Gene in murinen Stammzellen

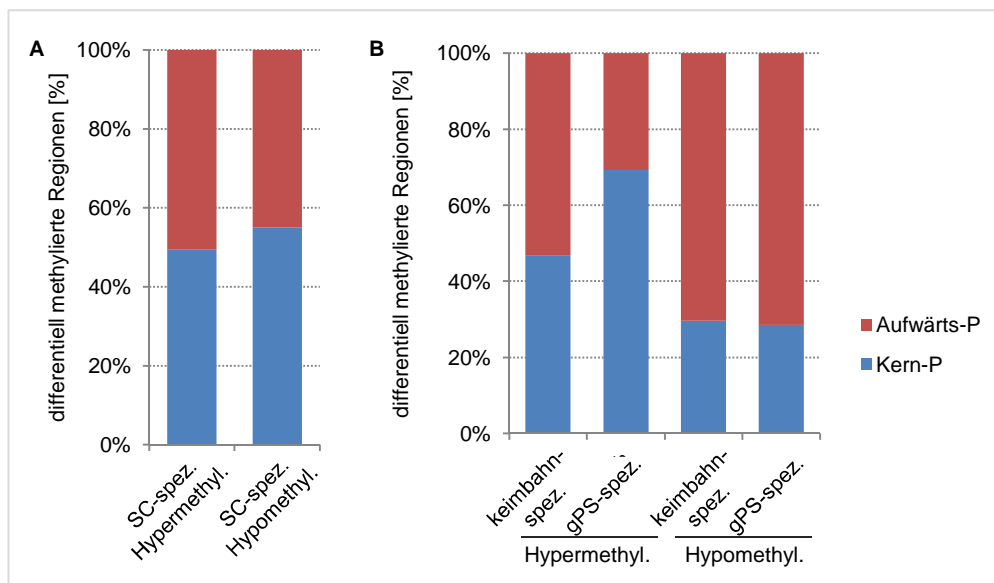


Abbildung 6.3-1: Differentielle Methylierung der gPS in Aufwärts- und Kern-Promotoren. Anteil differenziell methylierter Regionen in Kern- und Aufwärts-Promotoren mit (A) stammzellenspezifischer Hyper- bzw. Hypomethylierung gegenüber GSC und (B) keimbahn- und gewebespezifischer Hyper- bzw. Hypomethylierung der gPS gegenüber GSC und ESC. Hypomethylierte Regionen finden sich dabei vermehrt in Aufwärts-Promotoren.

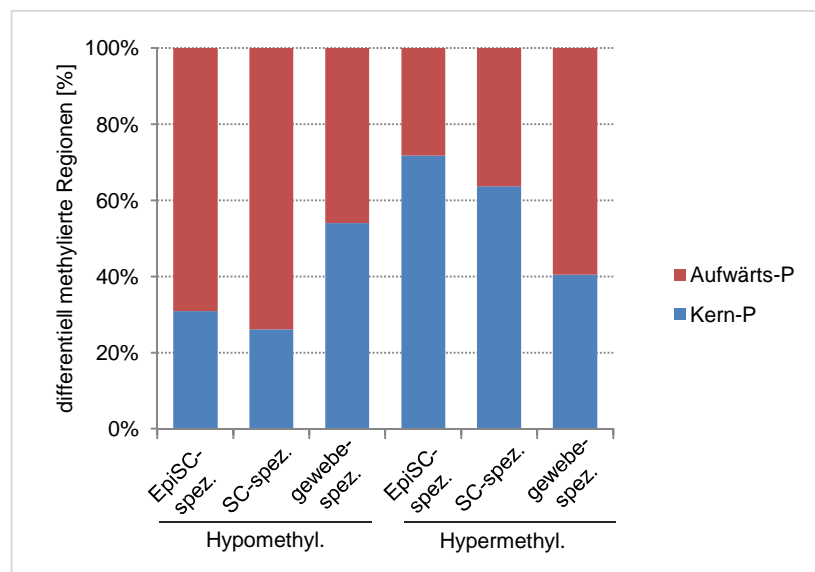


Abbildung 6.3-2: Differentielle Methylierung der EpiSC in Aufwärts- und Kern-Promotoren. Anteil differenziell methylierter Regionen in Kern- und Aufwärts-Promotoren mit EpiSC-, stammzell- und gewebespezifischer Hyper- bzw. Hypomethylierung der EpiSC. Hypomethylierte Regionen der Stammzellen finden sich vor allem in Aufwärts-Promotoren.

Tabelle 6.3-1 : GO-Annotationen überrepräsentierter biologischer Prozesse für hypomethylierte Gene der gPS. Gene mit statistisch signifikanter Hypomethylierung der gPS im Vergleich zu ESC (und GSC) wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Phosphorylierung		
GO:0006468-protein amino acid phosphorylation	1,4E-03	30
GO:0016310-phosphorylation	3,6E-03	31
GO:0006793-phosphorus metabolic process	5,1E-03	35
GO:0006796-phosphate metabolic process	5,1E-03	35
Regulation Phosphat-metabolischer Prozesse		
GO:0042325-regulation of phosphorylation	3,7E-03	17
GO:0051174-regulation of phosphorus metabolic process	5,2E-03	17
GO:0019220-regulation of phosphate metabolic process	5,2E-03	17
Sexuelle Reproduktion		
GO:0007276-gamete generation	3,7E-03	18
GO:0048232-male gamete generation	5,5E-03	15
GO:0007283-spermatogenesis	5,5E-03	15
GO:0032504-multicellular organism reproduction	7,4E-03	20
GO:0048609-reproductive process in a multicellular organism	7,4E-03	20
GO:0019953-sexual reproduction	1,7E-02	18
Zelluläre katabole Prozesse		
GO:0044265-cellular macromolecule catabolic process	1,6E-03	28
GO:0009057-macromolecule catabolic process	4,5E-03	28
GO:0051603-proteolysis involved in cellular protein catabolic process	1,7E-02	22
GO:0044257-cellular protein catabolic process	1,8E-02	22
GO:0043632-modification-dependent macromolecule catabolic process	1,8E-02	21
GO:0019941-modification-dependent protein catabolic process	1,8E-02	21
GO:0030163-protein catabolic process	2,6E-02	22

Tabelle 6.3-2 : GO-Annotationen überrepräsentierter biologischer Prozesse Gene mit keimbahn-spezifischer Hypomethylierung der gPS. Gene mit statistisch signifikanter Hypomethylierung der gPS und GSC im Vergleich zu ESC wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Zelluläre katabole Prozesse		
GO:0044265-cellular macromolecule catabolic process	5,1E-03	13
GO:0009057-macromolecule catabolic process	9,2E-03	13
GO:0051603-proteolysis involved in cellular protein catabolic process	1,4E-02	11
GO:0044257-cellular protein catabolic process	1,4E-02	11
GO:0030163-protein catabolic process	1,8E-02	11
GO:0019941-modification-dependent protein catabolic process	2,5E-02	10
GO:0043632-modification-dependent macromolecule catabolic process	2,5E-02	10
GO:0006508-proteolysis	9,5E-02	14
Imprintete Gene		
paternal imprint (Mest, Peg3, Plagl1, etc.)		12
isoform dependent (Gnas)		1
maternal imprint (Zim1, Commd1)		2

Tabelle 6.3-3 : GO-Annotationen überrepräsentierter biologischer Prozesse für Gene hypomethylierte Gene der gPS, exklusive keimbahn-spezifischer Muster. Gene mit statistisch signifikanter Hypomethylierung der gPS im Vergleich zu ESC und GSC (ohne Gene mit keimbahn-spezifischen Mustern) wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Phosphorylierung		
GO:0006468-protein amino acid phosphorylation	4,5E-03	21
GO:0006793-phosphorus metabolic process	8,0E-03	25
GO:0006796-phosphate metabolic process	8,0E-03	25
GO:0016310-phosphorylation	1,5E-02	21
Gewebe-Entwicklung		
GO:0060429-epithelium development	8,9E-03	12
GO:0002009-morphogenesis of an epithelium	1,2E-02	9
GO:0048729-tissue morphogenesis	6,1E-02	9
Organell-Lokalisation		
GO:0006903-vesicle targeting	6,9E-03	3
GO:0051650-establishment of vesicle localization	1,7E-02	3
GO:0051648-vesicle localization	2,4E-02	3
GO:0051640-organelle localization	6,0E-02	4
GO:0051656-establishment of organelle localization	1,1E-01	3

Tabelle 6.3-4: GO-Annotationen überrepräsentierter biologischer Prozesse für hypermethylierte Gene pluripotenter Zellen. Gene mit statistisch signifikanter Hypermethylierung der gPS und ESC im Vergleich zu GSC wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Sexuelle Reproduktion		
GO:0019953-sexual reproduction	8,4E-03	29
GO:0007283-spermatogenesis	1,3E-02	21
GO:0048232-male gamete generation	1,3E-02	21
GO:0007276-gamete generation	2,4E-02	24
GO:0032504-multicellular organism reproduction	5,0E-02	27
GO:0048609-reproductive process in a multicellular organism	5,0E-02	27

Tabelle 6.3-5: GO-Annotationen überrepräsentierter biologischer Prozesse für hypomethylierte Gene pluripotenter Zellen. Gene mit statistisch signifikanter Hypomethylierung der gPS und ESC im Vergleich zu GSC wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Antwort des Immun-Systems		
GO:0051249-regulation of lymphocyte activation	8,4E-04	22
GO:0002694-regulation of leukocyte activation	2,1E-03	22
GO:0050865-regulation of cell activation	2,6E-03	22
Regulation der Immun-Antwort		

GO:0002764-immune response-regulating signal transduction	4,8E-03	11
GO:0002757-immune response-activating signal transduction	8,2E-03	10
GO:0002253-activation of immune response	4,8E-02	12
Aktivierung der Immun-Antwort		
GO:0002757-immune response-activating signal transduction	8,2E-03	10
GO:0050851-antigen receptor-mediated signaling pathway	1,6E-02	8
GO:0002768-immune response-regulating cell surface receptor signaling pathway	1,6E-02	9
GO:0002429-immune response-activating cell surface receptor signaling pathway	3,2E-02	8

Tabelle 6.3-6: GO-Annotationen überrepräsentierter biologischer Prozesse für EpiSC-spezifische Hypermethylierung. Gene mit statistisch signifikanter Hypermethylierung der EpiSC im Vergleich zu ESC und MEF wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Ionen Homeostase		
GO:0006873-cellular ion homeostasis	5,6E-03	11
GO:0055082-cellular chemical homeostasis	6,8E-03	11
GO:0050801-ion homeostasis	1,2E-02	11
GO:0019725-cellular homeostasis	3,3E-02	11
GO:0048878-chemical homeostasis	5,3E-02	11
GO:0042592-homeostatic process	1,7E-01	13

Tabelle 6.3-7: GO-Annotationen überrepräsentierter biologischer Prozesse für gewebespezifische Hypermethylierung. Gene mit statistisch signifikanter Hypermethylierung der EpiSC und MEF im Vergleich zu ESC wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Meiose		
GO:0007127-meiosis I	3,10E-05	7
GO:0051327-M phase of meiotic cell cycle	2,95E-04	9
GO:0007126-meiosis	2,95E-04	9
GO:0051321-meiotic cell cycle	3,59E-04	9
Zell-Zyklus Phase		
GO:0000279-M phase	2,39E-04	17
GO:0022403-cell cycle phase	4,50E-04	18
GO:0022402-cell cycle process	6,34E-04	20
Reproduktion		
GO:0019953-sexual reproduction	9,74E-03	17
GO:0007283-spermatogenesis	2,54E-02	12
GO:0048232-male gamete generation	2,54E-02	12
GO:0007276-gamete generation	2,68E-02	14
GO:0032504-multicellular organism reproduction	3,41E-02	16
GO:0048609-reproductive process in a multicellular organism	3,41E-02	16

Tabelle 6.3-8: GO-Annotationen überrepräsentierter biologischer Prozesse für EpiSC-spezifische Hypomethylierung. Gene mit statistisch signifikanter Hypomethylierung der EpiSC im Vergleich zu ESC und MEF wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Immunsystem- und Organ-Entwicklung		
GO:0048534-hemopoietic or lymphoid organ development	5,6E-04	10
GO:0002520-immune system development	8,1E-04	10
GO:0030097-hemopoiesis	5,0E-03	8
Regionalisierung		
GO:0009952-anterior/posterior pattern formation	2,2E-03	7
GO:0003002-regionalization	3,0E-03	8
GO:0007389-pattern specification process	3,4E-03	9

Tabelle 6.3-9: GO-Annotationen überrepräsentierter biologischer Prozesse für stammzellspezifische Hypomethylierung. Gene mit statistisch signifikanter Hypomethylierung der EpiSC und ESC im Vergleich zu MEF wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Regulation RNA-metaboler Prozesse		
GO:0051252-regulation of RNA metabolic process	1,5E-10	64
GO:0006355-regulation of transcription, DNA-dependent	2,0E-10	63
GO:0045449-regulation of transcription	2,3E-08	77
Entwicklung des Organismus/Regionalisierung		
GO:0009952-anterior/posterior pattern formation	1,2E-08	18
GO:0003002-regionalization	1,6E-08	21
GO:0007389-pattern specification process	6,7E-08	23
Entwicklung des embryonalen Skeletts		
GO:0048704-embryonic skeletal system morphogenesis	5,6E-09	13
GO:0048706-embryonic skeletal system development	1,1E-08	14
GO:0048562-embryonic organ morphogenesis	1,9E-05	14
Negative Regulation der Gen-Expression		
GO:0000122-negative regulation of transcription from RNA polymerase II promoter	2,6E-08	21
GO:0045892-negative regulation of transcription, DNA-dependent	3,8E-08	24
GO:0051253-negative regulation of RNA metabolic process	4,4E-08	24
GO:0051172-negative regulation of nitrogen compound metabolic process	7,8E-08	27
GO:0010629-negative regulation of gene expression	1,2E-07	27
GO:0045934-negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	2,7E-07	26
GO:0016481-negative regulation of transcription	2,8E-07	25
GO:0031327-negative regulation of cellular biosynthetic process	3,4E-07	27
GO:0009890-negative regulation of biosynthetic process	4,1E-07	27
GO:0010558-negative regulation of macromolecule biosynthetic process	7,0E-07	26
GO:0010605-negative regulation of macromolecule metabolic process	7,2E-07	29
Positive Regulation der Gen-Expression		
GO:0010557-positive regulation of macromolecule biosynthetic process	2,5E-06	29
GO:0045941-positive regulation of transcription	3,3E-06	27
GO:0045935-positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	3,9E-06	28
GO:0010628-positive regulation of gene expression	4,9E-06	27
GO:0031328-positive regulation of cellular biosynthetic process	5,3E-06	29
GO:0009891-positive regulation of biosynthetic process	6,2E-06	29
GO:0051173-positive regulation of nitrogen compound metabolic process	6,4E-06	28
GO:0010604-positive regulation of macromolecule metabolic process	8,3E-06	31
GO:0045893-positive regulation of transcription, DNA-dependent	1,2E-05	24

GO:0045944-positive regulation of transcription from RNA polymerase II promoter	1,2E-05	22
GO:0051254-positive regulation of RNA metabolic process	1,3E-05	24
Protein-/Histon-Modifikation		
GO:0016577-histone demethylation	2,7E-03	3
GO:0008214-protein amino acid dealkylation	4,5E-03	3
GO:0006482-protein amino acid demethylation	4,5E-03	3
Stamzell-Entwicklung		
GO:0048863-stem cell differentiation	2,2E-03	5
GO:0019827-stem cell maintenance	7,5E-03	4
GO:0048864-stem cell development	8,7E-03	4
Entwicklung anatomischer Strukturen		
GO:0060173-limb development	8,6E-03	8
GO:0048736-appendage development	8,6E-03	8
GO:0035108-limb morphogenesis	2,5E-02	7
GO:0035107-appendage morphogenesis	2,5E-02	7
GO:0030326-embryonic limb morphogenesis	5,1E-02	6
GO:0035113-embryonic appendage morphogenesis	5,1E-02	6

Tabelle 6.3-10: GO-Annotationen überrepräsentierter biologischer Prozesse für dynamische Methylierungsunterschiede. Gene mit statistisch signifikanter ansteigender Methylierung ESC<EpiSC<MEF wurden mittels der DAVID-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Positive Regulation metabolischer Prozesse/RNA		
GO:0010604-positive regulation of macromolecule metabolic process	1,3E-04	29
GO:0045935-positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1,9E-04	25
GO:0051173-positive regulation of nitrogen compound metabolic process	2,8E-04	25
GO:0045941-positive regulation of transcription	4,3E-04	23
GO:0031328-positive regulation of cellular biosynthetic process	5,6E-04	25
GO:0010628-positive regulation of gene expression	5,9E-04	23
GO:0009891-positive regulation of biosynthetic process	6,3E-04	25
GO:0010557-positive regulation of macromolecule biosynthetic process	7,6E-04	24
GO:0045893-positive regulation of transcription, DNA-dependent	3,2E-03	19
GO:0051254-positive regulation of RNA metabolic process	3,5E-03	19
GO:0045944-positive regulation of transcription from RNA polymerase II promoter	4,1E-03	17
Entwicklung embryonaler Organe/Ohr		
GO:0043583-ear development	2,0E-04	10
GO:0042471-ear morphogenesis	7,8E-04	8
GO:0048839-inner ear development	1,6E-03	8
GO:0042472-inner ear morphogenesis	1,9E-03	7
Negative Regulation der Transcription		
GO:0000122-negative regulation of transcription from RNA polymerase II promoter	2,9E-03	13
GO:0045892-negative regulation of transcription, DNA-dependent	4,6E-03	15
GO:0051253-negative regulation of RNA metabolic process	4,9E-03	15
GO:0031327-negative regulation of cellular biosynthetic process	7,7E-03	18
GO:0009890-negative regulation of biosynthetic process	8,5E-03	18
GO:0010605-negative regulation of macromolecule metabolic process	8,5E-03	20
GO:0016481-negative regulation of transcription	9,8E-03	16
GO:0010629-negative regulation of gene expression	1,0E-02	17
GO:0010558-negative regulation of macromolecule biosynthetic process	1,3E-02	17
GO:0045934-negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1,7E-02	16
GO:0051172-negative regulation of nitrogen compound metabolic process	1,8E-02	16
Zellentwicklung		
GO:0000904-cell morphogenesis involved in differentiation	4,4E-03	12
GO:0000902-cell morphogenesis	1,1E-02	14
GO:0032989-cellular component morphogenesis	2,6E-02	14
Organ-Entwicklung		
GO:0060485-mesenchyme development	4,1E-03	6
GO:0014031-mesenchymal cell development	1,7E-02	5

GO:0048762-mesenchymal cell differentiation	2,0E-02	5
GO:0014032-neural crest cell development	3,0E-02	4
GO:0014033-neural crest cell differentiation	3,0E-02	4

Tabelle 6.3-11: GO-Annotationen überrepräsentierter biologischer Prozesse für dynamische Methylierungsunterschiede. Gene mit statistisch signifikanter ansteigender Methylierung EpiSC<ESC<MEF wurden mittels der *DAVID*-Software (Huang da et al. 2009) gegen die Gesamtheit der im *Mm8 Promotor-Array* abgedeckten Gene als Referenz evaluiert, um überrepräsentierte biologische Prozesse (GO-Ontologie) zu ermitteln. Kategorien ähnlicher biologischer Bedeutung wurden dabei durch das Programm in Gruppen zusammengefasst. Zur besseren Übersicht wurden die in den Abbildungen verwendeten Oberbegriffe der jeweiligen Kategorien mit angegeben.

Kategorie	p-Wert	Anzahl
Phosphat-Metabolismus		
GO:0016310-phosphorylation	6,4E-03	26
GO:0006793-phosphorus metabolic process	1,0E-02	29
GO:0006796-phosphate metabolic process	1,0E-02	29
GO:0006468-protein amino acid phosphorylation	1,2E-02	23
Protein-Katabolismus		
GO:0051603-proteolysis involved in cellular protein catabolic process	8,7E-03	20
GO:0044257-cellular protein catabolic process	9,3E-03	20
GO:0044265-cellular macromolecule catabolic process	9,3E-03	22
GO:0043632-modification-dependent macromolecule catabolic process	1,0E-02	19
GO:0019941-modification-dependent protein catabolic process	1,0E-02	19
GO:0009057-macromolecule catabolic process	1,1E-02	23
GO:0030163-protein catabolic process	1,3E-02	20

Tabelle 6.3-12: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für hypomethylierte Regionen der gPS, exklusive keimbahn-spezifische Muster. Siehe Tabelle 6.3-3 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Phosphorylierung	Cdkl1	Mapk9	Prkcn	Grm1
	Fgfr1	Map2k6	Prkacb	Grk1
	6230410P16Rik	Map4k2	Ptprn	Smtnl1
	Inpp5a	Map4k4	Ptprn2	Trp53rk
	Insr	Osm	Ptprs	Tgfa
	Lrrk1	Pml	Ror2	Tie1
Gewebe-Entwicklung	Raf1	Pou2f3	Dlgh1	Pthlh
	Elf5	Tbx18	Hhex	Sfn
	Notch2	Car9	Hoxa5	Pml
Organell-Lokalisation	Scrib	Map4k2		
	Hhex	Scrib		

Tabelle 6.3-13: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für Regionen mit keimbahn-spezifischer Hypomethylierung der gPS. Siehe Tabelle 6.3-2 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Zelluläre katabole Prozesse	Tmed10	2610207I05Rik	Klhl9	Trim33
	Edem3	Adam10	Peg10	Usp12
	Fbxw4	Dcps	Gm1019	Usp15
	Ibrdc2	Klk8	Rnf11	Usp16
				Usp29

Tabelle 6.3-14: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für pluripotenzspezifische Hypomethylierung. Siehe Tabelle 6.3-5 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Aktivierung der Immun-Antwort				
	Tnfrsf7	Gal	Lck	Stat6
	Cd38	Glmn	Lag3	Ticam1
	Cd40	Il27ra	AY078069	Tnsfs13
	Cd80	Il6	Prdx2	Zap70
	Nfam1	Lst1	Hsp90aa1	
	Casp3	Lat	Ptpn6	
Regulation der Immun-Antwort				
	Cd247	Nfam1	Lax1	Zap70
	Cd40	C1r	Card15	
	Cd79a	C2	Ptpn6	
	Cd79b	Lck	Ticam1	
Antwort des Immun-Systems				
	Cd247	Cd79b	Lax1	Ticam1
	Cd40	Nfam1	Card15	Zap70
	Cd79a	Lck	Ptpn6	

Tabelle 6.3-15: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für pluripotenzspezifische Hypermethylierung. Siehe Tabelle 6.3-4 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Sexuelle Reproduktion				
	Nobox	Eif2ak3	Pebp1	Sohlh2
	Adam3	Fabp9	Nhp2l1	Spata16
	Asz1	Gykl1	Psme4	Spata5
	Bmpr1b	Glrh	Rbp4	Sycp3
	Cldn11	Hook1	Rnf17	Zfp39
	Dazl	Kdr	Slco4c1	Zfa
	Dld	Mael	Spag16	Zpbp
	Esr1	Nkd1	Sohlh1	Zpbp2

Tabelle 6.3-16: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für EpiSC-spezifische Hypermethylierung. Siehe Tabelle 6.3-6 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol
Ionen Homeostase			
	Gpr12	Hspa1a	Rxfp3
	Rhcg	Hspa1b	Grin2a
	Cacng2	Hexa	Ckb
	Col11a2	Ndn	Ttc7
	Grik2	Gpsn2	

Tabelle 6.3-17: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für gewebespezifische Hypermethylierung. Siehe Tabelle 6.3-7 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Meiose				
	Dmc1h	Ehmt2	Sycp1	
	Rad51c	Mael	Topbp1	
	Adcy3	Msh4	Tdrd1	
Zell-Zyklus Phase				
	Dmc1h	Adcy3	Krt2-7	Ube2i
	Hormad1	Anapc13	Mael	Sycp1
	Rad51c	Cdkn1c	Msh4	Topbp1

	Ran	Ehmt2	Luzp5	Tdrd1
	Crebl1	C79407	Pkd2	Usp16
Reproduktion	Dmc1h	Acox1	Hsf1	Tdrd1
	D1Pas1	Tenr	Mael	Zfa
	Rad51c	Dnd1	Msh4	Zp3
	Tial1	Ehmt2	Tssk3	
	B4galt1	Eif2b4	Trim36	

Tabelle 6.3-18: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für EpiSC-spezifische Hypomethylierung. Siehe Tabelle 6.3-8 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol
Immunsystem- und Organ-Entwicklung		
	Atp6v0a1	Jag2
	Bax	Kdr
	Pou2f2	Pik3r1
	Hoxa3	Tcf21
	Hoxb4	Tnfsf11
Regionalisierung		
	Nkx6-2	Hoxa6
	Tbx18	Hoxa7
	Gna13	Hoxb4
	Hes7	Wnt3
	Hoxa3	

Tabelle 6.3-19: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für stammzellspezifische Hypomethylierung. Siehe Tabelle 6.3-9 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Regulation RNA-metabolischer Prozesse				
	Aff3	Cnbp1	Hoxa5	Nr2f2
	Bcl6b	Cbx4	Hoxa7	Mysm1
	Bcl3	Axud1	Hoxb4	Mafg
	Bmi1	Ebf2	Hoxb5	Sim2
	Cebpe	Ebf3	Hoxb8	Tcea3
	2610101J03Rik	En1	Hoxd11	Tcfap2a
	Hlx1	Fubp3	Hoxd12	Ube2b
	Htatip	Foxc2	Msx1	Zbtb38
	Jmjd3	Foxh1	Lef1	Zfp206
	Khsrp	Gtf2e2	Ntf3	Atbf1
	Lhx9	Gtf2h4	Nodal	Zfp13
	Lmx1b	Gtf2f1	Ncoa5	Zfp219
	Nkx2-3	Gcm1	Nfyc	Zfp27
	Pdlim1	Grlf1	Pax7	Zfp35
	1700023B02Rik	Gsc	Pitx1	Zfp36
	Sox15	Hand1	Pnn	Zfp369
	Sox6	Hsf4	Smarca5	Zfp503
	Thap7	Hells	Tdgf1	Zfp93
	Ctnnbip1	Hoxa2	Ybx1	
	Cdx1	Hoxa4	Pml	
Entwicklung des Organismus/Regionalisierung				
	Bmi1	Foxh1	Hoxb4	Nodal
	Lmx1b	Gsc	Hoxb5	Pax7
	Ctnnbip1	Hoxa2	Hoxb8	Tdgf1
	Cdx1	Hoxa4	Hoxd11	Nr2f2
	En1	Hoxa5	Hoxd12	Sim2
	Foxc2	Hoxa7	Lef1	
Entwicklung des embryonalen Skeletts				
	Bmi1	Hoxa2	Hoxb4	Pax7
	Hlx1	Hoxa4	Hoxb5	Tcfap2a
	Foxc2	Hoxa5	Hoxb8	
	Gsc	Hoxa7	Nodal	
Negative Regulation der Gen-Expression				

Bcl6b	Cbx4	Hoxb4	Nr2f2
Bcl3	Foxh1	Msx1	Bnip3l
Bmi1	Gsc	Lef1	Dynll1
Jmjd3	Hand1	Msh6	Sim2
1700023B02Rik	Hsf4	Nodal	Ube2b
Sox15	Hells	Piwil2	Atbf1
Sox6	Hoxa2	Smarca5	
Ctnnbip1	Hoxa7	Pml	
Positive Regulation der Gen-Expression			
Bcl3	Axud1	Hsf4	2610009E16Rik
Htatif	Ebf2	Hoxa2	Piwil2
Lmx1b	Ebf3	Hoxa7	Tdgf1
Nkx2-3	Fubp3	Msx1	Pml
Sox15	Foxc2	Lef1	Mysm1
Sox6	Foxh1	Ntf3	Ube2b
Cdx1	Gtf2f1	Nodal	Zbtb38
Cnbp1	Hand1	Nfyc	
Protein-/Histone-Modifikation			
Jmjd3			
Hsf4			
Ube2b			
Stamzell-Entwicklung			
Rif1	Pax7		
Lig4	Piwil2		
Nodal			
Entwicklung anatomischer Strukturen			
Lmx1b	Hoxd12	Pitx1	
En1	Msx1	Nr2f2	
Hoxd11	Lef1		

Tabelle 6.3-20: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für dynamische Methylierungsunterschiede, EpiSC<ESC<MEF. Siehe Tabelle 6.3-11 zum Vergleich. Aufgrund der zahlreichen Ergebnisse werden lediglich die Gen-Symbole angegeben.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Positive Regulation metabolischer Prozesse RNA				
	Cart1	Ctnnb1	Hoxa2	Tbx1
	Bcl10	Cln6	Hoxd9	Pnp
	Klf12	Cdk2	Ntf3	Six3
	Klf5	Ebf3	Card15	Tcf7l2
	Smad3	Evx1	Pitx2	Rela
	Pawr	Eya1	Pbx2	
	Sox9	Foxc1	Pias1	
	Ablim2	Gtf2f1	Rfx4	
Entwicklung embryonaler Organe/Ohr				
	Bcl2l11	Eya1	Insig2	Tbx1
	Clic5	Gbx2	Mcoln3	
	Dlx5	Hoxa2	Muted	
Negative Regulation der Transkription				
	Cart1	Klf12	Insig2	Trim27
	Dnajb5	Smad3	Jarid2	Rela
	Zfp336	Sox9	Lin28b	Zbtb7a
	Gipc1	Ctnnb1	Rfx5	
	Zfpn1a4	Eif2ak4	Nr2f2	
	Jundm2	Hoxa2	Six3	
Zellentwicklung				
	Rab3a	Ctnnb1	Dlx5	Clu
	Slitrk5	Clic5	Gbx2	Ltap
	Sox9	Col18a1	Hoxa2	
	Artn	Cyfip1	Ntf3	
Organ-Entwicklung				
	Cart1	Foxc1	Hoxa5	
	Sox9	Gbx2	Tbx1	

Tabelle 6.3-21: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für dynamische Methylierungsunterschiede, ESC<EpiSC<MEF. Siehe Tabelle 6.3-10 zum Vergleich.

Prozess	Gen-Symbol	Gen-Symbol	Gen-Symbol	Gen-Symbol
Phosphat-Metabolismus				
	Atp6v0a2	Camkk2	Gpd1l	Map2k1ip1
	Chek2	Csnk2a1	Lck	Prkaca
	Ndufc2	Cfl1	0610009K11Rik	D830007F02Rik
	Nek4	Cdkl3	Map3k7	Ttn
	Yes1	Eif2ak4	Pak4	Vrk1
	Acp5	Gsg2	Prdx2	
	Acvr1	6230410P16Rik	Pik3ca	
	Asz1	Gpd2	Pik4cb	
Protein-Katabolismus				
	Cdc23	Asb15	Man1b1	Rbx1
	Edem3	3110023E09Rik	0610009K11Rik	Rabgef1
	1500002O20Rik	Fem1c	C6.1AL	5730410I19Rik
	Sugt1	Rnps1	Tmepai	Ube1l2
	Senp5	Fbxl10	5830406J20Rik	Uhrf2
	Wsb2	Mgrn1	2500002L14Rik	

6.3.2 Analyse der DNA-Methylierung mittels Bisulfit-Sequenzierung

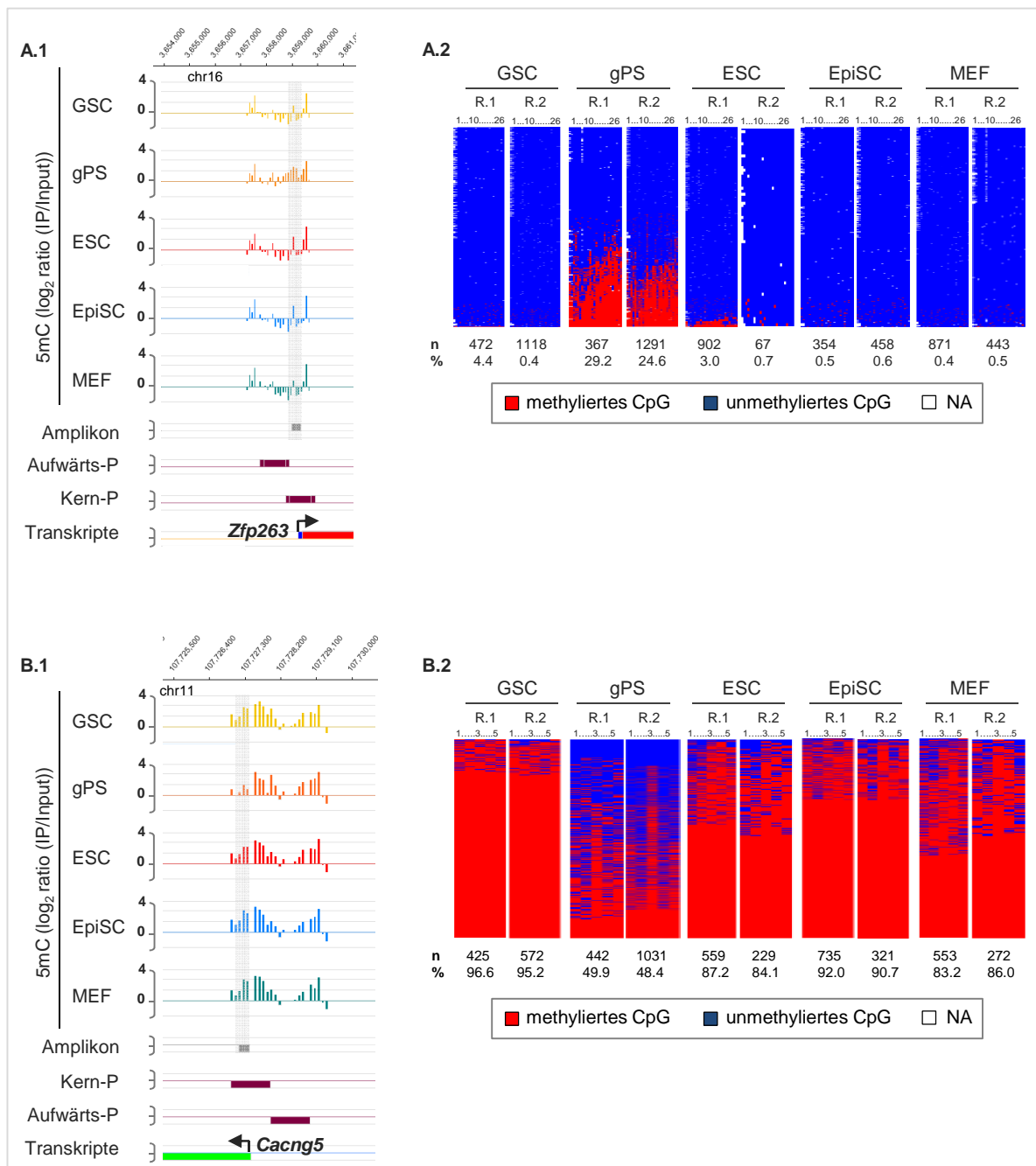


Abbildung 6.3-3: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Zfp263, Cacng5. (A.1, B.1) Anreicherungsprofile des Zfp263- und Cacng5-Promotors, mit \log_2 -Ratios relativ zu genomischen Koordinaten, funktionellen Regionen sowie Positionen der Bisulfit-Amplikons. (A.2, B.2) Resultate der Bisulfit-Sequenzierung dargestellt als Heatmaps der biologischen Replikate (R.1, R.2) pro Zelltyp mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen (n: Anzahl der analysierten Sequenzen, %: mittlere Bisulfit-Methylierung in %). Die Daten lassen eine gPS-spezifische Hyper- bzw. Hypomethylierung erkennen (Zfp263 in (A) bzw. Cacng5 in (B)).

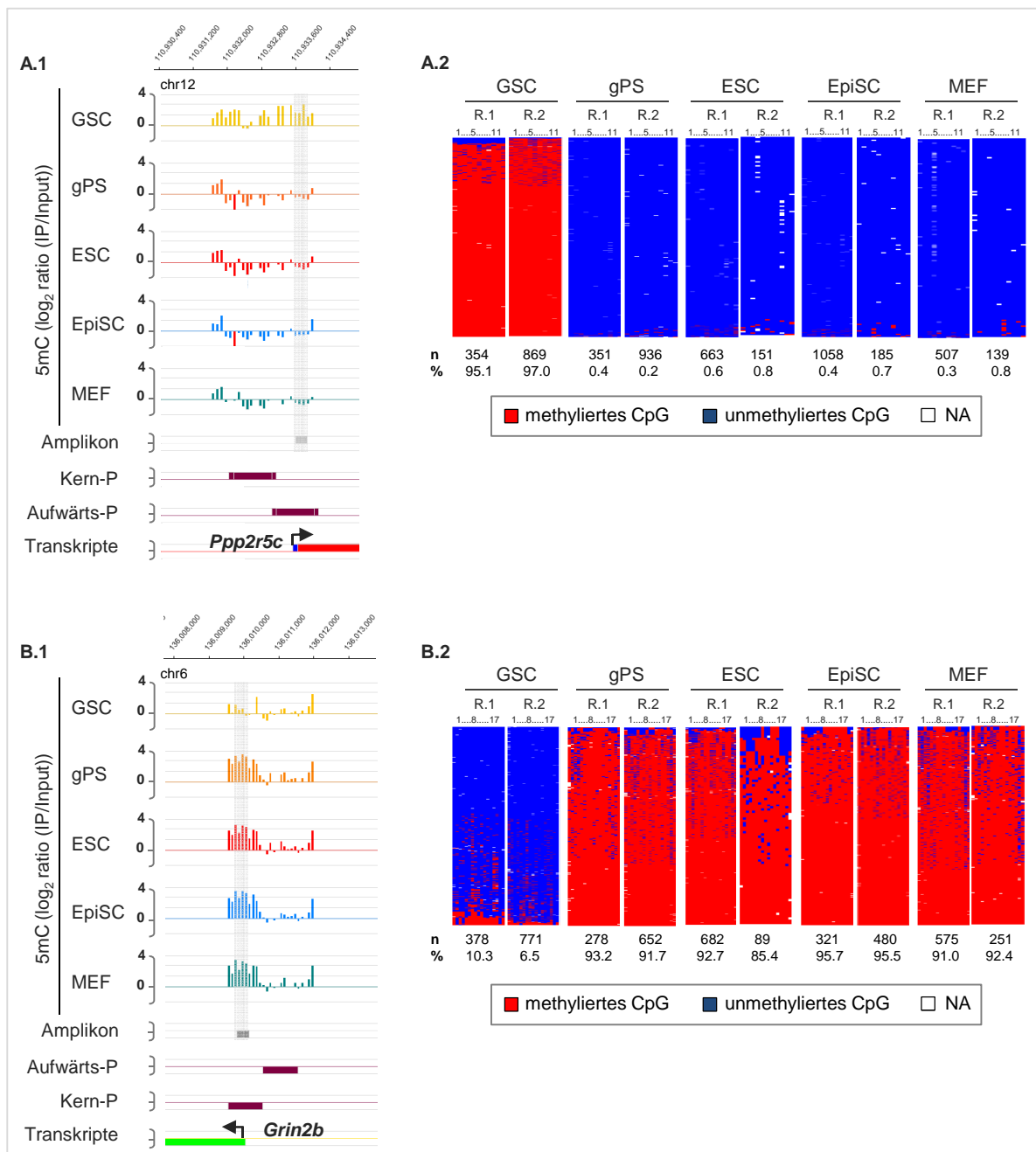


Abbildung 6.3-4: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Ppp2r5c, Grin2b. (A.1, B.1) Anreicherungsprofile des Ppp2r5c- und Grin2b-Promotors, mit log₂-Ratios relativ zu genomischen Koordinaten, funktionellen Regionen sowie Positionen der Bisulfit-Amplikons. (A.2, B.2) Resultate der Bisulfit-Sequenzierung dargestellt als Heatmaps der biologischen Replikate (R.1, R.2) pro Zelltyp mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen (n: Anzahl der analysierten Sequenzen, %: mittlere Bisulfit-Methylierung in %). Die Daten lassen eine GSC-spezifische Hyper- bzw. Hypomethylierung erkennen (Ppp2r5c in (A) bzw. Grin2b in (B)).

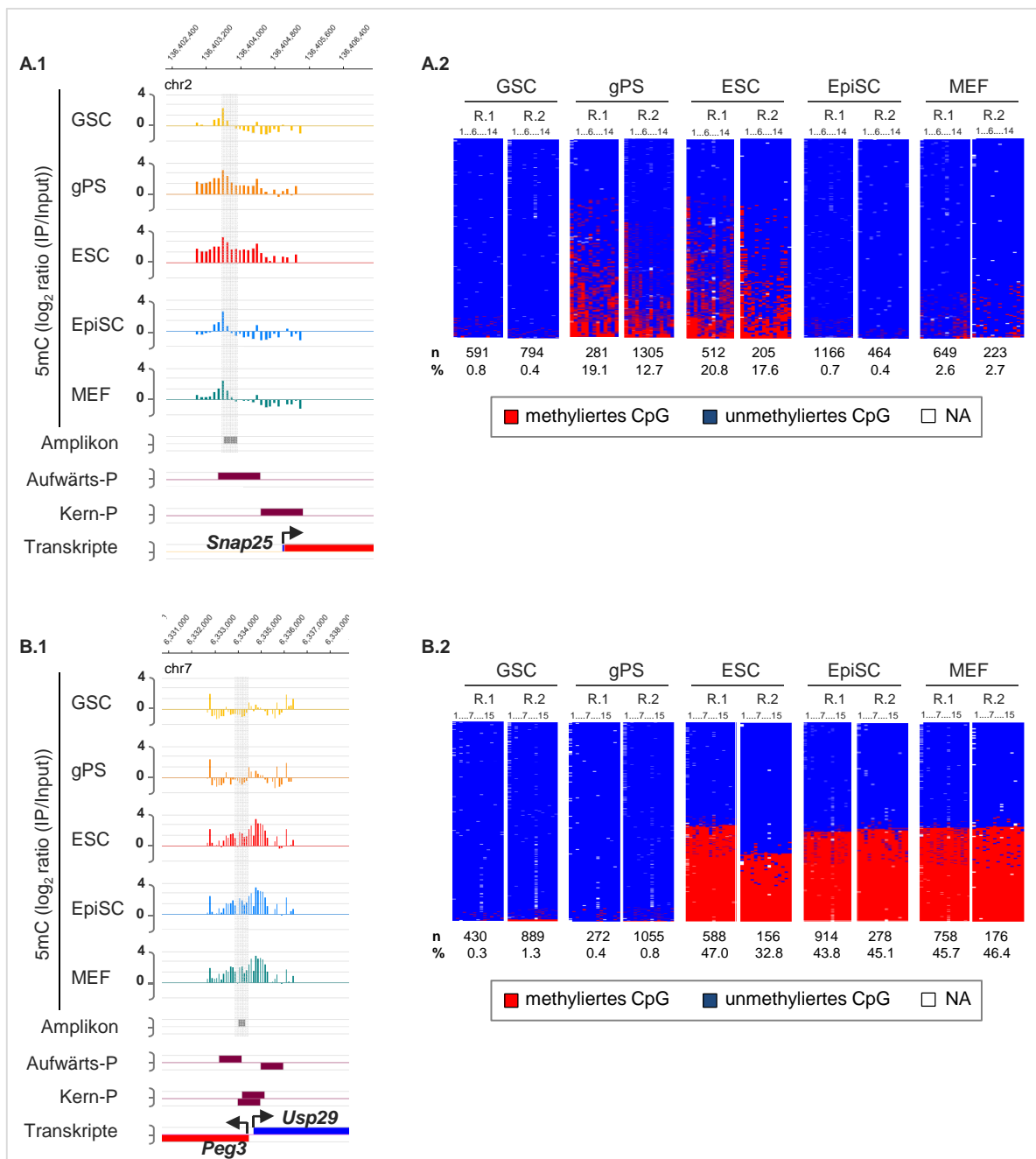


Abbildung 6.3-5: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Snap25, Peg3/Usp29. (A.1, B.1) Anreicherungsprofile des Snap25- und Peg3/Usp29-Promotors, mit \log_2 -Ratios relativ zu genomischen Koordinaten, funktionellen Regionen sowie Positionen der Bisulfit-Amplikons. (A.2, B.2) Resultate der Bisulfit-Sequenzierung dargestellt als Heatmaps der biologischen Replikate (R.1, R.2) pro Zelltyp mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen (n: Anzahl der analysierten Sequenzen, %: mittlere Bisulfit-Methylierung in %). Die Daten lassen eine Hypermethylierung pluripotenter ESC (und gPS) (Snap25 in (A)) bzw. eine keimbahnstetige Hypomethylierung der imprinteten Gene Peg3/Usp29 (B) erkennen.

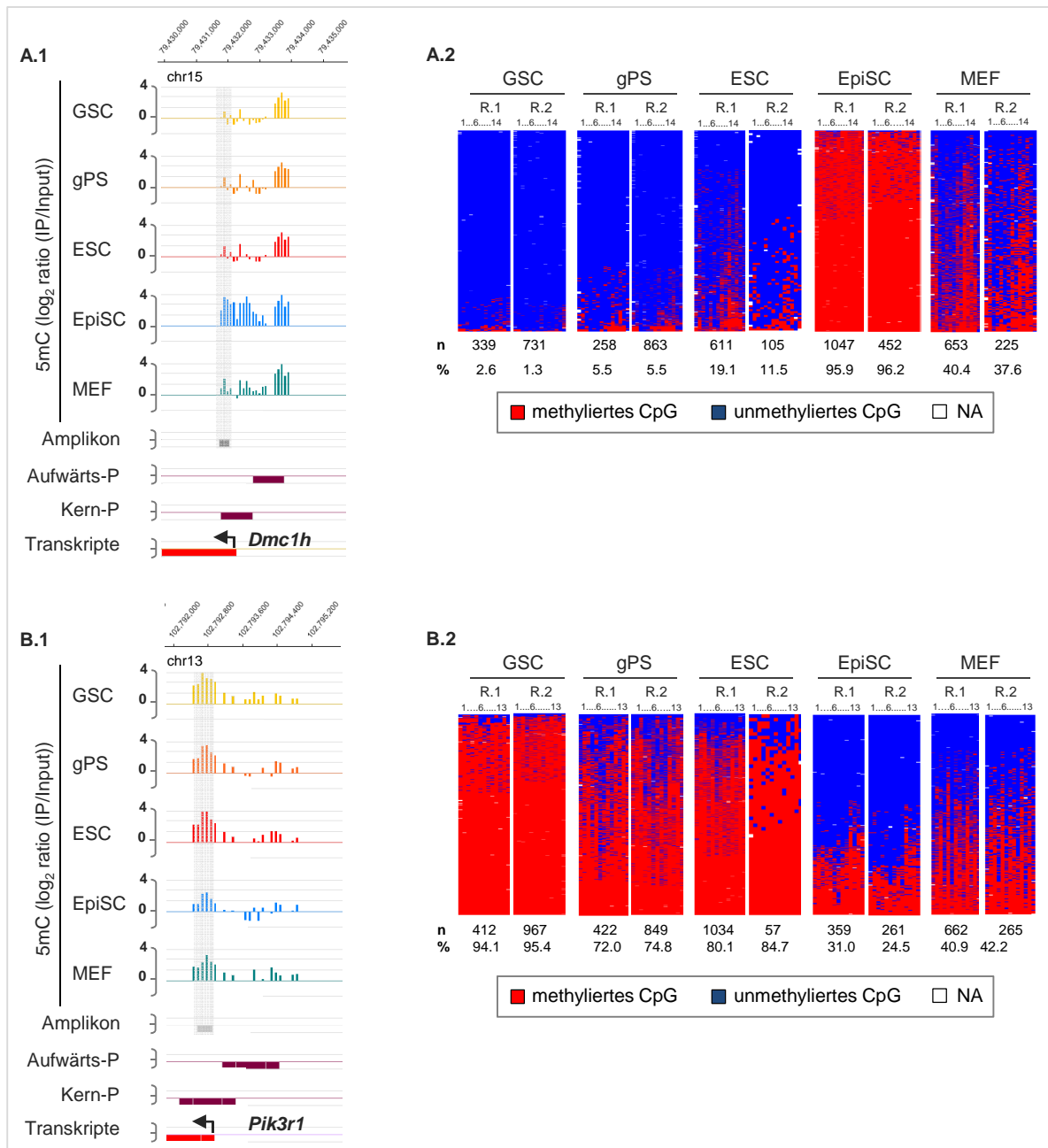


Abbildung 6.3-6: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: *Dmc1h*, *Pik3r1*. (A.1, B.1) Anreicherungsprofile des *Dmc1h*- und *Pik3r1*-Promotors, mit \log_2 -Ratios relativ zu genomischen Koordinaten, funktionellen Regionen sowie Positionen der Bisulfit-Amplikons. (A.2, B.2) Resultate der Bisulfit-Sequenzierung dargestellt als Heatmaps der biologischen Replikate (R.1, R.2) pro Zelltyp mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen (n: Anzahl der analysierten Sequenzen, %: mittlere Bisulfit-Methylierung in %). Die Daten lassen eine Hypermethylierung der EpiSC (und MEF) (*Dmc1h* in (A)) bzw. eine gewebespezifische Hypomethylierung (*Pik3r1* in (B)) erkennen.

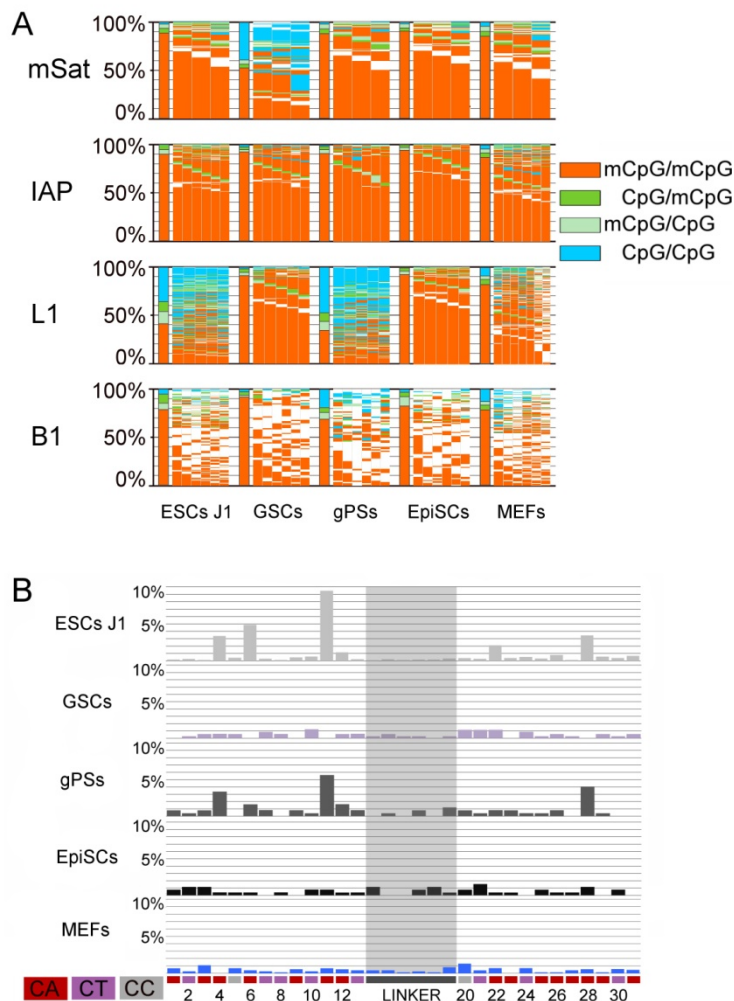


Abbildung 6.3-7: Methylierungsanalyse repetitiver Elemente mittels Bisulfit-Hochdurchsatz-Sequenzierung. Vier repetitive Elemente wurden in GSC, gPS, ESC(J1) und EpiSC mit der Hairpin-Bisulfit-Methode untersucht, wobei jeweils oberer und unterer Strang - verbunden über eine Linker-Sequenz - berücksichtigt wurden (Arand et al. in Revision). (A) Methylierungsmuster der repetitiven Elemente pro Zelltyp mit individuellen Sequenzen in der Vertikalen und einzelnen CpG-Positionen in der Horizontalen. gPS weisen ESC-ähnliche Muster auf, z.B. L1-Elemente mit der Verteilung methylierter und hemimethylierter (grün) CpGs. EpiSC sind in allen untersuchten Elementen ähnlich der MEF stark methyliert. (B) Asymmetrische Methylierung in mSat. gPS weisen positionsspezifische CA, CT- und CC-Methylierung ähnlich der ESC auf. EpiSC hingegen zeigen keine asymmetrische Methylierung. Abbildung und Daten stammen von der Doktorandin *Julia Arand*.

7 Abkürzungsverzeichnis

bp	Basenpaare
BSA	Rinderserumalbumin (aus dem engl.)
bzw.	beziehungsweise
°C	Celsius
C-	Carboxy-
5caC	5-Carboxylcytosin
ChIP	Chromatin-Immunopräzipitation
DMR	differentiell methylierte Region
DNA	Desoxyribonukleinsäure (aus dem engl.)
dNTPs	Desoxyribonukleosidtriphosphate
dsDNA	doppelsträngige DNA
<i>E. coli</i>	<i>Escherichia coli</i>
engl.	englisch
EpiSC	Epiblast-Stammzelle (aus dem engl.)
ESC	embryonale Stammzelle (aus dem engl.)
5fC	5-Formylcytosin
g	Gramm
gDNA	genomische DNA
gPS	aus der Keimbahn abstammende pluripotente Zelle (germline pluripotent cell)
GSC	Keimbahnstammzelle (Germline stem cell)
h	Stunde (hour)
5hmC	5-Hydroxymethylcytosin
H ₂ O	Wasser
Ig	Immunglobulin
iPSC	induzierte pluripotente Stammzelle (aus dem engl.)
kb	Kilobase
l	Liter
LB	lysogeny broth
M	Molar
Mb	Megabase
5mC	5-Methylcytosin
m	Milli
μ	mikro
MeDIP	Immunopräzipitation methylierter DNA (Methylated DNA Immunoprecipitation)
MEF	embryonale Fibroblasten der Maus (Mouse embryonic fibroblast)
min	Minute
miRNA	microRNA (micro Ribonukleinsäure)
mRNA	Boten-Ribonukleinsäure (messenger Ribonucleic acid)
n	nano
N-	Amino-
ncRNA	nicht-kodierende RNA (aus dem engl.)

nm	Nanometer
nt	Nukleotide
OD	optische Dichte
OH	Hydroxyl-
ORF	open reading frame (offener Leserahmen)
ori	origin of replication (Replikationsursprung)
P	Promotor
PCR	Polymerasekettenreaktion (aus dem engl.)
pH	potentia Hydrogenii
PBS	Phosphat gepufferte Salzlösung (aus dem engl.)
RNA	Ribonukleinsäure (aus dem engl.)
RNase	Ribonuklease
ROS	Sauerstoffradikale (reactive oxygen species)
rpm	Umdrehungen pro Minute (revolutions per minute)
RT	Raumtemperatur
SDS	Sodiumdodecylsulfat
sec	Sekunde
SNP	Einzelbasen-Polymorphismus (Single Nucleotide Polymorphism)
ssDNA	einzelsträngige DNA
<i>Taq</i>	<i>Thermus aquaticus</i>
TF	Transkriptionsfaktor
TSS	Transkriptionsstartstelle
U	Einheit (Unit)
UV	ultraviolett
3' UTR	3' untranslatierte Region
5' UTR	5' untranslatierte Region
W	Watt
WGA	genomrepräsentative Amplifikation (Whole Genome Amplification)
w/v	Gewicht pro Volumen (weight per volume)
v	Volumen
V	Volt

8 Abbildungsverzeichnis

Abbildung 1.1-1: DNA-Methylierung in Säugern.....	3
Abbildung 1.1-2: Schematische Darstellung des Histon-Codes.	6
Abbildung 1.1-3: Auflösung bivalenter Domänen während der Differenzierung von ESC.	9
Abbildung 1.2-1: (Pluripotente) Stammzellen,.....	10
Abbildung 1.2-2: Etablierung einer Zellkulturlinie aus embryonalen Stammzellen.	11
Abbildung 1.2-3: Stammzellen des Epiblasten im Vergleich zu ESC.	15
Abbildung 1.3-1: Zellschädigung in Abhängigkeit der Kühlrate bei Kryokonservierung.	17
Abbildung 1.4-1: Oxidativer Stress, zelluläre Schutzmechanismen und DNA-Methylierung.	20
Abbildung 2.2-1: Prinzip der Immunopräzipitation methylierter DNA (MeDIP).	44
Abbildung 2.2-2: Sequenzierung Bisulfit-behandelter DNA.....	48
Abbildung 2.2-3: Übersicht der aufeinanderfolgenden Teilschritte zur Analyse der MeDIP-Chip-Daten	52
Abbildung 3.1-1: Prinzip der Immunopräzipitation methylierter DNA (MeDIP).	64
Abbildung 3.1-2: Relative Anreicherung verschiedener Kontrollregionen in Abhängigkeit des Methylierungslevels und CpG-Gehalts.	65
Abbildung 3.1-3: Vergleich verschiedener Fragmentierungsmethoden.	66
Abbildung 3.1-4: Anreicherung in Abhängigkeit des AK:DNA-Verhältnisses.	68
Abbildung 3.1-5: Effekt der genomrepräsentativen Amplifikation.	69
Abbildung 3.1-6: Ablauf der Datenanalyse.	71
Abbildung 3.1-7: Diagnostische Plots der Qualitätskontrolle.....	73
Abbildung 3.1-8: Effekt verschiedener Prä-Prozessierungs-Ansätze,	75
Abbildung 3.1-9: Graphische Darstellung der Anreicherungsprofile zweier humaner Zelltypen.	77
Abbildung 3.1-10: ROC-Analyse zur Überprüfung der Peak-Detektion.	79
Abbildung 3.1-11: Überprüfung der Peak-Analyse zur Detektion differentieller Anreicherung.....	80
Abbildung 3.1-12: Einfluss der genomrepräsentativen Amplifikation auf die Anreicherung methylierter Regionen.	82
Abbildung 3.1-13: Vergleich der MeDIP-Chip-Daten unabhängiger Experimente des gleichen Zelltyps (Leukozyten).	84
Abbildung 3.2-1: Schematische Darstellung des Kryopreservations-Experiments,	86
Abbildung 3.2-2: Überprüfung der HepG2-Präzipitate vor und nach WGA.	87
Abbildung 3.2-3: Vergleich der HepG2-Proben untereinander auf Basis der Oligo-log ₂ -Ratios.	89
Abbildung 3.2-4: Anreicherungsprofile der HepG2-Proben.	90
Abbildung 3.2-5: Anreicherungsunterschiede nachfolgend Kryopreservation von HepG2-Zellen.	91
Abbildung 3.2-6: Unsystematische Anreicherungsunterschiede nachfolgend Kryopreservation von HepG2-Zellen.	92
Abbildung 3.2-7: Änderungen der DNA-Methylierung bzw. mRNA-Expression nachfolgend Kryopreservation betreffen unterschiedliche Gene.	93
Abbildung 3.2-8: Änderungen der Expression nach DMSO-Behandlung bzw. Kryopreservation.	94
Abbildung 3.2-9: Überrepräsentierte biologische Prozesse (Gen-Ontologie) differentiell exprimierter Gene nach Kryopreservation.	96
Abbildung 3.2-10: Anreicherungsprofile ausgewählter DPs im Vergleich zu quantitativen Methylierungsdaten.	97
Abbildung 3.2-11: DNA-Methylierung repetitiver Elemente.....	99
Abbildung 3.3-1: Schematische Darstellung zur Bearbeitung der murinen Proben.	100
Abbildung 3.3-2: Überprüfung der Immunopräzipitation muriner Proben.....	101
Abbildung 3.3-3: Überblick über MeDIP-Chip-Daten muriner Zelltypen.	103
Abbildung 3.3-4: Differentiell methylierte Promotor-Regionen in gPS vs. ESC und GSC.	105
Abbildung 3.3-5: Überrepräsentierte biologische Prozesse für hypo- und hypermethylierte Gene der gPS.....	107
Abbildung 3.3-6: Überrepräsentierte biologische Prozesse für hypo- und hypermethylierte Gene der pluripotenten Stammzellen.	107
Abbildung 3.3-7: Differentiell methylierte Promotor-Regionen in EpiSC vs. ESC und MEF.....	109
Abbildung 3.3-8: Überrepräsentierte biologische Prozesse für hypo- und hypermethylierte Gene der EpiSC.	111

Abbildung 3.3-9: Überrepräsentierte biologische Prozesse für dynamische Anreicherungs- unterschiede zwischen EpiSC, ESC und MEF.....	112
Abbildung 3.3-10: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung.	113
Abbildung 4.3-1: Modell zu möglichen Ursachen der in gPS beobachteten Hypomethylierung.	134
Abbildung 4.3-2: DNA-Methylierung vs. Gen-Expression in gPS und ESC.	136
Abbildung 4.3-3: Gen-Expression und DNA-Methylierung in EpiSC, ESC und MEF.	139
Abbildung 4.3-4: Modelle zur Interpretation der Methylierungsdaten im Kontext der Genexpression	141
Abbildung 6.1-1: Unspezifische Präzipitation in Zusammenhang mit RNase.....	157
Abbildung 6.1-2: Fragmentierung genomischer DNA mittels Sonifizierung.	158
Abbildung 6.1-3: Hierarchische Clusteranalyse	159
Abbildung 6.1-4: Unterschiedliche Signalverteilung in Abhängigkeit der Hybridisierung.	159
Abbildung 6.2-1: Bisulfit-Sequenzierung vier differentiell exprimierter Gene.	167
Abbildung 6.3-1: Differentielle Methylierung der gPS in Aufwärts- und Kern-Promotoren.....	168
Abbildung 6.3-2: Differentielle Methylierung der EpiSC in Aufwärts- und Kern-Promotoren.	168
Abbildung 6.3-3: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Zfp263, Cacng5.	179
Abbildung 6.3-4: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Ppp2r5c, Grin2b.	180
Abbildung 6.3-5: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Snap25, Peg3/Usp29.	181
Abbildung 6.3-6: Validierung der MeDIP-Chip-Daten mittels Bisulfit-Hochdurchsatz-Sequenzierung: Dmc1h, Pik3r1.	182
Abbildung 6.3-7: Methylierungsanalyse repetitiver Elemente mittels Bisulfit-Hochdurchsatz- Sequenzierung.....	183

9 Tabellenverzeichnis

Tabelle 2.1-1: Oligonukleotide zur Amplifikation Bisulfit-behandelter DNA.....	30
Tabelle 2.1-2: Adaptorsequenzen zur GS FLX-Sequenzierung,	31
Tabelle 2.1-3: Oligonukleotide zur Amplifikation genomischer DNA (qPCR).....	31
Tabelle 2.1-4: Plasmidspezifische Oligonukleotide.....	31
Tabelle 2.2-1: Parameter zur Amplifikation Bisulfit-behandelter DNA unter Verwendung der HOT FIREPol DNA Polymerase (HF).....	39
Tabelle 2.2-2: Parameter zur Amplifikation Bisulfit-behandelter DNA unter Verwendung der HotStarTaq DNA Polymerase (HS).....	39
Tabelle 2.2-3: Parameter zur Amplifikation Bisulfit-behandelter DNA unter Verwendung der HotStart-IT DNA Polymerase (HSIT).....	39
Tabelle 2.2-4: PCR-Bedingungen der verschiedenen Amplikons,.....	40
Tabelle 2.2-5: Parameter zur Amplifikation von Einzelkoloniematerial.....	40
Tabelle 2.2-6: ursprüngliche Parameter zur Amplifikation präzipitierten Materials,	41
Tabelle 2.2-8: Parameter der Sonifizierung.....	45
Tabelle 2.2-9: Versuchsbedingungen der in der vorliegenden Arbeit dargestellten Proben.....	47
Tabelle 2.2-10: Programm zur Bisulfit-Behandlung genomischer DNA	49
Tabelle 2.2-11: Reaktionsansatz zur Ligation von Bisulfit-PCR-Produkten in pGEM-T Vektor	49
Tabelle 2.2-12: Übersicht der pro Amplikon verwendeten Sequenzier-Technik	50
Tabelle 6.1-1: Kontrollregionen zur Überprüfung der MeDIP-Qualität mittels qPCR - <i>Homo sapiens</i>	156
Tabelle 6.1-2: Kontrollregionen zur Überprüfung der MeDIP-Qualität mittels qPCR - <i>Mus musculus</i>	156
Tabelle 6.1-3: Evaluierung Batman-transformierter MeDIP-Chip-Daten.....	158
Tabelle 6.2-1: Ausbeute der Immunopräzipitation methylierter DNA an behandelten und unbehandelten HepG2 Zellen.	160
Tabelle 6.2-2: Gene/Promotoren +/-1 kb um identifizierte DPs in <i>DMSO</i>	161
Tabelle 6.2-3: Gene/Promotoren +/-1 kb um identifizierte DPs in <i>f48h</i>	161
Tabelle 6.2-4: Gene/Promotoren +/-1 kb um identifizierte DPs in <i>120h</i>	162
Tabelle 6.2-5: Gene mit erhöhter Expression - DMSO-Effekt.....	162
Tabelle 6.2-6: Gene mit verminderter Expression - DMSO-Effekt.	162
Tabelle 6.2-7: Gene mit erhöhter Expression - Kryopreservations-Effekt.....	163
Tabelle 6.2-8: Gene mit verminderter Expression - Kryopreservations-Effekt.	163
Tabelle 6.2-9: Gene mit erhöhter Expression - später Effekt.....	163
Tabelle 6.2-10: Gene mit verminderter Expression - später Effekt.	164
Tabelle 6.2-11: GO-Annotationen überrepräsentierter biologischer Prozesse für herunterregulierte Gene mit DMSO-Effekt.	164
Tabelle 6.2-12: GO-Annotationen überrepräsentierter biologischer Prozesse für herunterregulierte Gene mit Kryopreservations-Effekt.	165
Tabelle 6.2-13: GO-Annotationen überrepräsentierter biologischer Prozesse für hochregulierte Gene mit spätem Effekt.	166
Tabelle 6.2-14: GO-Annotationen überrepräsentierter biologischer Prozesse für herunterregulierte Gene mit spätem Effekt.	166
Tabelle 6.3-1 : GO-Annotationen überrepräsentierter biologischer Prozesse für hypomethylierte Gene der gPS.....	169
Tabelle 6.3-2 : GO-Annotationen überrepräsentierter biologischer Prozesse Gene mit keimbahnspezifischer Hypomethylierung der gPS.	169
Tabelle 6.3-3 : GO-Annotationen überrepräsentierter biologischer Prozesse für Gene hypomethylierte Gene der gPS, exklusive keimbahnspezifischer Muster.	170
Tabelle 6.3-4: GO-Annotationen überrepräsentierter biologischer Prozesse für hypermethylierte Gene pluripotenter Zellen.	170
Tabelle 6.3-5: GO-Annotationen überrepräsentierter biologischer Prozesse für hypomethylierte Gene pluripotenter Zellen.	170
Tabelle 6.3-6: GO-Annotationen überrepräsentierter biologischer Prozesse für EpiSC-spezifische Hypermethylierung.	171

Tabelle 6.3-7: GO-Annotationen überrepräsentierter biologischer Prozesse für gewebespezifische Hypermethylierung.	171
Tabelle 6.3-8: GO-Annotationen überrepräsentierter biologischer Prozesse für EpiSC-spezifische Hypomethylierung.	172
Tabelle 6.3-9: GO-Annotationen überrepräsentierter biologischer Prozesse für stammzellspezifische Hypomethylierung.	172
Tabelle 6.3-10: GO-Annotationen überrepräsentierter biologischer Prozesse für dynamische Methylierungsunterschiede.	173
Tabelle 6.3-11: GO-Annotationen überrepräsentierter biologischer Prozesse für dynamische Methylierungsunterschiede.	174
Tabelle 6.3-12: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für hypomethylierte Regionen der gPS, exklusive keimbahnspezifische Muster.	174
Tabelle 6.3-13: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für Regionen mit keimbahnspezifischer Hypomethylierung der gPS.	174
Tabelle 6.3-14: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für pluripotenzspezifische Hypomethylierung.	175
Tabelle 6.3-15: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für pluripotenzspezifische Hypermethylierung.	175
Tabelle 6.3-16: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für EpiSC-spezifische Hypermethylierung.	175
Tabelle 6.3-17: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für gewebespezifische Hypermethylierung.	175
Tabelle 6.3-18: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für EpiSC-spezifische Hypomethylierung.	176
Tabelle 6.3-19: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für stammzellspezifische Hypomethylierung.	176
Tabelle 6.3-20: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für dynamische Methylierungsunterschiede, EpiSC<ESC<MEF.	177
Tabelle 6.3-21: Gene der mittels GO-Annotation identifizierten überrepräsentierten biologischen Prozesse für dynamische Methylierungsunterschiede, ESC<EpiSC<MEF.	178

10 Lebenslauf

Nina Gasparoni
Diplom-Biologin

Persönliche Daten

<i>Name</i>	Nina Gasparoni, geb. Pälme
<i>Geburtsdatum</i>	20.08.1980
<i>Geburtsort</i>	Duisburg-Homberg
<i>Familienstand</i>	verheiratet
<i>Staatsangehörigkeit</i>	deutsch

Ausbildung

Schulbildung:

1987-1991	Grundschule Weiler
1991-1992	Stefan - George - Gymnasium Bingen/ Rhein
1992-2000	Helmholtz - Gymnasium Zweibrücken, Abschluss Abitur (Durchschnittsnote 1,4)

Studium:

10/2000-12/2005	Human- und Molekularbiologie, Universität des Saarlandes, Saarbrücken Abschluss Diplom, Biologie (Dipl.Biol.) Diplom-Prüfungen und Diplomarbeit Hauptfach: Genetik Nebenfächer: Zellbiologie, Entwicklungsbiologie Diplomarbeit: <i>'Analysen zur Chromatin-Reprogrammierung in Zusammenhang mit DNA-Reparatur-Ereignissen in frühen Mausembryonen'</i>
-----------------	---

Beruflicher Werdegang

02/2006-05/2011 und seit 11/2011	Wissenschaftliche Mitarbeiterin, Institut für Genetik/Epigenetik (Universität des Saarlandes), Saarbrücken
Seit 02/2006	Doktorand, Institut für Genetik/Epigenetik (Universität des Saarlandes), Saarbrücken

Zusatzqualifikationen

Computerkenntnisse:

Windows 2000/XP/Vista, MS Office, ChromasPro, BiQ_Analyzer, Axiovision (Zeiss), Corel Photo Paint, R, python

Sprachkenntnisse:

Englisch

Französisch (Grundkenntnisse)

Spanisch (Grundkenntnisse)

Publikationen

Pärmke N, Santacruz D, Walter J (2011) Comprehensive analysis of DNA-methylation in mammalian tissues using MeDIP-chip. *Methods* **53**(2): 175-184.

Gasparoni N, Santacruz D, Ehrhart F, Zimmermann H, Walter J. *DNA methylation patterns largely unaffected by DMSO-based cryopreservation in HepG2 cell.* (in Vorbereitung)

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbst verfasst und keine anderen, als die angegebenen Hilfsmittel benutzt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Saarbrücken, im April 2012

Nina Gasparoni

