

Aus dem Institut für Experimentelle Pädiatrische Endokrinologie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

**Genetische und epigenetische Analyse einer konservierten
nicht-kodierenden Region des *NKX2-1*-Genes in Patienten
mit kongenitaler Hypothyreose**

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Sarah Assadeh Flemmerer

aus Berlin

Datum der Promotion: 04. März 2022

Inhaltsverzeichnis

Abbildungsverzeichnis	5
Tabellenverzeichnis.....	7
Abkürzungsverzeichnis	8
Abstract	10
Abstrakt	11
1 Einleitung	13
1.1 Die molekulare Ursache der Schilddrüsendysgenese	13
1.2 <i>NKX2-1</i> -Mutationen als Ursache der Gehirn-Lungen-Schilddrüsenerkrankung	16
1.3 Ausprägung des Gehirn-Lungen-Schilddrüsen-Syndroms durch Deletionen downstream von <i>NKX2-1</i>	17
1.3.1 Die Rolle des <i>MBIP</i> -Gens für die Entstehung des Gehirn-Lungen- Schilddrüsen-Syndroms.....	18
1.3.2 Störung von Enhancerelementen downstream von <i>NKX2-1</i> als Ursache des Gehirn-Lungen-Schilddrüsen-Syndroms.....	19
1.3.3 Mutationen in einem schilddrüsenspezifischen Enhancer als mögliche Erklärung der Schilddrüsendysgenese	24
1.4 Epigenetik	26
1.4.1 DNA-Methylierung	27
1.4.2 DNA-Methylierungsdefekte als Ursache der Schilddrüsendysgenese	31
1.5 Zielsetzung	32
1.5.1 Genetische Analyse der potentiellen Enhancer Region	34
1.5.2 Epigenetische Analyse der potentiellen Enhancer Region.....	35
2 Material und Methoden.....	37
2.1 Verwendete Geräte.....	37
2.2 Verwendetes Material	38
2.2.1 Verwendete Primer	40
2.3 Kohortenverzeichnis	42

2.4	Erklärung und Durchführung der angewandten Methoden.....	44
2.4.1	Genotypisierung	44
2.4.1.1	DNA Isolation aus Leukozyten	44
2.4.1.2	Polymerase Kettenreaktion	44
2.4.1.3	Agarosegelelektrophorese	46
2.4.1.4	ExoSAP-Aufreinigung.....	46
2.4.1.5	Sanger-Sequenzierung	47
2.4.1.6	Auswertung mit der Seq-Pilot-Software	49
2.4.2	Methylierungsanalyse.....	49
2.4.2.1	Bisulfitumwandlung	49
2.4.2.2	PCR der Bisulfit konvertierten DNA mit 5`biotinyliertem Rückwärtsprimer	53
2.4.2.3	Agaroseelektrophorese	54
2.4.2.4	Aufreinigung mit magnetischen Beads.....	54
2.4.2.5	Pyrosequenzierung	55
2.4.2.6	Auswertung mit PyroMark Q24 Software	58
3	Ergebnisse	59
3.1	Genotypisierung.....	59
3.1.1	Genetische Analyse der potentiellen Enhancer Region in drei Kohorten.....	60
3.1.2	Genetische Analyse von Schilddrüsen-DNA gesunder Personen.....	61
3.1.3	SNPs in der potentiellen Enhancer Region	62
3.1.4	SNPs als Ursache einer verminderten Transkriptionsfaktorbindung	62
3.2	Methylierungsanalyse	67
3.2.1	Pyrosequenzierung der potentiellen Enhancer Region	67
3.2.2	Gewebespezifität des Methylierungsmusters in der PER.....	69
3.2.3	Kohortenvergleich der Methylierung in der PER	70
3.2.3.1	Mittlere Methylierung der PER	71
3.2.3.2	Positionsspezifische Methylierungsunterschiede	72
3.2.3.3	CpG-positionsspezifische Methylierung	75
3.2.4	Kohortenvergleich der Methylierung bei monozygoten Zwillingen	77
3.2.4.1	Mittlere Methylierung	77
3.2.4.2	Positionsspezifische Methylierungsunterschiede	77
3.2.4.3	CpG-positionsspezifische Methylierung	78

3.2.4.4	Zwillingspaarspezifische Methylierung	79
3.2.5	Geschlechtsspezifische Methylierung	80
3.2.5.1	Kohortenunabhängige geschlechtsspezifische Methylierung.....	80
3.2.5.2	Kohortenabhängige geschlechtsspezifische Methylierung.....	81
3.2.5.3	Geschlechtsspezifische Methylierungsunterschiede zwischen den Kohorten	81
3.2.6	Belastbarkeit der Ergebnisse.....	82
3.2.6.1	Robustheit der Pyrogrammdatei	82
3.2.6.2	Runspezifischer Methylierungsunterschied.....	84
3.2.6.3	Abhängigkeit der Methylierung vom Alter des Patienten bei der Blutentnahme	87
3.2.6.4	Abhängigkeit der Methylierung vom Alter des gleichen Patienten	88
3.2.6.5	Abhängigkeit der Methylierung vom Alter der Blutprobe	88
4	Diskussion.....	92
4.1	Genotypisierung.....	93
4.1.1	Kohortenspezifische SNP-Verteilung in der PER.....	93
4.1.2	Beeinflussung der Transkriptionsfaktorbindung durch SNPs	96
4.2	Epigenetische Methylierungsanalyse der potentiellen Enhancer Region	99
4.2.1	Gewebespezifität der Methylierung der potentiellen Enhancer Region..	100
4.2.2	Abhängigkeit der Methylierung von der Zelllinie.....	103
4.2.3	Methylierungsintensität der PER im Kohortenvergleich	103
4.2.4	Altersabhängige Methylierungsveränderungen	105
4.2.5	Zusammenhang der Methylierung mit der Schilddrüsenfunktion	106
4.2.6	CpG-positionsspezifische Methylierungsunterschiede	106
4.2.7	DNA-sequenzspezifische Methylierungsunterschiede	108
4.2.8	Geschlechtsspezifische Methylierungsunterschiede.....	109
4.3	Zusammenfassung und Ausblick.....	110
5	Literaturverzeichnis.....	113
6	Eidesstattliche Versicherung	121
7	Curriculum vitae	123
8	Danksagung	124

Abbildungsverzeichnis

Abbildung 1-1: Schilddrüsenentwicklung	15
Abbildung 1-2: Deletionen mit NKX2-1 defizienten Phänotyp	17
Abbildung 1-3: Deletionen des MBIP-Gens	18
Abbildung 1-4: Funktionellen Domäne (TAD) des NKX2-1- und Mbip-Gens, sowie der nicht-kodierenden Region (NKR)	21
Abbildung 1-5: DNA-Interaktion zwischen NKX2-1, Mbip und der NKR downstream von NKX2-1	22
Abbildung 1-6: Potentielle Enhancer Region	26
Abbildung 1-7: Etablierung des Methylierungsmusters	29
Abbildung 1-8: Hypothese und Fragestellung	33
Abbildung 2-1: Ablauf der Bisulfidkonvertierung	51
Abbildung 2-2: Pyrosequenzierung	56
Abbildung 2-3: Pyrosequenzierter Bereich der potentiellen Enhancer Region	57
Abbildung 3-1: Auszug aus dem Genombrowser der PER	65
Abbildung 3-2: SNPs und SNVs in der CH-Kohorte	66
Abbildung 3-3: Ausschnitt aus der DNA-Sequenz der PER nach Bisulfidkonversion.	68
Abbildung 3-4: Pyrogramm der potentiellen Enhancer Region	69
Abbildung 3-5: Gewebespezifisches Methylierungsmuster der PER	70
Abbildung 3-6: Kohortenvergleich der mittleren Methylierung der PER	72
Abbildung 3-7: Positionsspezifischen Methylierungsunterschiede	73
Abbildung 3-8: Kohortenübergreifende positionsspezifische Methylierung	76
Abbildung 3-9: CpG-Positionen nach aufsteigender Methylierung	76
Abbildung 3-10: Monozygote Zwillinge (MZ), CpG-Positionen aufsteigend nach Methylierung sortiert	79
Abbildung 3-11: CpG-positionsspezifische Methylierung der monozygoten Zwillingspaare	80
Abbildung 3-12: Geschlechtsspezifische Methylierungsunterschiede	82
Abbildung 3-13: Signifikanzvergleich des Mittelwertes der Methylierung nach Datenauswahl	83

Abbildung 3-14: Signifikanzvergleich des Mittelwertes der Methylierung nach Datenauswahl pro CpG-Position.....	84
Abbildung 3-15: Runspezifischer Methylierungsunterschied	85
Abbildung 3-16: Kohortenübergreifender Vergleich der mittleren Methylierung verschiedener Pyrosequenzierungs-Runs	86
Abbildung 3-17: Kohortenspezifischer Vergleich der mittleren Methylierung verschiedener Pyrosequenzierungs-Runs	87
Abbildung 3-18: Korrelation Patientenalter bei Blutentnahme und Methylierung	88
Abbildung 3-19: Abhängigkeiten der Methylierung vom Alter des gleichen Patienten...	88
Abbildung 3-20: Kohortenspezifisches Alter der Blutproben.....	89
Abbildung 3-21: Abhängigkeit der Methylierung vom Alter der Blutprobe innerhalb der Kohorten.....	90
Abbildung 3-22: Kohortenspezifische Abhängigkeit der Methylierung vom Alter der Blutprobe	90
Abbildung 4-1: Zusammenhang Methylierungsverteilung und Schilddrüsenfunktion...	106

Tabellenverzeichnis

Tabelle 2-1: Verwendete Geräte	38
Tabelle 2-2: Verwendetes Material	40
Tabelle 2-3: Verwendete Primer	42
Tabelle 2-4: Kohortenverzeichnis.....	43
Tabelle 2-5: Monozygote diskordante Zwillingspaare	44
Tabelle 2-6: PCR-Ansatz	45
Tabelle 2-7: PCR-Ablauf	46
Tabelle 2-8: ExoSAP-Ablauf	47
Tabelle 2-9: Ansatz Sequenzier-PCR	48
Tabelle 2-10: Ablauf Sequenzier-PCR	48
Tabelle 2-11: Bisulfit PCR-Ansatz	53
Tabelle 2-12: Bisulfit PCR-Ablauf.....	53
Tabelle 3-1: Phänotyp SNP/ SNV Korrelation.....	61
Tabelle 3-2: Absolute Methylierungsdifferenz und Signifikanz.....	75
Tabelle 3-3: Methylierungsdifferenzen monozygoter Zwillingspaare	78
Tabelle 3-4: Durchschnittliche Methylierung pro Zwillingspaar pro CpG-Position	79
Tabelle 3-5: Geschlechtsspezifische Methylierungsunterschiede zwischen den Kohorten	80
Tabelle 3-6: Geschlechtsspezifische Methylierungsdifferenzen zwischen den Kohorten	82
Tabelle 3-7: Multiple Regression der Methylierung in Abhängigkeit vom Blutalter und der Kohorte.....	91

Abkürzungsverzeichnis

A.....	DNA Base Adenosin
ADP.....	Adenosindiphosphat
AMP.....	Adenosinmonophosphat
APS.....	Adenosin 5' Phosphosulfat
ATP.....	Adenosintriphosphat
C.....	DNA Base Cytosin
CTCF.....	Transkriptionsrepressor
CH.....	kongenitale Hypothyreose
CHIP.....	Chromatin-Immopräzipitation
3C.....	Chromosome conformation capture
4C.....	Circularized chromosome conformation capture
CNV.....	Copy number variants
CpG.....	CpG-Dinukleotid
ddATP.....	Didesoxyadenosintriphosphat
ddCTP.....	Didesoxycytosintriphosphat
ddGTP.....	Didesoxyguanosintriphosphat
ddNTP.....	Didesoxyribonukleosidtriphosphat
ddTTP.....	Didesoxythymidintriphosphat
DMR.....	Differentially methylated regions
dNDP.....	Desoxyribonukleosiddiphosphat
dNMP.....	Desoxyribonukleosidmonophosphat
dNTP.....	Desoxyribonukleosidtriphosphat
EDTA.....	Ethylendiamintetraacetat
EMSA.....	Electrophoresis mobility shift assay
FT4.....	freies Thyroxin
G.....	DNA Base Guanin
H3K9.....	Lysin 9 von Histon 3
H3K4me1.....	Monomethylierung des Lysin 4 von Histon 3
H3K4me3.....	Trimethylierung des Lysin 4 von Histon 3
H3K9me3.....	Trimethylierung des Lysin 9 von Histon 3
H3K27ac.....	Acetylierung des Lysin 27 von Histon 3
HTT.....	Hyperthyreotropinämie

IEPE Institut für experimentelle pädiatrische Endokrinologie
iPSCs Induzierte pluripotente Stammzellen
ITD Isolated thyroid dysgenesis
K Gesunde Kontrollen
MAF Minor Allel Frequenz
Meth. Diff. Methylierungsdifferenz
MZ Monozygote Zwillinge
MZPP Monozygote Zwillingspaare
NKR Nicht-kodierende Region
PBL Peripheren Blutzellen
PCR Polymerase Kettenreaktion
PER Potentielle Enhancer Region
PPi Pyrophosphat
rSNP Regulatory single nucleotide polymorphisms
SAP Shrimp Alkaline Phosphatase
SD Schilddrüse
SD-Gewebe Schilddrüsengewebe
SDDG Schilddrüsendysgenese
SNP Single nucleotide polymorphisms
SNV Single nucleotide variants
T DNA Base Thymin
TAD Topological associated domains
TD Thyroid dysgenesis
T-DMR Tissue-specific differentially methylated regions
TFBS Transkriptionsfaktorbindungsstellen
TSH Thyreoidea-stimulierendes Hormon
TSS Transkriptionsstartpunkt
U RNA Base Uracil

Abstract

Congenital hypothyroidism is the most frequent endocrine disorder in neonates. The most common cause is thyroid dysgenesis (TD), a structural disorder of thyroid development. Only in 5% of cases of TD, the genetic cause is known. The *NKX2-1* gene is involved in thyroid development, and its mutation can cause TD. It could be shown that deletions in a non-coding region downstream of *NKX2-1* cause an *NKX2-1* deficient phenotype. On this basis, a thyroid-specific enhancer in this non-coding region was suspected, whose genetic or epigenetic disorder could cause isolated thyroid dysgenesis (ITD). To analyze this non-coding region downstream of *NKX2-1*, one of the largest cohorts of patients with ITD and six pairs of monozygotic twins that are discordant for ITD were used. To classify and interpret the results one cohort of patients with hyperthyrotropinaemia, one cohort with healthy controls and thyroid tissue of healthy controls were analyzed.

In the genetic analysis, a new single nucleotide variant (SNV) an heterozygous rare SNPs were found only in patients with ITD. The epigenetic methylation analysis revealed tissue-specific hypomethylation of the non-coding region downstream of *NKX2-1* in thyroid tissue, which is an indication of the postulated thyroid-specific regulatory function. The methylation comparison of the cohorts and the pairs of monozygotic twins revealed a small statistically significant hypermethylation in patients with ITD.

This hypermethylation was hardly reproducible with monozygotic twins.

A gender-specific analysis revealed statistically significant hypermethylation in women with ITD. This finding could be related to the girlish nature of the disease.

Both the result of tissue-specific hypomethylation of thyroid tissue and the hypermethylation of patients with ITD are compatible with a tissue-specific regulatory function. A reproduction of of the results with larger and more homogenous cohorts is required for further evaluation and interpretation.

In summary, both the genetic analysis and the epigenetic methylation analysis provided indications for a potential relevance of the analyzed non-coding region downstream of *NKX2-1* for ITD.

Abstrakt

Die kongenitale Hypothyreose ist die häufigste angeborene endokrinologische Erkrankung. Die vorwiegende Ursache der CH ist eine strukturelle Störung der Schilddrüsenentwicklung, die Schilddrüsendysgenese. In nur 5 % der Fälle einer Schilddrüsendysgenese ist die genetische Ursache bekannt. Das *NKX2-1*-Gen ist unter anderem an der Schilddrüsenentwicklung beteiligt, und mutationsbedingte Störungen können Ursache der Schilddrüsendysgenese sein. Es konnte gezeigt werden, dass Deletionen in einem nicht-kodierenden Bereich downstream von *NKX2-1* zu einem *NKX2-1* defizienten Phänotyp führen. Auf dieser Grundlage wurde eine schilddrüsen-spezifische Enhancerfunktion in diesem nicht-kodierenden Bereich vermutet, deren genetische oder epigenetische Störung Ursache der isolierten Schilddrüsendysgenese sein könnte.

Für die Analyse der potenziellen Enhancer Region wurde eine der größten Kohorten an Patienten mit Schilddrüsendysgenese sowie sechs für die Schilddrüsendysgenese diskordante monozygote Zwillingspaare verwendet. Zur vergleichenden Einordnung wurden gesunde Kontrollen, Patienten mit Hyperthyreotropinämie sowie Schilddrüsen-gewebe gesunder Personen in die Analyse mit einbezogen.

In der genetischen Analyse konnte nur in Patienten mit isolierter Schilddrüsendysgenese ein neuer Single Nucleotide Variant (SNV), sowie heterozygot seltene SNPs gefunden werden. Die epigenetische Methylierungsanalyse der potentiellen Enhancer Region (PER) zeigte eine gewebespezifische Hypomethylierung im Schilddrüsen-gewebe, die ein Hinweis für die postulierte schilddrüsen-spezifische regulative Funktion ist. Der Methylierungsvergleich der Kohorten im Blut ergab eine geringe signifikante Hypermethylierung der Patienten mit isolierter Schilddrüsendysgenese, die sich in den diskordanten monozygoten Zwillingen nicht oder kaum reproduzieren ließ. Geschlechtsspezifisch zeigte sich eine signifikante Hypermethylierung der Frauen mit isolierter Schilddrüsendysgenese, was mit der Mädchenwendigkeit der CH in Zusammenhang stehen könnte.

Sowohl das Ergebnis der gewebespezifischen Hypomethylierung des Schilddrüsen-gewebes als auch die Hypermethylierung der Patienten mit isolierter Schilddrüsendysgenese sind mit gewebespezifischer regulativer Funktion im Sinne eines schilddrüsen-spezifischen Enhancers vereinbar. Zur weiteren Einordnung und Interpretation der Daten ist eine Reproduktion der Ergebnisse in größeren und

homogeneren Kohorten dringend erforderlich. Zusammenfassend zeigten sich sowohl in der genetischen Analyse als auch in der epigenetischen Methylierungsanalyse Hinweise für eine mögliche Beteiligung der potentiellen Enhancer Region an der Pathogenese der Schilddrüsendysgenese.

1 Einleitung

Die kongenitale Hypothyreose ist mit einer Inzidenz von 1: 3000 bis 1:4000 die häufigste angeborene endokrinologische Erkrankung [1, 2]. Aufgrund der zentralen Rolle der Schilddrüsenhormone für die Entwicklung des zentralen Nervensystems stellt die kongenitale Hypothyreose, wenn sie nicht sofort behandelt wird, weltweit die häufigste Ursache für geistige Retardierung im Kindesalter dar [3]. Aufgrund der anfangs dezenten und unspezifischen Symptome [2] und der verheerenden Folgen einer geistigen Retardierung bei einem zu späten Behandlungsbeginn ist die Erfassung der kongenitalen Hypothyreose (CH) Teil des Neugeborenen Screenings [4]. Die Therapie besteht in einer schnellst möglichen Substitution des fehlenden Schilddrüsenhormons [4]. Langzeitstudien konnten bei mehr als 90 Prozent der Patienten, die eine rechtzeitige und ausreichend hohe Substitutionstherapie erhielten, eine normale geistige Entwicklung beobachten [5, 6].

Die Ätiologie der kongenitalen Hypothyreose (CH) ist vielfältig. Die häufigste Form stellt die primäre CH dar [2]. Nach dem ursächlichen Mechanismus wird weiter eine Störung der Schilddrüsenhormonbiosynthese von einer strukturellen Störung der Schilddrüsenentwicklung, die als Schilddrüsendysgenese bezeichnet wird, unterschieden. Eine Störung der Schilddrüsenhormonbiosynthese ist für etwa 20% der Fälle einer primären CH verantwortlich, wohingegen mit 80% der Großteil der Fälle einer primären CH auf eine Schilddrüsendysgenese, einer strukturellen Störung der Schilddrüsenentwicklung, zurückzuführen ist [4, 7]. Die Schilddrüsendysgenese wird je nach Schwere der morphologischen und funktionellen Einschränkung weiter in die drei Unterformen Athyreose, Ektopie und Hypoplasie unterteilt [1]. Dabei ist die Athyreose die schwerste Form der Schilddrüsendysgenese ohne Nachweis von sichtbarem Schilddrüsenengewebe. Bei der Schilddrüsenektopie gelingt der Nachweis von Schilddrüsenengewebe meist sublingual außerhalb der normalen Position, wohingegen bei der Schilddrüsenhypoplasie Gewebereste in orthotoper Lage nachgewiesen werden können und die Beeinträchtigung der Schilddrüsenfunktion meist milder ausfällt [1, 2].

1.1 Die molekulare Ursache der Schilddrüsendysgenese

Screeningprogramme konnten ethnische Unterschiede in der Inzidenz der CH zeigen. So ist diese in den USA z.B. bei hispanischen Neugeborenen mit 1:2000 höher und bei

afroamerikanischen Neugeborenen mit 1:10000 deutlich geringer als die durchschnittliche über ethnische und geographische Grenzen hinweg nachgewiesene Inzidenz von 1:3000 bis 1:4000 [1, 8, 9]. Diese ethnischen Unterschiede sind hinweisend für eine genetische Beteiligung bei der Entstehung der CH [10].

Die genetische Basis der Schilddrüsendysgenese als häufigste Form der primären CH bleibt dennoch unklar. Neben den ethnischen Unterschieden der Inzidenz konnte als weiterer Hinweis für eine genetische Ursache der Schilddrüsendysgenese eine erhöhte Prävalenz an weiteren kongenitalen Defekten bei Patienten mit CH im Vergleich zur gesunden Bevölkerung nachgewiesen werden [11, 12]. Allerdings spricht das überwiegend sporadische Auftreten [4], das bevorzugte Auftreten bei Mädchen [13] und die große Diskordanz zwischen monozygoten Zwillingen [14, 15] gegen eine klassische mendelsche Vererbung als Ursache der Schilddrüsendysgenese. Eine Beteiligung von nicht genetischen Faktoren oder stochastischen Ereignissen während der frühen embryonalen Entwicklung bei der Entstehung der Schilddrüsendysgenese wird daher angenommen [11].

Die Schilddrüse entwickelt sich aus zwei unterschiedlichen embryonalen Geweben. Die das Schilddrüsenhormon produzierenden Follikelzellen entstehen aus endodermalem Gewebe des Mundbodens und bilden die mediane Schilddrüsenanlage. Die Calcitonin produzierenden C-Zellen dagegen entstehen aus ektodermalem Gewebe des ultimobranchialen Körpers und bilden die laterale Schilddrüsenanlage [16]. Für die Schilddrüsenhormonproduktion und somit auch für die CH sind einzig die Follikelzellen relevant. Deren terminale Differenzierung zu Schilddrüsenhormon produzierenden Zellen ist unter anderem abhängig von der Expression des TSH-Rezeptors, des Natrium-Iodid-Symporters, von Thyreoglobulin und der Thyreoperoxidase sowie von der regelrechten Organogenese mit Fusion der beiden Schilddrüsenanlagen und Lokalisation des Schilddrüsengewebes in der orthotopen Position [16, 17].

Mit Hilfe von Knockout Mausmodellen konnte die essentielle Rolle der Transkriptionsfaktoren *FOXE1*, *NKX2-1* und *PAX8* für eine regelrechte Schilddrüsenentwicklung sowie die Rolle des später in der Entwicklung exprimierten TSH-Rezeptors für die Differenzierung und Aufrechterhaltung der Follikelzellen identifiziert werden (siehe Abbildung 1-1) [16].

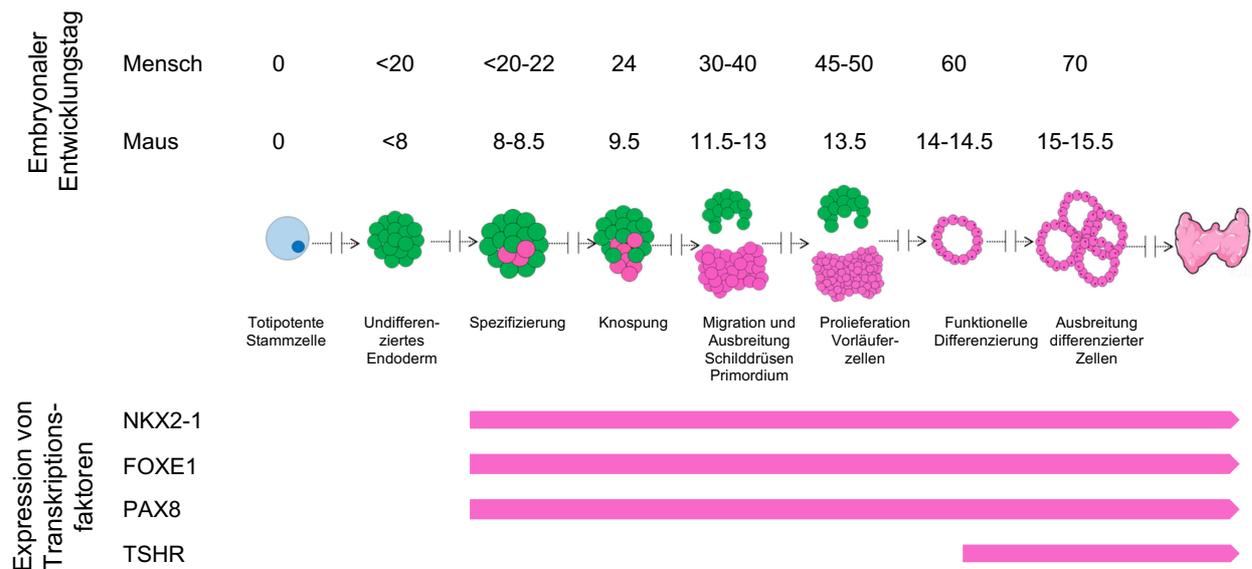


Abbildung 1-1: Schilddrüsenentwicklung

Zeitlicher Verlauf der Schilddrüsenentwicklung mit Expression relevanter Transkriptionsfaktoren. Weiter sind die verschiedenen Entwicklungsstadien im zeitlichen Verlauf für die Maus und den Menschen dargestellt. Ab dem 60. embryonalen Tag beginnt im Menschen während der Follikeldifferenzierung die Expression des TSH-Rezeptors. In Anlehnung an Mio et al 2020 [18].

Auf der Grundlage dieser tierexperimentellen Ergebnisse konnten entsprechende Mutationen in diesen vier Genen auch im Menschen beschrieben und mit einer Schilddrüsendysgenese in Verbindung gebracht werden [11, 16].

So konnten in ca. 5% der Fälle einer Schilddrüsendysgenese Mutationen in Transkriptionsfaktoren identifiziert werden, die an der Schilddrüsenentwicklung und Aufrechterhaltung beteiligt sind [11, 19]. Aufgrund der kritischen Rolle dieser Transkriptionsfaktoren auch für die Entwicklung anderer Organe tritt die Schilddrüsendysgenese in diesen Fällen meist im Rahmen einer syndromalen Erkrankung auf. Mutationen in den Transkriptionsfaktoren *PAX8*, *FOXE1* und *NKX2-1* konnten als genetische Ursache der syndromalen Schilddrüsendysgenese identifiziert werden [20]. Allerdings werden vereinzelt *PAX8*-Mutationen wie auch *TSH-Rezeptor*-Mutationen als ursächliche Mutationen einer isolierten, nicht syndromalen Schilddrüsendysgenese beschrieben [2, 19, 21]. Mittlerweile gibt es neben den oben genannten Mutationen eine Vielzahl von Genen, für die eine direkte oder indirekte Assoziation mit der Schilddrüsendysgenese beschrieben werden konnte [10].

Nichtsdestotrotz sind monogene Ursachen insgesamt nur in 5% der Fälle einer Schilddrüsendysgenese zu finden. Für den Großteil der Fälle, insbesondere einer

isolierten Schilddrüsendysgenese, bleibt die zu Grunde liegende molekulare Ursache weiter unbekannt [2, 11].

1.2 *NKX2-1*-Mutationen als Ursache der Gehirn-Lungen-Schilddrüsenerkrankung

Einer der oben genannten Transkriptionsfaktoren ist das *NKX2-1*-Gen, dessen Mutation für einen Teil der syndromalen Schilddrüsendysgenese verantwortlich ist.

Das *NKX2-1*-Gen liegt auf dem Chromosom 14q13 [22, 23] und gehört zu der NK2-Genfamilie, die durch das Vorhandensein einer NK2-Homeobox charakterisiert ist [24, 25]. Das *NKX2-1*-Gen kodiert den gleichnamigen Transkriptionsfaktor. Dieser bindet an Promotorregionen der DNA und aktiviert so spezifische Gene der Schilddrüse, Lunge und des zentralen Nervensystems [24, 26]. Mutationen im *NKX2-1*-Gen konnten als genetische Ursache des Gehirn-Lungen-Schilddrüsen-Syndroms nachgewiesen werden [27], welches durch ein variables phänotypisches Spektrum entsprechender Organfunktionsstörungen charakterisiert ist [23]. In seiner vollen Ausprägung tritt es in Form einer kongenitalen Hypothyreose, einem akuten Atemnotsyndrom sowie einer benignen hereditären Choreoathetose in Erscheinung [28, 29]. Bei etwa der Hälfte der Patienten tritt das Syndrom in dieser phänotypischen kompletten Trias auf [30]. Die andere Hälfte der Patienten zeigt eine variable phänotypische Ausprägung, wobei eine Bewegungsstörung meist in Form einer benignen hereditären Choreoathetose als führendes Symptom in allen Patienten zu finden ist. Eine Affektion der Lunge oder der Schilddrüse ist dagegen phänotypisch variabel ausgeprägt [31].

In der Schilddrüse ist *NKX2-1* an der Entwicklung und Aufrechterhaltung der Schilddrüsenfunktion beteiligt [28] und reguliert die Expression der für die Schilddrüsenhormonbiosynthese essentiellen Gene, zu denen unter anderen Thyroglobulin, der TSH-Rezeptor, die Schilddrüsenperoxidase sowie Pendrin gehören [28, 31]. Eine Vielzahl von Mutationen im *NKX2-1*-Gen konnte somit in Patienten mit syndromaler Schilddrüsendysgenese beschrieben werden [28].

1.3 Ausprägung des Gehirn-Lungen-Schilddrüsen-Syndroms durch Deletionen downstream von *NKX2-1*

In einer Studie des Instituts für experimentelle pädiatrische Endokrinologie (IEPE) von Thorwarth et al. 2014 mit umfassender Genotyp-Phänotyp-Korrelation einer großen Kohorte von Patienten mit Gehirn-Lungen-Schilddrüsen-Syndrom konnten neben 15 neuen intragenetischen Punktmutationen im *NKX2-1*-Gen 10 Patienten mit heterozygoten Deletionen des Chromosoms 14q13 identifiziert werden. Acht dieser heterozygoten Deletionen umfassten das *NKX2-1*-Gen, was eine plausible Erklärung für die phänotypische *NKX2-1*-Defizienz mit Ausprägung des Gehirn-Lungen-Schilddrüsen-Syndroms dieser Patienten darstellt.

Die übrigen zwei heterozygoten Deletionen umfassten allerdings nicht das *NKX2-1*-Gen, sondern lediglich das downstream gelegene *MBIP*-Gen. Dennoch zeigte sich in beiden Patienten die phänotypische Ausprägung einer *NKX2-1*-Defizienz in Form des Gehirn-Lungen-Schilddrüsen-Syndroms mit einer Choreoathetose und in einem der Patienten zusätzlich einer milden Schilddrüsenfunktionsstörung [31].

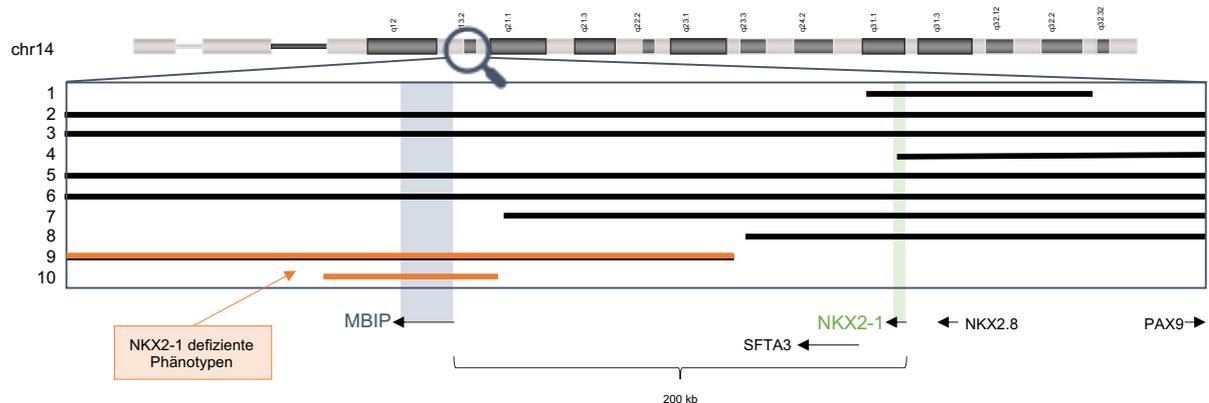


Abbildung 1-2: Deletionen mit *NKX2-1* defizienten Phänotyp

Die Abbildung zeigt die in der Studie von Thorwarth et al. 2014 [31] gefundenen Deletionen in räumlichem Zusammenhang mit dem *NKX2-1*- und *MBIP*-Gen. In Orange markiert sind die Deletionen, die nicht das *NKX2-1*-Gen betreffen, deren Träger aber einen *NKX2-1* defizienten Phänotyp zeigen. Die Abbildung ist in Anlehnung an Thorwarth et al. 2014 [31].

Weitere dieser Deletionen, downstream des *NKX2-1*-Gens mit Deletion des *MBIP*-Gens in Patienten mit einem *NKX2-1* defizienten Phänotyp, wurden beschrieben [29, 30, 32, 33]. Der zugrunde liegende Mechanismus der *NKX2-1*-Defizienz bei diesen Patienten ist bisher allerdings nicht bekannt.

1.3.1 Die Rolle des *MBIP*-Gens für die Entstehung des Gehirn-Lungen-Schilddrüsen-Syndroms

Das *MBIP*-Gen ist downstream von *NKX2-1* gelegen und kodiert für das inhibitorische Bindeprotein MAP3K12 [29, 31], das Teil eines Histonacetylasekomplexes ist, der an der Chromatinmodifikation beteiligt ist [30].

Eine mögliche Erklärung für den *NKX2-1* defizienten Phänotyp durch Deletionen, die ausschließlich das *MBIP*-Gen und nicht das *NKX2-1*-Gen betreffen, ist einerseits, dass eine Störung des *MBIP*-Gens direkt in die Pathogenese des Gehirn-Lungen-Schilddrüsen-Syndroms involviert ist, oder andererseits, dass *MBIP* über eine Beeinflussung der *NKX2-1*-Expression zu einer *NKX2-1*-Defizienz führt [29].

Invernizzi et al. 2018 konnten eine verminderte Expression von *NKX2-1* bei Patienten mit einer entsprechenden Deletion downstream des *NKX2-1*-Gens, die das *MBIP*-Gen umfasst, zeigen. Dies könnte für eine sekundäre Störung der *NKX2-1*-Expression durch die gefundenen Deletionen sprechen, die somit eine *NKX2-1*-Haploinsuffizienz vortäuschen [29].

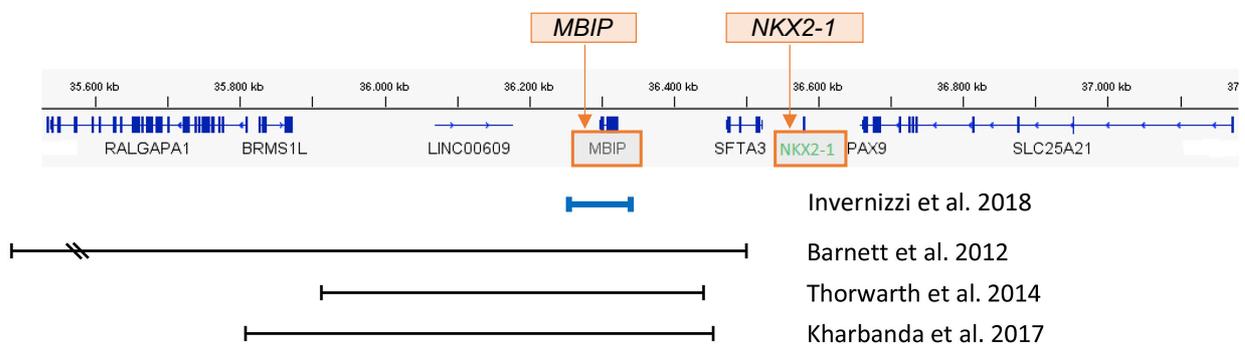


Abbildung 1-3: Deletionen des *MBIP*-Gens

Dargestellt sind Deletionen in Patienten mit *NKX2-1* defizientem Phänotyp, die das *MBIP*-Gen, aber nicht das *NKX2-1*-Gen umfassen. Die Abbildung ist in Anlehnung an Invernizzi et al. 2018 [29] verfasst, deren gefundene Deletion in Blau dargestellt ist.

Weiter konnte eine Studie des IEPEs von Thorwarth et al. 2014 im Mausmodell eine Coexpression von *NKX2-1* und *Mbp* im ZNS, in der Lunge und der Schilddrüse nachweisen. Eine daraufhin analysierte Kohorte von 74 Patienten mit Gehirn-Lungen-Schilddrüsen-Syndrom ohne *NKX2-1*-Mutationen zeigte allerdings in keinem der Patienten eine *MBIP*-Mutation [31]. Dies wiederum spricht gegen eine direkte Rolle von *MBIP*-Mutationen bei der Entstehung des Gehirn-Lungen-Schilddrüsen-Syndroms [29].

Bisherige Erkenntnisse über das MBIP-Protein zeigen auch keinen Hinweis auf eine Beteiligung an der Schilddrüsenentwicklung oder Expressionregulation von schilddrüsenspezifischen Genen [29, 30].

1.3.2 Störung von Enhancerelementen downstream von *NKX2-1* als Ursache des Gehirn-Lungen-Schilddrüsen-Syndroms

Viel diskutiert wird neben der Rolle des *MBIP*-Gens bei der Entstehung des Gehirn-Lungen-Schilddrüsen-Syndroms durch Deletionen, downstream des *NKX2-1*-Gens, eine mögliche Zerstörung von regulativen Elementen wie z.B. Enhancern, die zu einer Expressionsbeeinflussung des *NKX2-1*-Gens führen [30, 31].

Der Störung regulativer Elemente und Funktionen durch Deletionen, Insertionen und strukturelle Varianten wie Inversionen und Translokationen im nicht-kodierenden Genom wird für die Klärung der genetischen Ursache von Erkrankungen eine wichtige Rolle beigemessen und ist Gegenstand aktueller Forschung [34]. Es konnte gezeigt werden, dass die räumliche und funktionelle Organisation des Genoms maßgeblich an der Genregulation beteiligt ist [35]. Mit sogenannten *chromosome conformation capture* (3 C) Experimenten können Chromatininteraktionen des Genoms analysiert werden. Die Weiterentwicklung dieser Experimente ermöglicht inzwischen eine genomweite Analyse der Chromatininteraktionen [36]. Mit Hilfe dieser Experimente konnten Dixon et al. 2012 sogenannte „topological associated domains“ (TADs) nachweisen [35]. Dabei handelt es sich um megabasen große DNA-Bereiche, die das Genom in funktionelle Einheiten unterteilen [37]. Die TADs als funktionelle Einheiten an sich sind während der Differenzierung und Entwicklung stabil vorhanden, während sich die Organisation innerhalb einer TAD zelltypspezifisch unterscheidet. DNA-Interaktionen innerhalb einer TAD werden durch eine Schleifenbildung der DNA vermittelt, die als „DNA-Looping“ bezeichnet wird. Das DNA-Looping wird durch die lokale Chromatinstruktur erleichtert und bringt distal gelegene transkriptionsfaktorgebundene regulative Elemente näher zu ihren Zielgenen. Typischerweise treten diese Interaktionen zwischen Enhancern und Promotoren eines Zielgens auf. Diese funktionellen Enhancer-Promotor-Loops innerhalb der TAD können einerseits zwischen verschiedenen Säugetierarten hoch konserviert vorliegen, können aber andererseits auch von gewebespezifischen Transkriptionsfaktoren abhängig sein und somit gewebespezifisch auftreten [38]. Die

Begrenzung der TADs als funktionelle regulative Einheit durch Proteine wie CTCF und Cohesin scheint dabei eine entscheidende Rolle für die Aufrechterhaltung regelrechter Chromatininteraktionen zu spielen [34]. Lupianez et al. 2016 konnten im Mausmodell zeigen, dass die verschiedenen Formen von Extremitätenfehlbildungen durch strukturelle Varianten im EPHA4-Locus durch die Lokalisation der strukturellen Varianten innerhalb der TAD bedingt sind. So kommt es durch eine Störung der TAD-Grenzen zu TAD übergreifenden ektope Enhancer-Gen-Interaktionen, die je nach entsprechender Genaktivierung zu den verschiedenen Extremitätenfehlbildungen führen. Die entsprechenden krankheitsauslösenden ektope Interaktionen konnten in humanen Zellen aus betroffenen Patienten reproduziert werden [39]. Eine isolierte Organisationsstörung des Genoms konnte somit als krankheitsverursachend nachgewiesen werden. Für eine derartige genomische Regulationsstörung wurde der Begriff „Enhancer adoption“ geprägt. Er beschreibt die ektope Expression eines Gens durch einen Enhancer, der nicht dem entsprechenden Gen zugehörig ist [40].

Unveröffentlichte Ergebnisse des Instituts für experimentelle pädiatrische Endokrinologie (IEPE) von Thorwath et al. konnten zeigen, dass sich *NKX2-1*, *Mbip* und die nicht-kodierende Region (NKR) in einer gemeinsamen TAD befinden (siehe Abbildung 1-4).

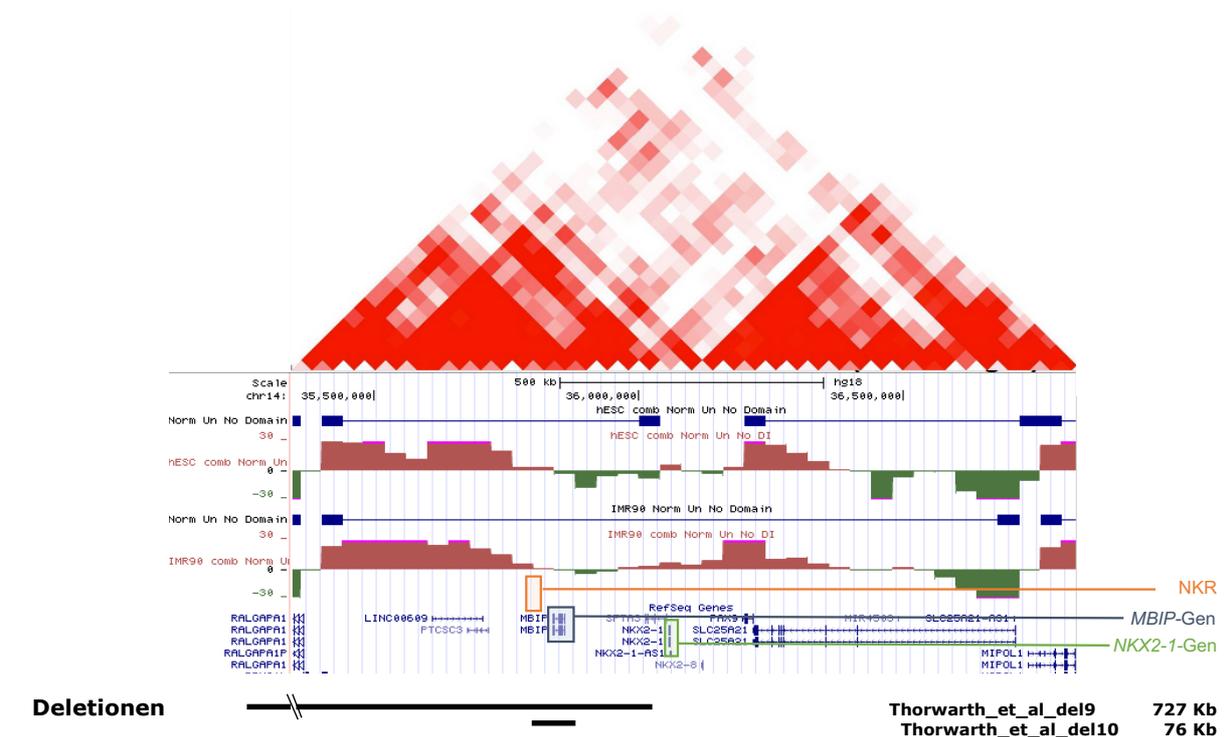


Abbildung 1-4: Funktionellen Domäne (TAD) des *NKX2-1*- und *Mbip*-Gens, sowie der nicht-kodierenden Region (NKR)

Dargestellt ist ein Ausschnitt des Chromosoms 14q13 im UCSC Genome Browser (NCBI36/hg18) mit dem *NKX2-1*-Gen und dem *Mbip*-Gen, sowie den in der Studie von Thorwarth et al. 2014 nachgewiesenen Deletionen downstream von *NKX2-1* und die nicht-kodierende Region (NKR), die Teil dieser Deletionen ist. Die funktionelle Chromatininteraktion der TAD wird als rotes Dreieck dargestellt, in dem die zwei kleineren Dreiecke die sogenannten sub-TADs darstellen.

Weiter konnten unveröffentlichte Daten aus *circularized chromosome conformation capture* (4C) Experimenten in Mausgewebe des IEPEs von Thorwarth et al. eine Interaktion zwischen dem Transkriptionsstartpunkt (TSS) des *NKX2-1*-Gens, dem *Mbip*-Gen und der nicht-kodierenden Region (NKR) downstream des *Mbip*- und des *NKX2-1*-Gens zeigen. Sowohl der TSS von *NKX2-1* als auch das *Mbip*-Gen „loopen“ demnach zu dieser downstream gelegenen nicht-kodierenden Region (NKR), die Teil der Deletionen in allen bisher beschriebenen Patienten mit einem *NKX2-1* defizienten Phänotyp ohne entsprechende *NKX2-1*-Mutation ist.

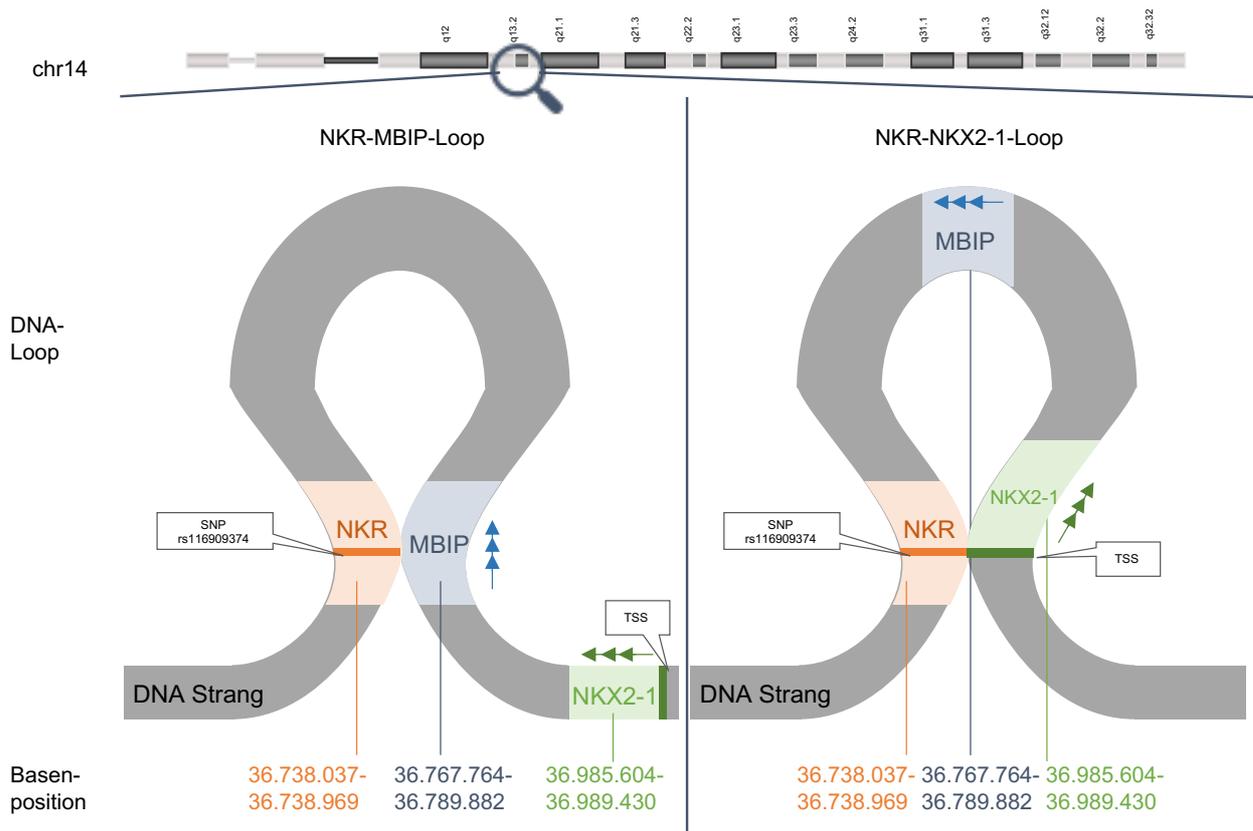


Abbildung 1-5: DNA-Interaktion zwischen *NKX2-1*, *Mbip* und der NKR downstream von *NKX2-1*

Dargestellt ist schematisch die Interaktion zwischen dem *NKX2-1*-Gen, dem *Mbip*-Gen und der nicht-kodierenden Region (NKR) downstream von *NKX2-1* nach unveröffentlichten Ergebnissen aus 4C-Experimenten des IEPEs von Thorwarth et al. Beide Gene und die NKR sind farblich markiert. Die Pfeile geben jeweils die Transkriptionsrichtung der Gene an. Der farblich hervorgehobene Balken innerhalb der NKR markiert den mit dem TSH-Wert assoziierten SNP rs116909374. Der farblich hervorgehobene Balken im *NKX2-1*-Gen markiert den Transkriptionsstartpunkt (TSS) des Gens. Die Basenposition gibt die genaue Position des *NKX2-1*-Gens, des *Mbip*-Gens und der NKR auf dem Chromosom 14q13 im UCSC Genome Browser (NCBI37/hg19) an.

Diese Ergebnisse des IEPEs von Thorwarth et al. liefern deutliche Hinweise für eine funktionelle Relevanz dieser nicht-kodierenden Region (NKR) in Bezug auf die Regulation der benachbarten Gene.

Neben den Ergebnissen des IEPEs von Thorwarth et al. konnten Gudmundsson et al. 2012 in dieser nicht-kodierenden Region (NKR) die Assoziation eines SNPs mit dem TSH-Wert nachweisen [41], die in einer Metaanalyse von Taylor et al. 2015 bestätigt wurde [42]. Als sogenannte „Single nucleotide polymorphisms“ (SNPs) werden einzelne Basenpositionen der genomischen DNA bezeichnet, für die verschiedene Allele existieren, deren selteneres Allel in mindestens einem Prozent der gesunden Individuen

einer Population zu finden ist [43]. Für den in der nicht-kodierenden Region (NKR) gelegenen SNP rs116909374 ist das seltenere T-Allel (Minor Allel) mit einem niedrigen TSH-Wert und einem erhöhten FT4-Wert assoziiert [41, 42]. Absolut betrachtet, trat das Minor Allel T in der Studie von Gudmundsson et al. 2012 in 4,4% der insgesamt 3190 bezüglich des in dieser Studie untersuchten Schilddrüsenkarzinoms gesunden Kontrollen auf [41]. Die Metaanalyse von Taylor et al. 2015 zeigte mit einer Minor Allel Frequenz (MAF) von 4,3% bei insgesamt 15037 untersuchten Proben eine mit den Ergebnissen von Gudmundsson et al. 2012 vergleichbare MAF für den mit dem TSH-Wert assoziierten SNP rs116909374 [42]. Die Minor Allel Frequenz (MAF) beschreibt dabei die Häufigkeit des Auftretens des zweit häufigsten Allels einer SNP. Wie auch in den vorliegenden Studien wird in vielen genomweiten Assoziationsstudien auf die Bestimmung des Haplotyps verzichtet, da das Einbeziehen des Haplotyps einerseits die Analyse erschwert und andererseits die Unterschiede der Ergebnisse im Vergleich zur alleinigen Bestimmung der Allelfrequenz gering sind [44].

Bei genauerer Betrachtung der Ergebnisse der 4C-Experimente des IEPES von Thorwarth et al. scheint die Chromatininteraktion des *Mbip*-Gens sowie des *NKX2-1*-Gens mit der nicht-kodierenden Region (NKR) downstream von *NKX2-1* auf die Umgebung dieses SNPs lokalisiert zu sein (siehe Abbildung 1-5).

Smemo et al. 2014 konnten unter anderem mit 4C-Experimenten zeigen, dass adipositasassoziierte SNPs in einer nicht-kodierenden Intronsequenz des *FTO*-Gens eine funktionelle Verbindung zu dem Megabasenpaaren entfernt liegenden *IRX3*-Gen haben und die Ausprägung der SNP-Allele die Expression von *IRX3* beeinflussen. Weiter konnte gezeigt werden, dass sich die adipositasassoziierten SNPs mit Enhancern überschneiden, so dass vermutet wird, dass es durch Allelvarianten der SNPs zu einer Störung der Enhanceraktivität und damit zu einer Veränderung der Expression des Zielgens kommen könnte [45].

Dieser zusätzliche Befund stützt die Annahme, dass es sich auch bei der nicht-kodierenden Region (NKR) des oben genannten *NKX2-1*-Gens um ein schilddrüsenspezifisches regulatives Element z.B. in Form eines schilddrüsenspezifischen Enhancers handeln könnte.

Die Korrelation des SNPs rs116909374 mit dem TSH- und FT4-Wert könnte, ähnlich wie es Weedon et al. 2014 für die Pankreasagenesie zeigen konnte [21], die Folge einer SNP-Allel abhängigen Beeinflussung eines schilddrüsenspezifischen Enhancers in der nicht-kodierenden Region (NKR) sein.

1.3.3 Mutationen in einem schilddrüsenspezifischen Enhancer als mögliche Erklärung der Schilddrüsendysgenesie

Ähnlich wie bei der Schilddrüsendysgenesie blieb die genetische Ursache der *isolierten* Pankreasagenesie trotz des vermehrten Auftretens in konsanguinen Familien und damit der Vermutung eines autosomal-rezessiven Defekts lange unentdeckt. Als Ursache der *syndromalen* Form der Pankreasagenesie konnten hingegen Mutationen in Genen, die unter anderem an der Pankreasentwicklung beteiligt sind, gefunden werden [21].

Weedon et al. 2014 konnten für die bis dahin fehlende genetische Grundlage der *isolierten* Form der Pankreasagenesie einen pankreasspezifischen Enhancer identifizieren, dessen mutationsbedingte Zerstörung ursächlich für die isolierte Form der Pankreasagenesie ist. Der durch Weedon et al. 2014 entdeckte Enhancer liegt in einem nicht-kodierenden Bereich downstream des unter anderem für die Pankreasentwicklung essentiellen *PTF1A*-Gens. Mutationen im *PTF1A*-Gen waren bereits als Ursache der *syndromalen* Pankreasagenesie bekannt. Zu der Entdeckung des pankreasspezifischen Enhancers führte unter anderem die Analyse von epigenetischen Chromatinmodifikationen, die mit Enhancerfunktionen assoziiert sind. So ist der 400 Basenpaare große Enhancer eine evolutionär konservierte Region, die einen hohen Anteil der enhancertypischen Chromatinmodifikationen H3K4me1 und H3K27ac zeigt und an den Transkriptionsfaktoren binden [21].

In der nicht-kodierenden Region (NKR) downstream des *NKX2-1*-Genes wurde aufgrund der bisherigen Ergebnisse des IEPEs von Thorwarth et al. ein schilddrüsenspezifischer Enhancer vermutet. Auch in weiteren Studien wird die Ursache des *NKX2-1* defizienten Phänotyps durch Deletionen downstream von *NKX2-1* durch eine Störung von Enhancerelementen diskutiert [29, 30, 32]. Es erfolgte daher die Analyse der nicht-kodierenden Region (NKR), die nachfolgend als potentielle Enhancer Region (PER) bezeichnet wird, bezüglich regulativer Elemente und der Bindung von Transkriptionsfaktoren im UCSC Genome Browser (NCBI GRCh37/hg19). Die

Betrachtung der potentiellen Enhancer Region (PER) im UCSC Genombrowser zeigt, dass es sich um einen 800 Basenpaare großen, evolutionär hochkonservierten Bereich der DNA handelt (siehe Abbildung 1-6). Pennacchio et al. 2006 konnten zeigen, dass es sich bei hochkonservierten nicht-kodierenden DNA-Bereichen häufig um gewebespezifische Enhancerelemente handelt [46]. Die Betrachtung der epigenetischen Chromatinmodifikationen zeigt, dass die PER mit der enhancertypischen Histonmodifikation H3K4me1 sowie in geringerem Ausmaß auch mit der in aktiven Promotoren vorkommenden Histonmodifikation H3K4me3 assoziiert ist (siehe Abbildung 1-6) [47]. H3K4me1 ist typischerweise in den flankierenden Nucleosomen von Enhancern zu finden, wohingegen H3K4me3 promotorassoziiert auftritt [48]. Weiter fand sich die Histonmodifikation H3K27ac, welche zusätzlich zu den oben genannten Histonmodifikationen bei Aktivierung der entsprechenden Enhancer oder Promotorelemente gefunden wird [48]. Es gibt keinen klaren Konsens darüber, welche Kombination der zahlreichen Histonmodifikationen für die Vorhersage von Enhancerelementen am besten ist [48]. Eine gängige Kombination ist die in der PER gefundene mit dem Vorhandensein von H3K4me1 und H3K27ac, geringen Mengen an H3K4me3 sowie dem Fehlen von H3K27me3 [48, 49]. Diese Charakteristika bestärken die Annahme einer möglichen Enhancerfunktion, wenn auch die Histonmodifikation H3K27ac nicht in hohem Maße in der PER zu finden ist. Zusätzlich weist die Bindung mehrerer Transkriptionsfaktoren an die PER auf ihre funktionelle Relevanz hin (siehe Abbildung 1-6).

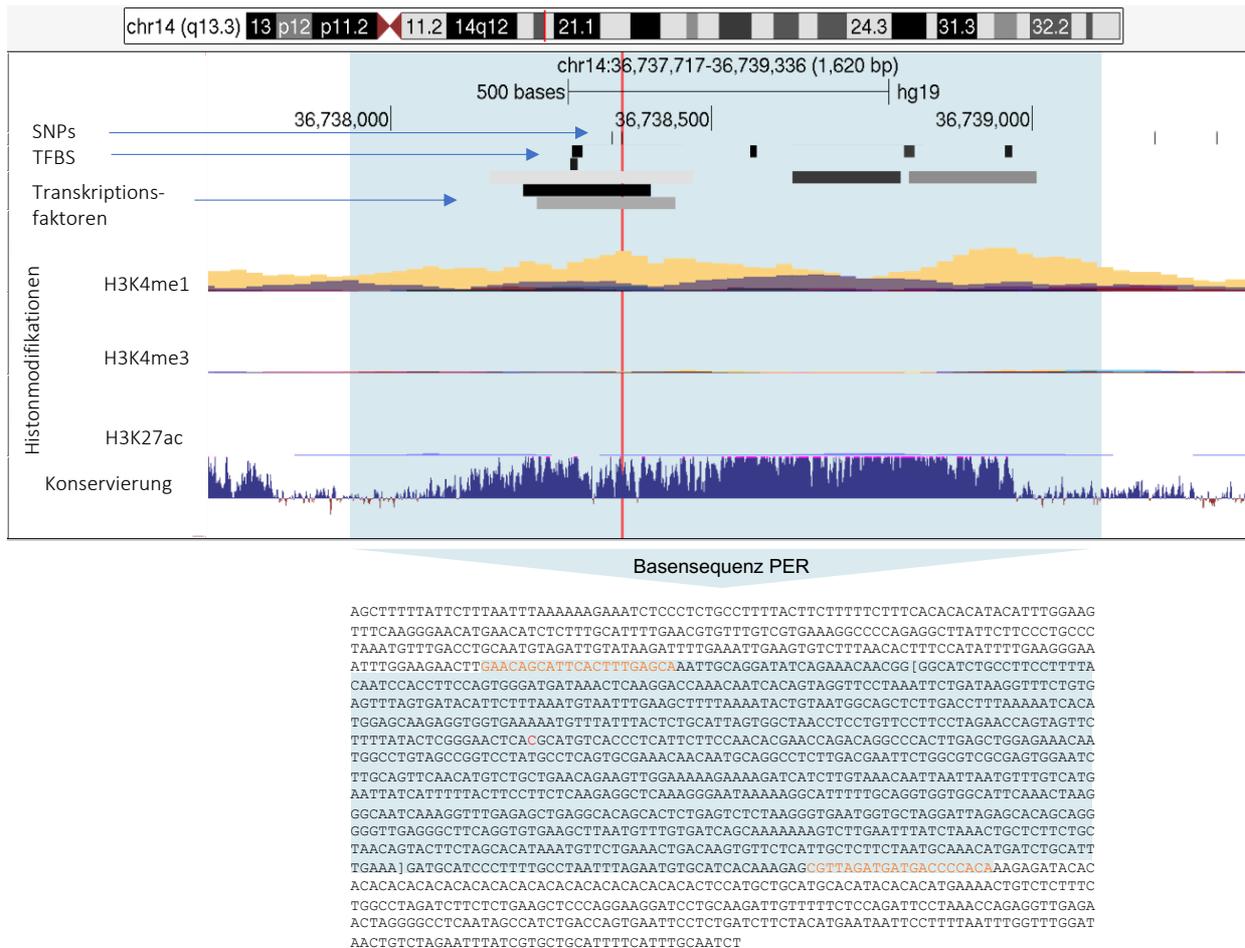


Abbildung 1-6: Potentielle Enhancer Region

Gezeigt ist ein Ausschnitt aus dem UCSC Genome Browser (NCBI GRCh37/hg19). In Grau hinterlegt ist die potentielle Enhancer Region (PER) chr14:36,738,037-36,738,969. Die senkrechte Linie kennzeichnet den mit dem TSH-Wert assoziierten SNP rs116909374. Weiter sind die SNPs mit einer MAF > 1% (dbSNP153), die potentiell bindenden Transkriptionsfaktoren, die Transkriptionsfaktorbindungsstellen (TFBS), die enhancerassozierten Histonmodifikationen und der Grad der Konservierung dargestellt. Im unteren Teil der Abbildung ist die Basensequenz der PER dargestellt. Orange markiert sind die verwendeten PCR-Primer. Die rot markierte Cytosin-Base kennzeichnet die Position des mit dem TSH-Wert assoziierten SNPs rs116909374.

1.4 Epigenetik

Der Begriff Epigenetik ist definiert als vererbare Genfunktionsänderungen, die ohne eine Veränderung der DNA-Sequenz erfolgen [50].

Es sind unterschiedliche Mechanismen bekannt, wie diese sequenzunabhängigen Genfunktionsänderungen vermittelt werden. Die kovalente DNA-Methylierung an der fünften Position der Cytosinbase an sogenannten CpG-Positionen, vermittelt durch Methyltransferasen, führt z.B. zu einer Inaktivierung der Genexpression, wohingegen unmethylierte CpG-Positionen mit einer aktiven Genexpression assoziiert sind [51].

Eine weitere wichtige Funktion bei der epigenetischen Genomregulation spielen Histonmodifikationen. Sie sind entscheidend an der chromatinabhängigen Genregulation beteiligt [52]. Verschiedene kovalente Modifikationen wie z.B. unter anderem die Methylierung, Acetylierung oder Phosphorylierung an Aminosäureresten der Histone werden beschrieben. Die Histonmodifikationen können einerseits mit einer Aktivierung und andererseits mit einer Repression der Transkription assoziiert sein [53]. Ihre Funktion wird dabei unter anderem durch die Rekrutierung von enzymatischen Proteinen sowie durch die Beeinflussung der Chromatinstruktur an sich vermittelt [54]. Die Wirkung der einzelnen Histonmodifikationen kann dabei allerdings z.T. je nach Kontext variieren [53]. Die Methylierung des Histonrestes H3K9 führt z.B. zu einer Repression in Promotorregionen und zu einer Aktivierung der Transkription in kodierenden Regionen [55]. In seiner repressiven Funktion ist H3K9me3 z.B. maßgeblich an der Heterochromatinbildung beteiligt [54]. Die Histonmodifikation H3K4me3 verhindert die Bindung repressiver Proteine an das Chromatin und ist dadurch mit aktiver Transkription assoziiert [54]. Die Wirkung der Histonmodifikationen ist insgesamt sehr komplex und dynamisch. So konnte sowohl eine Funktionsbeeinflussung der Histonmodifikationen untereinander [54] als auch eine Beeinflussung ihrer Wirkung durch die Methylierung der DNA gezeigt werden [56]. Dabei wird z.B. die Proteinbindung an die Histonmodifikation H3K9me3 durch methylierte DNA verstärkt [56].

Eine Beteiligung von weiteren epigenetischen Faktoren bei der Expressionsregulation von Genen wird beschrieben, so z.B. eine Beteiligung von nicht-kodierender RNA an der Transkriptionsrepression [52].

1.4.1 DNA-Methylierung

Die am besten untersuchte epigenetische Modifikation ist die enzymatische, kovalente DNA-Methylierung an der fünften Position der Base Cytosin durch DNA-Methyltransferasen [57]. Sie tritt vorwiegend an CpG-Positionen auf und ist wichtig für die Entwicklung, Differenzierung und Aufrechterhaltung der Zellidentität [50]. CpG-Positionen treten vor allem in kurzen CG-reichen DNA-Bereichen auf, die als CpG-Inseln bezeichnet werden und vorwiegend in Promotorregionen von Genen zu finden sind [58-61]. Der Großteil der CpG-Positionen im Genom ist methyliert. CpG-Inseln hingegen zeigen eine reduzierte Methylierung. Eine Methylierung der CpG-Inseln wird mit einer

Repression der Promotoraktivität und der entsprechenden Genfunktion in Zusammenhang gebracht [61, 62]. Die Methylierungsmuster in differenzierten somatischen Zellen sind im Allgemeinen stabil und [57] werden während der DNA-Replikation mit Hilfe von DNA-Methyltransferasen übertragen [50]. Krankheitsassoziierte Störungen der Genexpression durch Veränderungen der DNA-Methylierung werden beschrieben. Javierre et al. 2010 konnten in einer Studie mit monozygoten Zwillingen zeigen, dass eine veränderte DNA-Methylierung in Patienten mit systemischen Lupus erythematodes die Expression von krankheitsrelevanten Genen verändert [63].

Die Etablierung des Methylierungsmusters erfolgt während der Keimzellentwicklung und der frühen embryonalen Entwicklung und ist ein hoch dynamischer Prozess, bei dem es zu einer Veränderung der epigenetischen Information kommt. Der zugrunde liegende Mechanismus ist hierbei eine aktive oder passive Demethylierung, gefolgt von einer jeweils späteren zell- und gewebespezifischen Remethylierung. Diese Methylierungsveränderungen passieren einerseits während der Keimzellentwicklung und andererseits in der frühen Embryonalentwicklung [57]. Es wird angenommen, dass diese dynamischen Änderungen des Methylierungsmusters während der Entwicklung eine wichtige Rolle für die spätere gewebespezifische Differenzierung der Zellen spielen und somit einen erheblichen Einfluss auf die regelrechte embryonale Entwicklung und Zelldifferenzierung haben (siehe Abbildung 1-7) [57, 60].

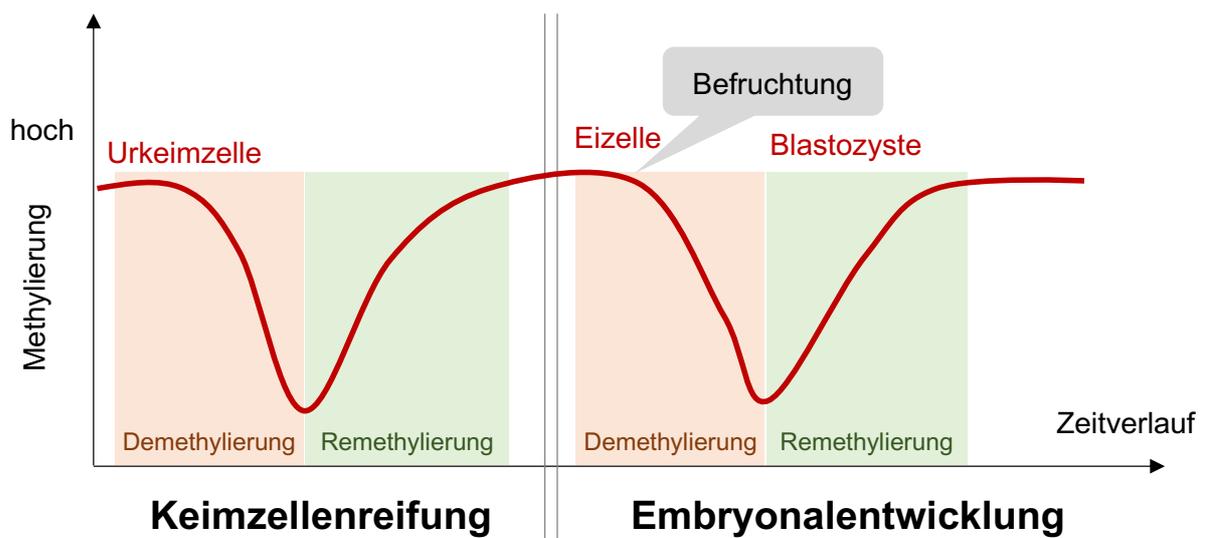


Abbildung 1-7: Etablierung des Methylierungsmusters

Dargestellt ist schematisch die Etablierung der DNA-Methylierung am Beispiel der weiblichen Keimzelle. Die Abbildung ist in Anlehnung an [57].

Neben dem Einfluss der DNA-Methylierung auf die in CpG-Inseln liegenden Promotorregionen von Genen ist eine Hypomethylierung von CpG-Positionen auch in anderen aktiven regulatorischen Genbereichen wie z.B. Enhancern zu finden. An 15-20% der CpG-Positionen im menschlichen Genom zeigen sich gewebespezifische Unterschiede in der Methylierungsintensität. Diese Regionen werden als „differentially methylated regions“ (DMRs) bezeichnet. Gewebespezifische DMRs sind durch einen zelllinienspezifischen Erwerb oder Verlust von DNA-Methylierung während der embryonalen Entwicklung bedingt. Sie zeigen zelltypspezifische regulatorische Funktionen z.B. in Form von Enhancern und haben Bindestellen für gewebespezifische Transkriptionsfaktoren [61]. Eine wechselseitige Beeinflussung der Transkriptionsfaktoren und der DNA-Methylierung wird vermutet. So wird vermutet, dass einerseits die gewebespezifische Hypomethylierung der DMRs durch eine spezifische Transkriptionsfaktorbindung vermittelt wird und andererseits die DNA-Methylierung Einfluss auf die Bindungsaffinität und Spezifität von Transkriptionsfaktoren nimmt [64].

Ein Beispiel für ein weitverbreitetes epigenetisches Phänomen, bei dem es unter anderen durch allelspezifische Methylierungsveränderungen zu der alleinigen Expression des maternalen oder des paternalen Allels eines Genes kommt, ist das Imprinting von Genen.

So kann ein Gen durch DNA-Methylierung und Histonmodifikationen z.B. auf dem paternalen Allel inaktiviert sein, sodass nur das maternale Gen exprimiert wird und anders herum [65].

Bell et al. 2000 konnten zeigen, dass eine Hypermethylierung der CTCF-Bindungsstelle zu einem Verlust der CTCF-Bindung führt und damit zu einem Verlust der enhancerblockierenden Wirkung von CTCF. Die Folge ist die Aktivierung eines durch Imprinting inaktivierten paternalen Gens, was wiederum Auslöser der Imprinting Erkrankung des Beckwith-Wiedemann-Syndroms ist [66]. Dies ist ein Beispiel für eine gestörte Protein-DNA-Interaktion, die zu einer krankheitsauslösenden Enhanceraktivierung führt.

Neben dem Phänomen des Imprintings von Genen durch allelspezifische Methylierungsveränderungen sind metastabile Epiallele ein weiteres relevantes epigenetisches Phänomen der Expressionsregulation von Allelen, die durch Methylierungsunterschiede bedingt sind. An metastabilen Epiallelen erfolgt die Remethylierung in der frühen embryonalen Entwicklung anders als z.B. beim Imprinting von Genen zufällig [67]. Die dadurch entstandenen Methylierungsunterschiede haben dabei Einfluss auf die phänotypische Allelausprägung [67]. Ein bekanntes tierexperimentelles Modell ist die Agouti Maus. Es konnte gezeigt werden, dass der Grad der Methylierung des A^{vy} -Allels, welches durch das Vorhandensein eines Retrotransposons gekennzeichnet ist, Einfluss auf die Fellfarbe der Maus hat [68]. Das Retrotransposon im A^{vy} -Allel besitzt einen Promotor, der die Expression des Agouti-Gens induzieren kann [67]. Da das Agouti-Gen ein Antagonist der Melanocortin-Rezeptoren ist, resultiert aus der ektopen Expression des Agouti-Gens eine gelbe hypopigmentierte Fellfarbe. Eine Hypermethylierung des Retrotransposons führt zu einer graduellen Inaktivierung des Promotors und somit zu einer phänotypischen graduellen Annäherung an die Fellfarbe des Wildtyps [68]. Es wird vermutet, dass metastabile Epiallele an der Entstehung von humanen Erkrankungen beteiligt sind, deren Vererbungsmodi bisher ungeklärt sind [67]. Bisher konnte z.B. für den menschlichen BMI, für den eine starke Erbllichkeit vermutet wird, nur ein kleiner Anteil durch genetische Varianten erklärt werden. Kühnen et al. 2016 konnten Hinweise dafür finden, dass die quantitative Methylierung einer im Sinne eines vermuteten metastabilen Epiallels einer variabel methylierten Region des *POMC*-Gens mit dem BMI assoziiert ist [69].

Monozygote Zwillinge stammen von derselben zygotischen Zelle ab und haben daher ein nahezu identisches Genom; dennoch bestehen phänotypische Unterschiede bis hin zur Diskordanz bezüglich einer Vielzahl von Krankheiten wie z.B. der kongenitalen Hypothyreose [70]. Epigenetische Modifikationen, wie z.B. die DNA-Methylierung oder Histonmodifikationen, werden durch dynamische Veränderungen während der Keimzellreifung und frühen embryonalen Entwicklung sowie durch Umweltfaktoren, hormonelle Veränderungen und stochastische Ereignisse beeinflusst und sind daher deutlich instabiler als genomische Faktoren. Es wird daher angenommen, dass phänotypische Unterschiede von monozygoten Zwillingen durch epigenetische Unterschiede bedingt sind [71]. Kaminsky et al. 2009 konnten bei monozygoten Zwillingspaaren in verschiedenen Geweben aller drei Keimblätter Methylierungsunterschiede nachweisen [71], was für eine mögliche Beteiligung von Methylierungsunterschieden bei der Entstehung der Diskordanz spricht. Rakyán et al. 2011 konnten Methylierungsunterschiede zwischen monozygoten Zwillingen in für den Autoimmunprozess bei Typ 1 Diabetes relevante CD14+ Monozyten nachweisen, die mit krankheitsrelevanten Genen assoziiert sind. Weiter konnte gezeigt werden, dass diese entsprechenden Methylierungsunterschiede bei Kindern mit Nachweis von diabetestypischen Antikörpern schon vor der Manifestation der Erkrankung zu finden sind. Dieser Befund spricht für eine mögliche Krankheitsrelevanz der gefundenen Methylierungsunterschiede [72]. Zusammenfassend sind diskordante monozygote Zwillinge somit von großer Relevanz für die Detektion möglicher krankheitsverursachender epigenetischer Störungen.

1.4.2 DNA-Methylierungsdefekte als Ursache der Schilddrüsendysgenese

Auch für die Schilddrüsendysgenese wird eine Diskordanz von monozygoten Zwillingen beschrieben [14]. Für die Ursache der Schilddrüsendysgenese wird daher eine Beteiligung von epigenetischen Faktoren diskutiert [10]. Die in der potentiellen Enhancer Region (PER) des *NKX2-1*-Gens vermutete schilddrüsenspezifische Enhancerfunktion könnte somit neben einer möglichen genetischen Störung des Enhancers auch durch DNA-Methylierungsdefekte in ihrer Funktion gestört werden. Eine Hypermethylierung der PER mit einer gestörten Bindungsaffinität von schilddrüsenspezifischen Transkriptionsfaktoren könnte die Folge sein. Auch die in 4C-Experimenten des IEPes von Thorwarth et al. nachgewiesenen DNA-Interaktionen zwischen der PER, dem *NKX2-1*-Gen und dem *Mbip*-Gen könnten durch einen Methylierungsdefekt beeinflusst werden.

1.5 Zielsetzung

In Zusammenschau der bisherigen Erkenntnisse könnte es sich bei der konservierten nicht-kodierenden Region (PER) des *NKX2-1*-Gens um einen schilddrüsenspezifischen Enhancer handeln, dessen Störung zur isolierten Form der Schilddrüsendysgenese führt und nicht zu der durch intragenische *NKX2-1*-Mutationen ausgelösten syndromalen Form der Schilddrüsendysgenese im Rahmen des Gehirn-Lungen-Schilddrüsen-Syndroms. Der in der potentiellen Enhancer Region (PER) lokalisierte, mit dem TSH- und FT4-Wert assoziierte SNP rs116909374, könnte dabei je nach Allelausprägung zu einer Störung des postulierten schilddrüsenspezifischen Enhancers führen und damit die Expression von schilddrüsenspezifischen Genen beeinflussen, wie es Smemo et al. 2014 für einen adipositasassoziierten SNP [45] und Weedon et al. 2014 für die mutationsbedingte Störung eines Enhancers als Ursache der isolierten Pankreasagenese [21] zeigen konnten. Da das Minor Allel T des SNPs rs116909374 mit einem erniedrigten TSH-Wert und einem erhöhten FT4-Wert korreliert, ergibt sich für die vorliegende Arbeit die Hypothese, dass das Minor Allel T häufiger bei schilddrüsengesunden Kontrollen und das Major Allel C häufiger bei CH-Patienten zu finden ist. Für die vorliegende Arbeit ist daher zu erwarten, dass die Minor Allel Frequenz (MAF) der schilddrüsengesunden Kontrollen annähernd übereinstimmend ist mit den Ergebnissen von Gudmundsson et al. 2012 [41] und Taylor et al. 2015 [42] und die CH-Patienten eine geringere MAF aufweisen.

Basierend auf diesen Annahmen ist die zentrale Hypothese der Arbeit folgende:

Der Genomabschnitt chr14:36,738,037-36,738,969 stellt einen potentiellen schilddrüsenaktiven Enhancer des *NKX2-1*-Gens dar, der durch genetische oder epigenetische Veränderungen funktionell beeinträchtigt werden könnte mit der Folge einer Schilddrüsenstörung.

Eine graphische Zusammenfassung der zentralen Hypothese und Zielsetzung der Arbeit erfolgt in der Abbildung 1-8. Die Hypothese wurde in Analogie zu Weedon et al. 2014 entwickelt. Diese konnten als Ursache der *isolierten* Pankreasagenese die Störung eines pankreasspezifischen Enhancers identifizieren, der in einer nicht-kodierenden Region downstream des für die Pankreasentwicklung essentiellen Transkriptionsfaktor PTF1A

lokalisiert ist. In Analogie zur Schilddrüsendysgenese waren *PTF1A*-Mutationen bereits als Ursache für die *syndromale* Pankreasagenese bekannt.

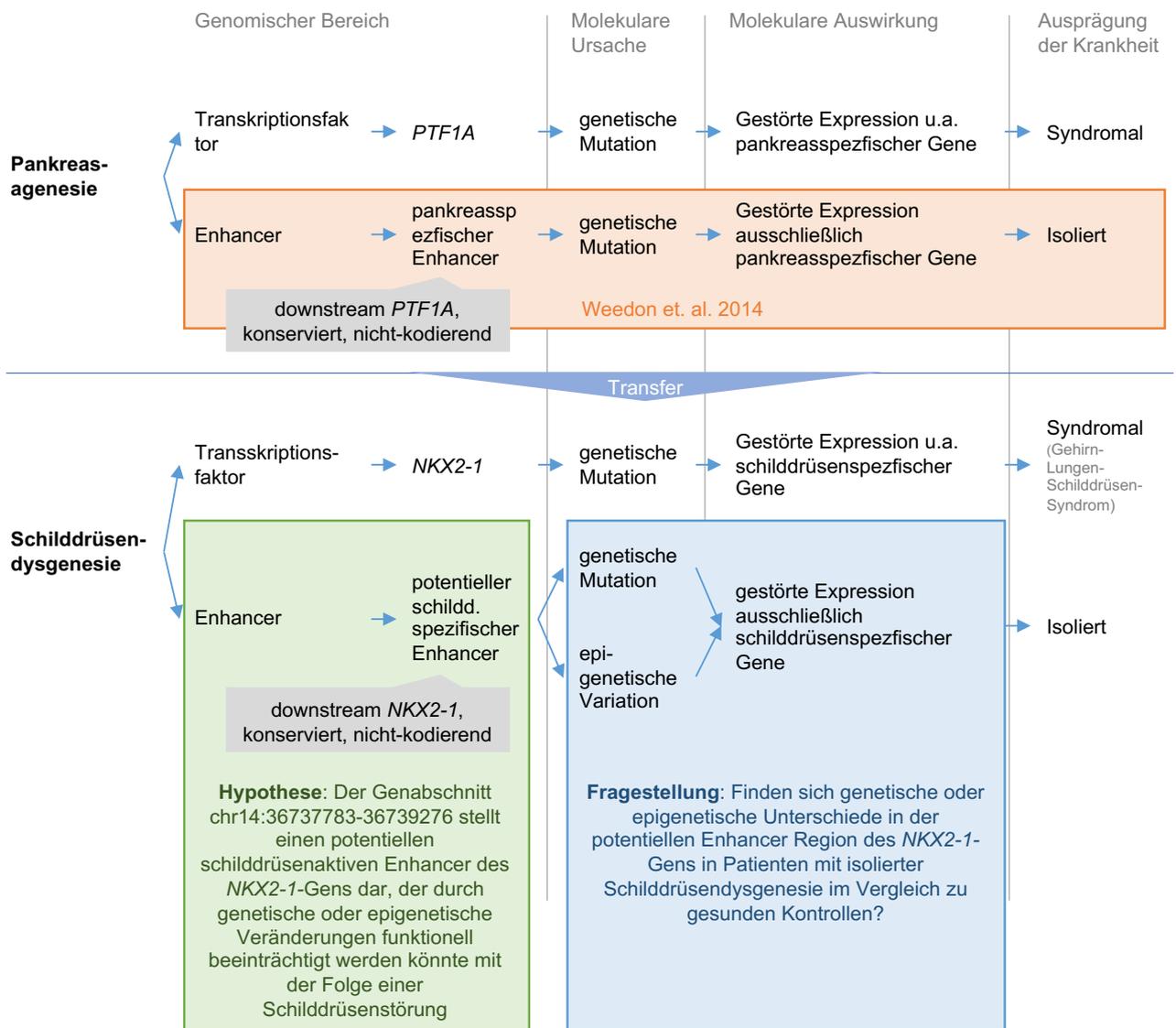


Abbildung 1-8: Hypothese und Fragestellung

Dargestellt ist, vergleichend mit den Ergebnissen von Weedon et al. 2014 [21], die zentrale Hypothese und Fragestellung der Arbeit.

Zur Beantwortung der zentralen Fragestellung der Arbeit erfolgte eine genetische und epigenetische Analyse der potentiellen Enhancer Region (PER) in einer der größten Kohorten an Patienten mit isolierter Schilddrüsendysgenese. Zur weiteren Einordnung der Ergebnisse wurden zusätzlich sechs für die isolierte Schilddrüsendysgenese diskordante monozygote Zwillingspaare in die Analyse mit einbezogen.

Um einen möglichen Zusammenhang des mit dem TSH-Wert assoziierten SNPs rs116909374 mit einer milden Störung der Schilddrüsenfunktion zu analysieren wie auch Mutationen im *NKX2-1*-Gen, die im Rahmen des Gehirn-Lungen-Schilddrüsen-Syndroms häufig zu einer Schilddrüsenstörung in Form einer Hyperthyreotropinämie führen [10, 19], wurden zusätzlich Patienten mit einer Hyperthyreotropinämie in die Analyse mit einbezogen. Die milde Störung der Schilddrüse im Rahmen der Hyperthyreotropinämie ist durch eine Erhöhung des TSH-Wertes bei normwertigen peripheren Schilddrüsenhormonen gekennzeichnet [4]. So könnte die Schwere der Schilddrüsenfunktionsstörung z.B. mit der Größe oder dem Ort der Mutation in der potentiellen Enhancer Region (PER) korrelieren. Weiter könnte es auch sein, dass eine isolierte Störung der PER zu einer Hyperthyreotropinämie führt und es erst im Rahmen einer weiteren Mutation zu der isolierten Schilddrüsendysgenese als schwere Beeinträchtigung der Schilddrüsenfunktion kommt.

Zur vergleichenden Einordnung der Ergebnisse wurden zusätzlich eine Kohorte schilddrüsengesunder Personen sowie das krankheitsrelevante Gewebe der Schilddrüse in schilddrüsengesunden Personen analysiert.

Um eine gute Einordnung und vergleichende Interpretation der Ergebnisse bezüglich möglicher krankheitsrelevanter Befunde zu gewährleisten, wurden somit insgesamt drei verschiedene Kohorten für die Schilddrüsendysgenese diskordante monozygote Zwillinge und Schilddrüsengewebe von schilddrüsengesunden Personen in die Analyse mit einbezogen (siehe Tabelle 2-4).

1.5.1 Genetische Analyse der potentiellen Enhancer Region

Im ersten Teil der Arbeit erfolgte die Analyse auf mögliche genetische Mutationen oder SNP-Varianten in der PER mittels Sanger-Sequenzierung. Ziel hierbei war die Beantwortung folgender Fragen:

1. Finden sich genetische Unterschiede in der potentiellen Enhancer Region des *NKX2-1*-Gens in Patienten mit isolierter Schilddrüsendysgenese oder Patienten mit Hyperthyreotropinämie im Vergleich zu gesunden Kontrollen?
 - a. In Form von Mutationen
 - b. In Form von SNP-Variationen

- c. Unterscheidet sich die Allel-Ausprägung des mit dem TSH-Wert assoziierten SNP rs116909374 zwischen den Kohorten?
- d. Besteht ein Zusammenhang zwischen der Allel-Ausprägung des mit dem TSH-Wert assoziierten SNP rs116909374 und dem morphologischen Schilddrüsenphänotyp bei Patienten mit isolierter Schilddrüsendysgenese?

1.5.2 Epigenetische Analyse der potentiellen Enhancer Region

Da für die isolierte Schilddrüsendysgenese ein mendelscher Vererbungsmodus als alleinige Ursache eher unwahrscheinlich ist, wird immer wieder die Möglichkeit der Beteiligung von Epimutationen bei der Entstehung der isolierten Schilddrüsendysgenese diskutiert [14]. Daher erfolgte im zweiten Teil dieser Arbeit die Analyse des epigenetischen Methylierungsmusters der potentiellen Enhancer Region mittels Pyrosequenzierung zur Beantwortung folgender Fragen:

1. Finden sich epigenetische Methylierungsunterschiede in der potentiellen Enhancer Region des *NKX2-1*-Gens in Patienten mit isolierter Schilddrüsendysgenese oder Patienten mit Hyperthyreotropinämie im Vergleich zu gesunden Kontrollen sowie zwischen für die isolierte Schilddrüsendysgenese diskordanten monozygoten Zwillingen?
 - a. Im Mittelwert der Methylierung der neun CpG-Positionen
 - b. CpG-positionsspezifisch
2. Gibt es kohortenübergreifend Unterschiede in der Methylierung der verschiedenen CpG-Positionen in der potentiellen Enhancer Region des *NKX2-1*-Gens?
3. Tritt das Methylierungsmuster der potentiellen Enhancer Region des *NKX2-1*-Gens ubiquitär auf oder gibt es gewebespezifische Unterschiede?
4. Geht das bevorzugte Auftreten der CH bei Mädchen mit geschlechtsspezifischen Methylierungsunterschieden z.B. in der PER des *NKX2-1*-Gens einher?

Zur Beantwortung der Fragen erfolgte eine genetische und epigenetische Analyse der potentiellen Enhancer Region mittels Sanger-Sequenzierung und Pyrosequenzierung.

Analysiert wurden Patienten mit isolierter Schilddrüsendysgenese sowie sechs für die isolierte Schilddrüsendysgenese diskordante monozygote Zwillingspaare, Patienten mit Hyperthyreotropinämie sowie schilddrüsende Personen und Schilddrüsende von schilddrüsenden Personen.

Nachfolgend werden die angewandten Methoden erklärt, die Ergebnisse dargestellt und diese schließlich diskutiert.

2 Material und Methoden

Als Erstes erfolgt die Zusammenstellung der verwendeten Geräte sowie die Darstellung der für die Analysen verwendeten Kohorten. Im zweiten Teil erfolgt schließlich die Erklärung der verwendeten Methoden. Dabei werden zuerst die verwendeten Methoden der Sanger-Sequenzierung dargestellt und anschließend die Methodik der Pyrosequenzierung erklärt.

Die vorliegende Arbeit wurde von der örtlichen Ethikkommission genehmigt. Ethikantragsnummer: EA2/131/11.

2.1 Verwendete Geräte

Gerät	Bezeichnung	Hersteller
Pipetten	Matrix Multipipette	Thermo Scientific, MA, USA
	Eppendorf Research Plus	Eppendorf, Hamburg
Waage	Kern PCB 3500-2	Kern & Sohn, Balingen
Fluorometer	Quantus	Promega, Madison, WI, USA
pH-Meter	Seven easy	Mettler Toledo, Gießen
Vakuumpumpe	2522C-02	Welch, Niles, USA
Mikrowelle	MW 802	Exquisit, Kaarst
Tiefkühlschrank	Forma -86C ULT Freezer	Thermo Scientific, MA, USA
Wasseraufbereitungssystem	H2O (MilliQ) Barnstead GenPure Standard	Thermo Scientific, MA, USA
Tischzentrifugen	Centrifuge 5417R/ 5430R	Eppendorf, Hamburg
Minizentrifuge	Sprout mini centrifuge 12 V	Heathrow Scientific, IL, USA
	Combispin FVL-2400 N	Peglab, Erlangen
Vortexer	Vortex Genie 2	Bender & Hobein
Magnetrührer	Magnetrührer IKA-Combimag RCT	Janke & Kunkel, Staufen
Magnetständer	Diagenode	Diagenode, Belgien
Elektrophoresekamera	Whatman compact XS/S/XL	Biometra, Göttingen
Spannungsquelle	Microcomputer Electrophoresis power supply 500/500	Renner, Göttingen
Geldokumentationssystem	Gene Flash Network	Syngene, Cambridge, UK
	Transilluminator Bioview UST-20M-8E	Biostep, Burkhardtsdorf

Bildwandler	Gerix 1000	Biostep, Burkhardtsdorf
DNA - Präparationsstation	Maxwell 16	Promega, Madison, WI, USA
Thermocycler	Biometra UNO – Thermoblock Mastercycler gradient Mastercycler egradient S Mastercycler nexus gradient	Biometra, Göttingen Eppendorf, Hamburg Eppendorf, Hamburg Eppendorf, Hamburg
Heizblock	UNO-Thermoblock Thermomixer F 1.5/ F1.6A/ FP	Biometra, Göttingen Eppendorf, Hamburg
Probenaufbereitung station	PyroMark Q24 Vacuum Workstation (220 V)	Qiagen, Hilden
Sequenzierer	3130 XL Genetic Analyzer PyroMark Q24	Applied Biosystems, Weiterstadt Qiagen, Hilden

Tabelle 2-1: Verwendete Geräte

2.2 Verwendetes Material

Bezeichnung	Hersteller
Oligonukleotide	
Primer: Sequenzen siehe Tabelle 2-3	Sigma Aldrich, Taufkirchen
EasyLadder I	Bioline, Luckenwalde
PyroMark Control Oligo	Qiagen, Hilden
Enzyme	
MangoTaq-DNA-Polymerase	Bioline, Luckenwalde
Exonuklease I	New England Biolabs, MA, USA
USB Shrimp Alkaline Phosphatase	Affymetrix, Santa Clara, CA, USA
Enzypuffer	
Mango Taq Reaction Buffer 5x colourless	Bioline, Luckenwalde
Mango Taq Reaction Buffer 5x coloured	Bioline, Luckenwalde
MgCl ₂	Bioline, Luckenwalde
5x Sequencing Buffer	Applied Biosystems, Weiterstadt
PyroMark Denaturation Solution	Qiagen, Hilden
PyroMark Wash Buffer (10x)	Qiagen, Hilden
PyroMark Annealing Buffer	Qiagen, Hilden
PyroMark Binding Buffer	Qiagen, Hilden

Reagenziensysteme

ABI Prism BigDye Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems, CA, USA
Maxwell 16 Tissue DNA Purification Kit	Promega, Madison, WI, USA
EZ DNA Methylation Kit	Zymo Research, CA, USA
PyroMark CpG Assay	Qiagen, Hilden
EpiTect Bisulfite Kit	Qiagen, Hilden
PyroMark Gold Q24 Reagents	Qiagen, Hilden
AMPure® XP- Kit	Agencourt, Krefeld
Streptavidin Sepharose High Performance	GE Healthcare, Schweden
QuantiFluor dsDNA System	Promega, Madison, WI, USA

Sonstige Reagenzien

dNTPs	Bioline, Luckenwalde
Ethidiumbromid	Sigma-Aldrich, Taufkirchen
H2O (HPLC)	Roth, Karlsruhe
EDTA Pulver	Biorad, Feldkirchen
NaOH Granulat	Roth, Karlsruhe
Borsäure	Roth, Karlsruhe
Tris-Base	Roth, Karlsruhe
100x Tris-EDTA-Puffer pH 8.0	AppliChem
Ethanol 96%	Roth, Karlsruhe
Natriumacetat 1,5 M pH 5,5 Pulver	Roth, Karlsruhe
Bromphenolblau	Sonstige
Xylencyanol	Sonstige
Glycerol	Sonstige

Lösungsansätze für die Gelelektrophorese

5x TBE Zusammensetzung	
Tris	54 g/l
H3BO3 (Borsäure)	27,5 g
EDTA (0,5 M)	20 ml/l
in H2O	
pH 8	
Loading-DYE Zusammensetzung	
Bromphenolblau	0,25%
Xylencyanol	0,25%
Glycerol	25,00%
in H2O	

Außerordentliches Verbrauchsmaterial

Eppis Multiply-Pro 0,5ml Biosphere	Sarstedt, Nümbrecht
Eppis Micro tube 1,5ml	Sarstedt
8er Streifen Multiply strip 0,2 ml Kette	Sarstedt
PyroMark Q24 Cartridge	Qiagen, Hilden
PyroMark Q24 Plate	Qiagen, Hilden
PyroMark Q24 Vacuum Prep Troughs	Qiagen, Hilden
Agarose NEEO Ultra-Qualität	Roth, Karlsruhe

Verwendete Software

Sequence Pilot JSI	JSU Medical Systems ,MA, USA
GraphPad Software Version 6-9	GraphPad Software, CA, USA
Primer3web 4.0.0	Michelstadt
Pyrosequencing Assay Design Software	Qiagen, Hilden
Bisulfite-Primer-Seeker	Zymo Research, CA, USA
UCSC Genomebrowser (GRCh37/hg19)	University of California, CA, USA
Microsoft Office	Microsoft , WA, USA

Tabelle 2-2: Verwendetes Material

2.2.1 Verwendete Primer

Sämtliche verwendeten Oligonukleotide wurden von der Firma Sigma Aldrich (Taufkirchen) synthetisiert und sind in folgender Tabelle 2-3 aufgeführt.

Name	Sequenz (5' --> 3')	Annealing Temperatur	Amplikon Größe
PCR-Primer für die Sanger-Sequenzierung			
TSH_SNP1F	GAACAGCATTCACTTTGAGCA	59,5°C 63,1°C	bis 937 bp
TSH_SNP1R	TGTGGGGTCATCATCTAACG		937 bp
TSH_SNP2F	AGGATTAGAGCACAGCAGGG		937 bp
TSH_SNP2R	TTCACCACCTCTTGCTCCAT		937 bp
TSH_SNP3F	ATGCAGGCCTCTTGACGAAT		937 bp
TSH_SNP3R	CAGGAGGTTAGCCACTAATGC		937 bp
PCR-Primer für die Bisulfitsequenzierung			
Pyro_Sarah_F_SNP_Z1_F	ATTTTAAAAATTATATGGAGTAAGA GGTGGTG	62,2°	309 bp
Pyro_Sarah_F_SNP_Z1_R	AAAATAATCTTTTCTTTTCCAACCTTC TATTC AAC	62,2°	310 bp

Pyro_Sarah_F_SNP_Z2 _F	ATTGTAATGGTAGTTTTTGATTTTAA AAAATTATATGG		328 bp
Pyro_Sarah_F_SNP_Z2 _R	AAAATAATCTTTTCTTTTCCAACCTC TATTCAAC		328 bp
Pyro_Sarah_F_SNP_Z3 _F	TAGTAGGTTTTTAAATTTTGATAAGG TTTTTGTGAG		284 bp
Pyro_Sarah_F_SNP_Z3 _R	CTACAAACCATTATTTCTCCAACCTCA AATAAACC		284 bp
Pyro_Sarah_F_SNP_P4 _F	TTATATGGAGTAAGAGGTGGTGAAA	57,4°C	292 bp
Pyro_Sarah_F_SNP_P4 _R	ATCTTTTCTTTTCCAACCTTCTATTCA		292 bp
Pyro_Sarah_F_SNP_P5 _F	ATGGAGTAAGAGGTGGTGAAAA		288 bp

Für die Pyrosequenzierung biotinylierte Bisulfit-PCR-Primer

Pyro_Sarah_F_SNP_Z1 b_R1	AAAATAATCTTTTCTTTTCCAACCTC TATTCAAC		
Pyro_Sarah_F_SNP_P4 b_R1	ATCTTTTCTTTTCCAACCTTCTATTCA	57,4°C	292 bp
Pyro_Sarah_F_SNP_Z1 b_F	ATTTTTAAAAATTATATGGAGTAAGA GGTGGTG		
Pyro_Sarah_F_SNP_Z1 b_R2	AAAATAATCTTTTCTTTTCCAACCTC TATTCAAC		
Pyro_Sarah_F_SNP_P4 b_F	TTATATGGAGTAAGAGGTGGTGAAA		

Sequenzierprimer für die Sanger-Sequenzierung

TSH_SNP1F	GAACAGCATTCACTTTGAGCA	937 bp
TSH_SNP1R	TGTGGGGTCATCATCTAACG	937 bp
TSH_SNP2F	AGGATTAGAGCACAGCAGGG	937 bp
TSH_SNP2R	TTCACCACCTCTTGCTCCAT	937 bp

Sequenzierprimer für die Pyrosequenzierung

Pyro_Sarah_F_SNP_S1 _F	GTTTTTTTTTAGAATTAGTAGTTTT	96 bp
Pyro_Sarah_F_SNP_S2 _F	GTATGTTATTTTTATTTTTTAATA	
Pyro_Sarah_F_SNP_S3 _F	TATTTGAGTTGGAGAAATAATGGTT	
Pyro_Sarah_F_SNP_S4 _F	GAAATAATAATGTAGGTTT	26 bp
Pyro_Sarah_F_SNP_S5 _F	TAGAATTAGTAGTTTTTTTATATT	
Pyro_Sarah_F_SNP_S6 _F	AGAATTAGTAGTTT	

Tabelle 2-3: Verwendete Primer

Grün hinterlegte Primer wurden für die Analysen verwendet. Im Primernamen steht der letzte Großbuchstabe für die Richtung, F steht für Vorwärtsprimer, R für Rückwärtsprimer

2.3 Kohortenverzeichnis

In der Tabelle 2-4 sind die zur Analyse verwendeten Kohorten aufgeführt. Für die absolute Anzahl der Patienten ist zu beachten, dass in der Methylierungsanalyse die gesunden Zwillinge den gesunden Kontrollen zugeordnet wurden und die Zwillinge mit Schilddrüsendysgenese der CH-Kohorte.

Geschl.	Alter	Anzahl	Kohorte	Diagnose	Morphologie
weiblich	22,5	16	Kontrollen	SD gesund	normal
männlich	24,3	14	Kontrollen	SD gesund	normal
gesamt	23,3	30	Kontrollen	SD gesund	
weiblich	8,3	6	HTT	HTT	normal
männlich	5,3	7	HTT	HTT	normal
n.a.	12,2	1	HTT	HTT	normal
gesamt	6,8	14	HTT	HTT	
weiblich	10,2	32	CH	SDDG	Athyreose
weiblich	11,9	5	CH	SDDG	Ektopie
weiblich	7,5	20	CH	SDDG	Hypoplasie

weiblich	8,5	2	CH	SDDG	Hemithyroidea
weiblich	3,9	8	CH	SDDG	n.a.
gesamt	9,0	67	CH	SDDG	
weiblich					
männlich	11,7	9	CH	SDDG	Athyreose
männlich	5,3	1	CH	SDDG	Ektopie
männlich	7,1	12	CH	SDDG	Hypoplasie
männlich	9,3	1	CH	SDDG	Hemithyroidea
männlich	5,6	2	CH	SDDG	Hemihypoplasie
männlich	13,0	6	CH	SDDG	n.a.
gesamt	9,5	31	CH	SDDG	
männlich					
gesamt	9,2	98	CH	SDDG	
<i>Gesamt</i>					
männlich	12,1	52			
weiblich	11,6	89			
n.b.	12,2	1			
Gesamt	11,8	142			

Tabelle 2-4: Kohortenverzeichnis

Abkürzungen: Geschl.: Geschlecht; Alter: durchschnittliches Alter in Jahren der jeweiligen Kohorte zum Zeitpunkt der Blutentnahme; n.a.: nicht angegeben; SD gesund: Schilddrüsengesund; HTT: Hyperthyreotropinämie; CH: kongenitale Hypothyreose; SDDG: Schilddrüsendysgenesie

Die Zwillingspaare sind monozygot und hinsichtlich der primären kongenitalen Hypothyreose diskordant. Die nicht betroffenen Zwillinge sind schilddrüsengesund.

Zwillingspaar	Geschlecht	Diagnose Zwilling	erkrankter Zwilling	Schilddrüsenmorphologie erkrankter Zwilling
1	weiblich	Schilddrüsendysgenese	nicht bekannt	
2	weiblich	Schilddrüsendysgenese	Ektopie	
3	männlich	Schilddrüsendysgenese	nicht bekannt	
4	männlich	Schilddrüsendysgenese	Hypoplasie	
5	weiblich	Schilddrüsendysgenese	Athyreose	
6	weiblich	Schilddrüsendysgenese	nicht bekannt	

Tabelle 2-5: Monozygote diskordante Zwillingspaare

2.4 Erklärung und Durchführung der angewandten Methoden

Nachfolgend werden die angewandten Methoden erklärt. Dabei werden im ersten Teil die für die genetische Analyse verwendeten Methoden dargestellt und im zweiten Teil die für die epigenetische Analyse verwendeten Methoden erklärt.

2.4.1 Genotypisierung

Für die Genotypisierung der potentiellen Enhancer Region (PER) erfolgte die Sanger-Sequenzierung aller verwendeten DNA-Proben der drei Patientenkohorten. Analysiert wurde dabei die PER des *NKX2-1*-Gens sowie der in diesem Bereich liegende SNP rs116909374 aufgrund seiner Korrelation mit der Höhe des TSH-Serumwertes [41, 42]. Weiter diente die Genotypisierung dem Ausschluss mutationsbedingter Methylierungsunterschiede, die für die nachfolgende Pyrosequenzierung von Bedeutung sind.

2.4.1.1 DNA Isolation aus Leukozyten

Die Extraktion der Patienten-DNA erfolgte aus EDTA-Vollblutproben mit Hilfe des Maxwell® Tissue DNA Purification Kits der Firma Promega.

2.4.1.2 Polymerase Kettenreaktion

Die Polymerase Kettenreaktion (PCR) ist ein Verfahren der in vitro Vervielfältigung spezifischer DNA-Abschnitte. Für die PCR-Reaktion wird ein entsprechendes DNA-Template, eine thermostabile DNA-Polymerase, die vier Basen der DNA (Adenin, Cytosin, Guanin und Thymin) als Desoxyribonukleosidtriphosphate (dNTP) und ein

Primerpaar benötigt. Das Leitenzym der PCR ist eine thermostabile DNA-Polymerase, die komplementär zur Template-DNA die entsprechenden dNTPs miteinander verknüpft und so das gewünschte PCR-Produkt generiert. Der sogenannte Primer, ein sequenzspezifisches kurzes DNA-Fragment, definiert das zu amplifizierende DNA-Fragment, indem es komplementär an die Ziel-DNA bindet und der DNA-Polymerase als Startpunkt dient.

Die Synthese des DNA-Fragmentes erfolgt in drei Teilreaktionen, die sich thermisch und zeitlich unterscheiden:

1. Denaturierung: thermische Aufspaltung des DNA-Doppelstranges
2. Annealing: Anlagerung der Primer an den jeweiligen DNA-Einzelstrang
3. Elongation: Replikation des definierten DNA-Abschnittes durch die DNA-Polymerase

Die Abfolge dieser drei Teilreaktionen entspricht einem Zyklus. Jeder Zyklus führt zur Verdopplung des spezifischen DNA-Fragmentes. Die Anzahl der Zyklen kann je nach gewünschter Ausbeute variabel gewählt werden [73-76].

Durchführung:

Im ersten Schritt der Genotypisierung erfolgte die Amplifikation der PER mittels PCR nach dem folgenden Schema:

Reagenz	Konz.	Vol. (µl)
H ₂ O		15,53
Polymerase-Puffer	5x	5,00
MgCl ₂	50 mM	0,75
dNTPs	10 mM	0,13
Vorwärtsprimer	10 µM	1,25
Rückwärtsprimer	10 µM	1,25
Taq Polymerase	5 U/µl	0,1
Template DNA		1,00

Tabelle 2-6: PCR-Ansatz

Reaktionsschritt	Temperatur	Zeit	
Denaturierung	95°C	4:00 min	
Denaturierung	95°C	0:30 min	
Annealing	58°C – 59,5°C	0:30 min	} 35-39 Zyklen
Elongation	72°C	1:30 min	
Elongation	72°C	8:00 min	
	10°C	∞	

Tabelle 2-7: PCR-Ablauf

Die zur Etablierung der PCR verwendeten Primerpaare sind Tabelle 2-3 zu entnehmen. Verwendet wurde der Primer TSH_SNP_1.

2.4.1.3 Agarosegelelektrophorese

Die Agarosegelelektrophorese dient der Auftrennung und Visualisierung der in der PCR synthetisierten DNA-Fragmente. Sie ist die am weitesten verbreitete Methode zur Analyse von PCR-Produkten [74]. Ein bis zwei Mikroliter des PCR-Produktes wurden mit einem Farbstoff (Dye) gemischt auf ein Agarosegel aufgetragen, welches je nach Agarosekonzentration eine unterschiedlich dichte Struktur bildet und so die Laufgeschwindigkeit und Trennschärfe der DNA-Fragmente bestimmt. Für das vorliegende PCR-Produkt wurde ein 1%iges Agarosegel verwendet. Die Auftrennung der DNA-Fragmente erfolgt mittels elektrischer Spannung. Dabei wandert die negativ geladene DNA im Agarosegel der Größe nach zum positiven Pol der Elektrophoresekammer [77]. Die Laufstrecke des DNA-Fragments ist dabei invers proportional zu dessen Molekulargewicht [77]. Zur Visualisierung des PCR-Produktes wurde das Agarosegel mit 1 µl Ethidiumbromid versetzt, das in die DNA interkaliert und durch Anregung mit UV-Licht fluoresziert. Zur Quantifizierung der Fragmentgröße erfolgte die Auftragung des DNA-Größenmarkers Easy ladder 1 (Bioline), der DNA-Fragmente bekannter Größe enthält.

2.4.1.4 ExoSAP-Aufreinigung

Vor der Sequenzier-PCR erfolgte die Elimination von überschüssigen Substraten und Nebenprodukten der vorausgegangenen PCR, die zu einer Störung der anschließenden Sequenzier-PCR führen könnten. Im ersten Schritt kommt es zu einer Degradierung von verbliebenen einzelsträngigen DNA-Fragmenten und Primern durch die Exonuklease I (Exo) sowie einer Hydrolyse der überschüssigen dNTPs durch die Shrimp-Alkaline-

Phosphatase (SAP). Im zweiten Schritt kommt es zu einer selbstinduzierten Inaktivierung des ExoSAP-Gemisches [76]. Dazu wurden 8 µl des PCR-Produktes mit 0,18 µl Exonuklease I (Exo) (20U/µl) und 0,32 µl (1U/µl) Shrimp-Alkaline-Phosphatase (SAP) unter folgenden Bedingungen behandelt:

Temperatur	Zeit
37° C	30 min
65°C	15 min
4°C	∞

Tabelle 2-8: ExoSAP-Ablauf

2.4.1.5 Sanger-Sequenzierung

Die Sequenzanalyse erfolgte mit der Sequenziermethode nach Sanger unter Verwendung des 3130xL Genetic Analyzer (Thermo Fischer, Waltham, MA, USA), basierend auf einer Fluoreszenzdetektionsmethode. Die für die Sanger-Sequenzierung benötigten Abbruchfragmente wurden in einer Sequenzierreaktion mittels PCR durch Zugabe von fluoreszenzmarkierten Dideoxynukleotiden (ddNTPs) generiert, wobei jedes der vier Dideoxynukleotide (ddATP, ddGTP, ddTTP und ddCTP) mit einem Fluorophor unterschiedlicher Wellenlänge markiert war. Die Auftrennung der Fragmente nach Molekulargewicht erfolgt mittels Kapillarelektrophorese, an deren Ende die optische Detektion der basenspezifischen Fluoreszenzemission und deren Digitalisierung zu Sequenzdiagrammen erfolgt [78, 79].

2.4.1.5.1 Sequenzier-PCR

Nach der Aufreinigung des PCR-Produktes erfolgte die Sequenzierreaktion nach dem Ansatz in Tabelle 2-9 unter Verwendung des PCR-Primerpaares TSH_SNP1 (siehe Tabelle 2-3).

Reagenz	Konzentration	Volumen (µl)
H ₂ O		5,5
Big Dye ® Terminator v1.1, v3.1 Sequencing Buffer	(5x) 5x	1,5
Big Dye ® v3.1 Ready Reaction Mix		0,5
Sequenzierprimer	10 µM	0,5
ExoSAP behandelte Template DNA		2,0

Tabelle 2-9: Ansatz Sequenzier-PCR

Reaktionsschritt	Temperatur	Zeit
Denaturierung	95°C	5:00 min
Denaturierung	95°C	0.30 min
Annealing	55°C	0:15 min
Elongation	60°C	4:00 min
Elongation	60°C	7:00 min
	4°C	∞

Tabelle 2-10: Ablauf Sequenzier-PCR

2.4.1.5.2 Natriumacetat-Ethanol-fällung

Nach der Sequenzierreaktion erfolgte zur Reinigung der Sequenzier-PCR eine Natriumacetat-Ethanol-fällung des PCR-Produktes. Bis zur Analyse am 3130xL Genetic Analyzer (Thermo Fischer, Waltham, MA, USA) wurden die so erzeugten DNA-Pellets bei -20°C aufbewahrt.

Durchführung:

Das PCR-Produkt wurde mit 2 µl Natriumacetat und 20 µl 96%igem Alkohol versetzt und bei 14°C mit 10500 rpm zentrifugiert und anschließend der Überstand aus dem Tube abgenommen. Im nächsten Schritt wurde 200 µl 70%iger Ethanol in den Tube gegeben und wie im ersten Schritt zentrifugiert und der Überstand abgenommen. Anschließend wurde erneut zentrifugiert und das DNA-Pellet zuletzt bei 37°C getrocknet.

2.4.1.6 Auswertung mit der Seq-Pilot-Software

Die Auswertung der durch den 3130xL Genetic Analyzer (Thermo Fischer, Waltham, MA, USA) generierten Sequenzdiagramme erfolgte mit der Software-Sequence-Pilot (JSU medical systems Corp., Boston, MA, USA).

2.4.2 Methylierungsanalyse

Während im ersten Teil der Arbeit die Genotypisierung der potentiellen Enhancer Region mit Hilfe der Sanger-Sequenzierung erfolgte, wurde im zweiten Teil der Arbeit das epigenetische Methylierungsmuster der potentiellen Enhancer Region mittels Pyrosequenzierung analysiert.

2.4.2.1 Bisulfitumwandlung

Die Methylierung von Cytosin-Basen der DNA ist die am besten untersuchte epigenetische Modifizierung der DNA und erfolgt fast ausschließlich an sogenannten CpG-Positionen. Dabei handelt es sich um eine Cytosin-Base, auf die eine Guanin-Base folgt [80]. Die Bisulfitkonvertierung stellt die Grundlage für die Quantifizierung der an Position fünf des Pyruvatrings methylierten Cytosinbasen (5-Methylcytosin) der DNA dar. Erstmals Erwähnung fand die Methode durch Frommer et al. 1992. Das Prinzip dieser Methode beruht auf der chemischen Behandlung der zu analysierenden DNA mit Natriumbisulfit [80].

Die Bisulfitkonversion ist einzelstrangspezifisch [81]. Somit kann nur einzelsträngige DNA konvertiert werden. Deshalb erfolgt zuvor die Denaturierung der doppelsträngigen DNA. Die chemische Konvertierungsreaktion erfolgt anschließend in drei Teilschritten (siehe Abbildung 2-1).

Durch die Bisulfitkonvertierung werden nicht methylierte Cytosin-Basen zu Uracil konvertiert. 5-Methylcytosin Basen dagegen sind von der chemischen Konvertierung durch Bisulfit ausgeschlossen und bleiben als Cytosin-Base in der DNA erhalten [80]. Nach der Bisulfitkonvertierung erfolgt die PCR vermittelte Amplifikation des DNA-Fragments. Dabei kommt es im neu synthetisierten DNA-Strang zum Austausch der durch Bisulfitkonvertierung generierten DNA-Base Uracil zu Thymin und damit letztendlich zu einer Konversion von nicht methylierten Cytosin-Basen zu Thymin. Jedes im neu synthetisierten DNA-Strang noch vorhandene Cytosin ist somit auf eine ursprüngliche Methylierung der Cytosin-Base zurückzuführen [81] (siehe Abbildung 2-1).

Nach der Bisulfitkonversion sind die DNA-Stränge nicht mehr komplementär und die DNA liegt einzelsträngig vor [82]. Daher ist die Amplifikation anschließend nur noch strangspezifisch mit jeweils unterschiedlichen bisulfit-spezifischen PCR-Primern möglich [83].

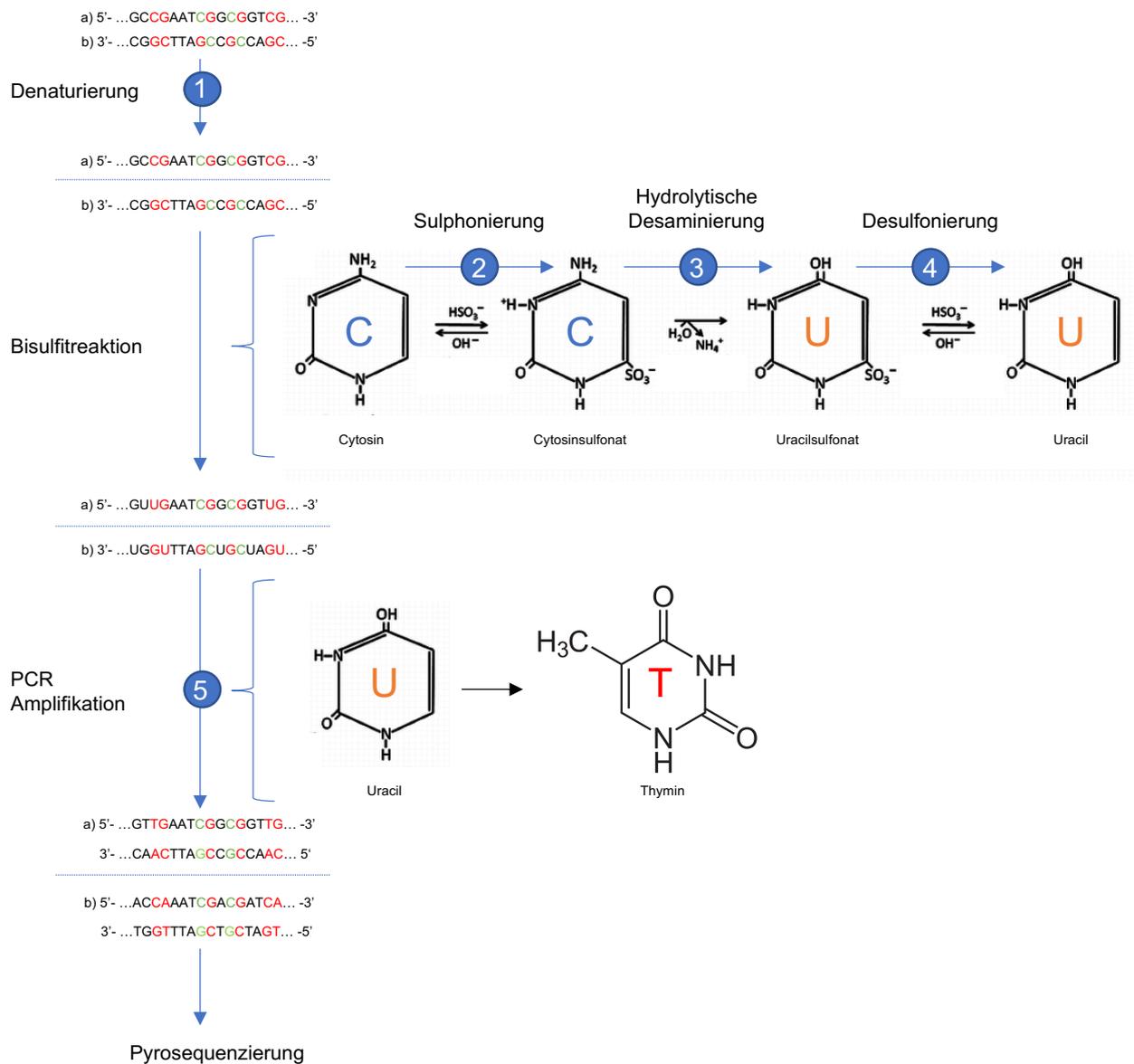


Abbildung 2-1: Ablauf der Bisulfitkonvertierung

Im ersten Schritt erfolgt die Denaturierung des DNA-Doppelstrangs um die DNA der Bisulfitbehandlung zugänglich zu machen. Im zweiten Schritt der Sulfonierung kommt es zur chemischen Reaktion der Cytosin-Base mit Bisulfit. Das Produkt der Reaktion ist Cytosinsulfonat. Im dritten Schritt erfolgt die hydrolytische Desaminierung, sodass Uracilsulfonat entsteht. Im letzten Schritt entsteht schließlich durch Desulfonierung die Base Uracil. Bei der anschließenden Amplifikation (Schritt fünf) wird das Uracil durch die Base Thymin ersetzt. In dem dargestellten Beispiel-DNA-Fragment sind methylierte Cytosinbasen einer CpG-Position grün dargestellt, nicht methylierte CpG-Positionen rot und alle anderen DNA-Basen schwarz. In Anlehnung zur Abbildung [83].

Die Bisulfitkonversion der DNA führt zu einer erheblichen chemischen und physikalischen Veränderung der DNA, vor allem durch den niedrigen pH-Wert und die hohen Temperaturen, die für die Reaktion notwendig sind. Die Folge sind unter anderem DNA-

Strangbrüche, die zu einer verkürzten Fragmentlänge führen, was die anschließende Amplifizierung der DNA stark erschweren kann und zu einer verminderten Produktmenge führt [80, 84]. Auch die spezifische Primerbindung wird durch den Verlust des größten Teils der Cytosin-Basen und damit der Komplexität der DNA erschwert und resultiert in einer erhöhten unspezifischen Primerbindung. Lange Primersequenzen von 18-30 Basenpaaren können die spezifische Primerbindung erhöhen und werden daher für die Bisulfit-PCR empfohlen [81, 85]. Insgesamt ist die PCR gestützte Amplifikation von bisulfitkonvertierter DNA daher deutlich anspruchsvoller als bei Verwendung unbehandelter nativer DNA.

Durchführung:

Die Bisulfitkonvertierung erfolgte mit dem EZ-Methylation-Kit (Zymo Research, Irvine, CA, USA) unter Beachtung des zugehörigen Anwendermanuals. Es wurden jeweils 500 ng der zu untersuchenden DNA umgewandelt. Die Quantifizierung der DNA-Konzentration zur Bestimmung der gewünschten Menge von 500 ng erfolgte mit dem Quantus Fluorometer (Promega, Madison, WI, USA).

Zur Denaturierung wurde die DNA mit 5 µl M-Dilution-Buffer versetzt, auf ein Gesamtvolumen von 50 µl mit Wasser aufgefüllt und durch Pipettieren gemischt. Der Ansatz wurde anschließend für 15 min bei 37°C inkubiert.

Währenddessen erfolgte die Vorbereitung der CT-Conversion-Reagent. Hierzu wurde die CT-Conversion-Reagent mit 750 µl Wasser und 210 µl M-Dilution-Puffer versetzt und für 10 min bei 25°C und 1400 rpm gemischt. Anschließend wurde die zuvor denaturierte DNA mit 100 µl der CT-Reagent gemischt und für die Konvertierungsreaktion bei 50°C für 12-16 Stunden inkubiert. Die Probe wurde dann für mindestens 10 min bis maximal 20 Stunden auf 4°C abgekühlt. Die weiteren Schritte erfolgten in dem Filtereppi Zymo-spin IC colum. Es wurden 400 µl M-Binding-Buffer vorgelegt und die Probe hinzugefügt. Durch 6-maliges Invertieren erfolgte die Mischung beider Komponenten. Anschließend wurde der Zymo-Spin-IC-colum mit 14000 rpm für 30 Sekunden zentrifugiert.

Die Desulphonierung erfolgte mit 200 µl Desulphonation-Buffer für 20 Minuten bei 25°C. Anschließend wurde die Probe mit 14000 rpm zentrifugiert und durch zweimaliges Waschen mit jeweils 200 µl M-Wash-Buffer gereinigt. Zuletzt wurde die DNA aus dem Filter des Zymo-spin- IC -colum in 20-25 µl M-Elution-Buffer gelöst.

2.4.2.2 PCR der Bisulfit konvertierten DNA mit 5`biotinyliertem Rückwärtsprimer

Nach der oben beschriebenen Bisulfitkonvertierung erfolgte die Amplifizierung eines knapp 300 Basenpaar großen DNA-Fragmentes der potentiellen Enhancer Region, in der sich neun CpG-Positionen befinden (siehe Abbildung 2-3). Zur Etablierung der Bisulfit-PCR wurden entsprechende PCR-Primer mit der Pyrosequencing-Assay-Design-Software (Qiagen, Hilden) und der Software Bisulfite-Primer-Seeker (Zymo Research, Irvine, CA, USA) entworfen. Die für die Etablierung der Bisulfit-PCR verwendeten Primer sind der Tabelle 2-3 zu entnehmen. Die Amplifizierung erfolgte schließlich mit dem Primerpaar P4bR/PF.

Zunächst wurde die PCR nach Tabelle 2-11 angesetzt. Der verwendete Rückwärtsprimer P4bR wurde am 5' Ende biotinyliert, um eine spätere Isolierung und Aufreinigung der DNA für die Pyrosequenzierung zu ermöglichen. Die Amplifizierung erfolgte nach den Reaktionsbedingungen in Tabelle 2-12.

Reagenz	Konzentration	Volumen (µl)
Polymerase Puffer	5x	10
MgCl ₂	50 mM	1,5 - 3,0
dNTPs	10 mM	0,25
Vorwärtsprimer	10 µM	1,25
Rückwärtsprimer	10 µM	1,25
Taq Polymerase	5 U/µl	0,25
Bisulfit-behandelte DNA		2,0 - 4,0
H ₂ O		Add 50

Tabelle 2-11: Bisulfit PCR-Ansatz

Reaktionsschritt	Temperatur	Zeit	
1. Denaturierung	95°C	3:00 min	} 39-40 Zyklen
2. Denaturierung	95°C	1:00 min	
3. Annealing	60,9-62,6 °C	1:00 min	
4. Elongation	72°C	2:00 min	
5. Elongation	72°C	8:00 min	
6.	4°C	∞	

Tabelle 2-12: Bisulfit PCR-Ablauf

Der 50 µl PCR-Ansatz wurde für die PCR-Reaktion im Thermocycler auf acht Aliquots zu je 6,2 µl aufgeteilt, um durch das geringere Reaktionsvolumen bessere thermische Bedingungen zu schaffen und damit eine ausreichende Produktausbeute zu erreichen.

2.4.2.3 Agaroseelektrophorese

Zur Qualitäts- und Quantitätskontrolle der Bisulfit-PCR erfolgte eine Gelelektrophorese (siehe 2.4.1.3) mit einem 1- oder 2 %igen Agarosegel, 0,5 µl Ethidiumbromid unter Einsatz von 3-6 µl des PCR-Produktes, welches zur Darstellung im Agarosegel mit 1 µl eines Farbstoffes (Dye) (siehe Tabelle 2-2) vermischt wurde. Zur Größenbestimmung des PCR-Produktes erfolgte die Auftragung des DNA-Größenmarkers Easy Ladder I.

2.4.2.4 Aufreinigung mit magnetischen Beads

Vor der Pyrosequenzierung erfolgte eine Aufreinigung des Bisulfit-PCR-Produktes mit Hilfe des AMPure XP-Kits, um eine Störung der nachfolgenden Pyrosequenzierung durch kurze DNA-Fragmente, in diesem Fall aneinander gelagerte Primer, zu verhindern. Dabei binden magnetische Beads je nach Konzentration DNA-Fragmente unterschiedlicher Größe. Mit Hilfe eines Magnetständers und mehrerer Reinigungsschritte wurden die störenden kleinen DNA-Fragmente aus dem PCR-Produkt entfernt.

Um das knapp 300 Basenpaar große PCR-Produkt aufzureinigen, wurde die 0,75-fache Menge an Bead-Flüssigkeit im Verhältnis zum PCR-Produkt verwendet, zum PCR-Produkt hinzu pipettiert und durch maximal zehnmaliges Pipettieren vorsichtig gemischt. Nach einer Inkubationszeit von 10 Minuten bei Raumtemperatur erfolgte die Trennung des an den magnetischen Beads haftenden PCR-Produkts mit Hilfe eines Magnetständers, der das magnetisch markierte DNA-Fragment an die Tubenwand zieht. Die überschüssige Lösung konnte so verworfen werden und das an der Tubenwand haftende PCR-Produkt zweimalig mit 80%igem Ethanol gewaschen und anschließend bei Raumtemperatur getrocknet werden. Als nächstes wurde das an die magnetischen Beads gebundene Produkt in Eluation-Puffer gelöst, wobei sich die Magnetbeads von der DNA lösen und das PCR-Produkt schließlich erneut mit Hilfe des Magnetständers von den Magnetbeads getrennt in ein neues Tube pipettiert wurde.

2.4.2.5 Pyrosequenzierung

Die Methylierung der DNA ist die am besten untersuchte epigenetische Modifizierung der DNA [86]. Sie erfolgt fast ausschließlich an Cytosin-Basen von CpG-Positionen. Die Grundlage für die Analyse der DNA-Methylierung ist die vorausgegangene Bisulfitkonvertierung der DNA, wodurch alle noch vorhandenen Cytosin-Basen auf eine ursprüngliche Methylierung zurückzuführen sind, da diese nicht durch Bisulfit konvertiert werden [83]. Eine gut etablierte Methode zur quantitativen epigenetischen Methylierungsanalyse von CpG-Positionen ist die sogenannte Pyrosequenzierung. Erstmals beschrieben wurde diese Methode 1998 von Ronaghi et al. [87]. Das grundlegende Prinzip beruht auf der Sequenzierung durch Synthese. Die Nukleotide werden dabei einzeln an den zu synthetisierenden Strang angefügt. Das bei jedem neu eingebauten Nukleotid freiwerdende Pyrophosphat wird dabei über eine Enzymkaskade in ein Lichtsignal transformiert, welches proportional zu der Anzahl der eingebauten Basen ist und als Sequenz-Peak im Pyrogramm dargestellt wird (siehe Abbildung 2-2) [88].

Bisulfit-PCR-Produkt

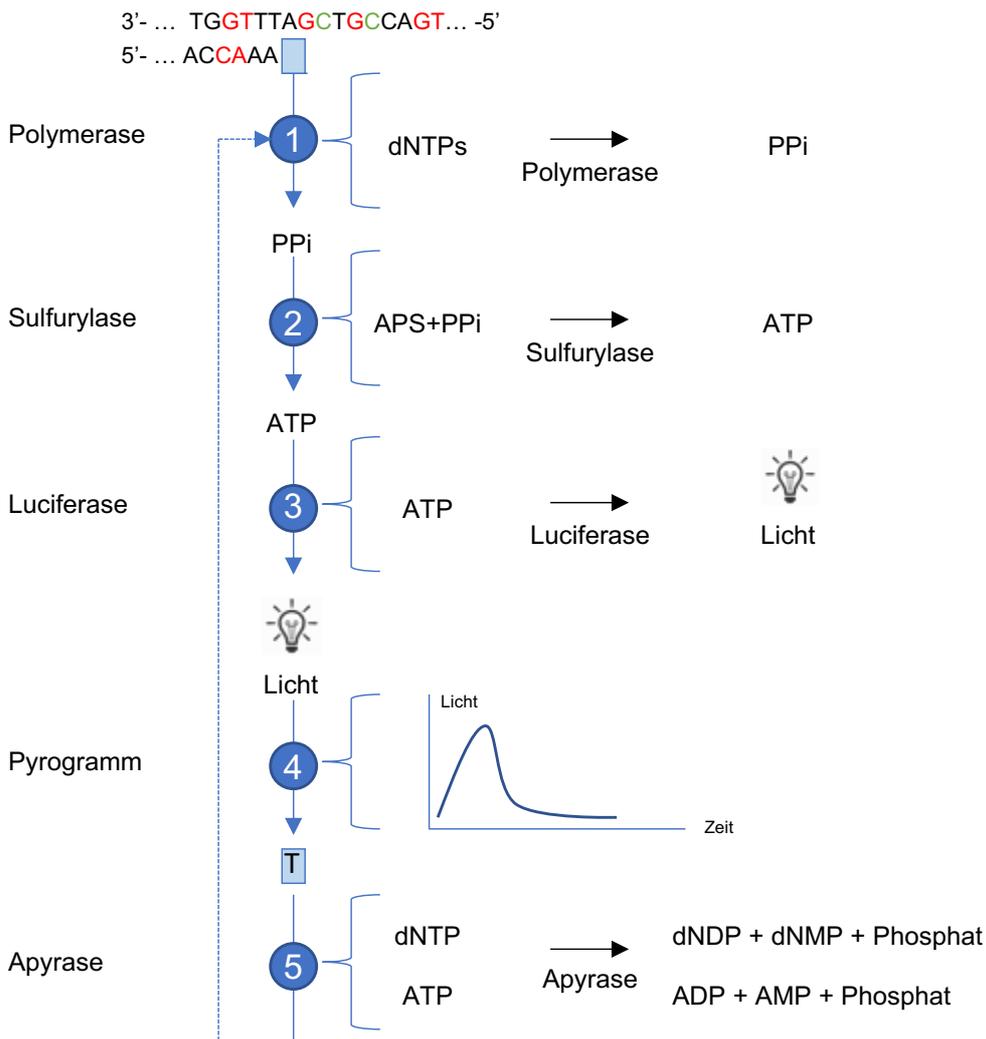


Abbildung 2-2: Pyrosequenzierung

Die Nukleotide (dNTPs) werden nach einer assayspezifisch festgelegten Reihenfolge zu der Probe hinzugegeben. Die im Enzym-Substratgemisch enthaltene Polymerase baut die Nukleotide unter Abspaltung von Pyrophosphat ein. Das entstandene Pyrophosphat (PPi) wird zusammen mit dem Substrat Adenosin 5' Phosphosulfat (APS) durch die Sulfurylase in Adenosintriphosphat (ATP) umgesetzt, welches als Substrat der Luziferase in einem proportionalen Lichtsignal resultiert. Dieses Lichtsignal wird detektiert und als Peak in einem Pyrogramm dargestellt. Die Apyrase katalysiert den Abbau der nicht eingebauten Nukleotide und ATP, sodass das nächste Nukleotid eingebaut werden kann. Weitere Abkürzungen: Desoxyribonukleosidtriphosphat (dNTP); Desoxyribonukleosiddiphosphat (dNDP); Desoxyribonukleosidmonophosphat (dNMP); Adenosindiphosphat (ADP); Adenosinmonophosphat (AMP). Abbildung angelehnt an [89].

Die Durchführung der Pyrosequenzierung erfolgte unter Beachtung des PyroMark Anwender-Manuals. Dabei wurde das aus der Bisulfit-PCR erhaltene biotinylierte PCR-Produkt im ersten Schritt an Steptavidin-Beads gebunden, an einer Waschstation

aufgereinigt und denaturiert, um das für die Sequenzierung benötigte einzelsträngige DNA-Template zu erhalten.

Hierzu wurden im ersten Schritt jeweils 40 µl Binding Buffer (Quiagen) mit 1 µl Sepharose (General Health) und 39 µl des PCR-Produkt mit Wasser auf ein Gesamtvolumen von 80 µl aufgefüllt und bei 1400 rpm und Raumtemperatur für 10 min geschüttelt.

Mit Hilfe einer Vakuumpumpe wurde das an Sepharose gebundene einzelsträngige DNA-Template angesaugt und zur Reinigung und Denaturierung jeweils 5 sek. in 70%igen Ethanol, 5 sek. in Denaturation Buffer und 10 sek. in Washbuffer gehalten. Anschließend erfolgt die Hybridisierung des einzelsträngigen aufgereinigten DNA-Templates mit 25 µl des in Annealing-Buffer gelösten 0,3 µM Sequenzierprimers für 2min bei 80°C und 15 min bei Raumtemperatur. Zusammen mit der assayspezifischen Menge an Nukleotiden und Enzym-Substrat-Mix erfolgte dann die Sequenzierung im Pyrosequenzierer Q24 (Qiagen, Hilden).

In der vorliegenden Arbeit wurden die in Abbildung 2-3 dargestellten neun CpG-Positionen der potentiellen Enhancer Region untersucht. Die neun CpG-Positionen wurden hierbei in zwei Pyrosequenzierassays mit unterschiedlichen Sequenzierprimern geteilt, um durch eine kürzere Sequenzierlänge eine hohe Qualität der Pyrogramme zu erreichen. Die zur Etablierung der Pyrosequenzierung verwendeten Sequenzprimer sind der Tabelle 2-3 zu entnehmen.

```
TTATATGGAGTAAGAGGGTGGTAAA AATGTTTATTTATTTTGTATTAGTGGTTAATTT  
TTT GTTTTTTTT TAGAATTAGTAGTTTT TTTATATT CG GGAATTTA CG TATGTTATTTT  
TATTTTTTTAATA CG AATTAGATAGGTTTATTTGAGTTGGAGAAATAATGGTTTGTAG  
T CG GTTTTATGTTTTAGTG CGAAATAATAATGTAGGTTI TTTGA CG AATTTTGG CG T  
CGCG AGTGGAATTTTGTAGTTTAATATGTTTGT TGAATAGAAGTTGGAAAAAGAAAA  
GAT
```

Abbildung 2-3: Pyrosequenzierter Bereich der potentiellen Enhancer Region

Ausschnitt aus der Bisulfit konvertierten potentiellen Enhancer Region mit den zu untersuchenden CpG-Positionen Bisulfit-PCR-Primerpaar (P4bR/P4F) grün, Pyrosequenzierprimer rosa (S1F) und violett (S4F), pyrosequenzierte CpG-Positionen rot, CpG-Position die den mit dem TSH-Wert assoziierten SNP rs116909374 beinhaltet, türkis.

2.4.2.6 Auswertung mit PyroMark Q24 Software

Die qualitative Kontrolle der Pyrosequenziererergebnisse erfolgte mit Hilfe der dafür vorgesehenen PyroMark Q24 Software.

3 Ergebnisse

Ziel der Arbeit war es, eine mögliche Pathogenese der Schilddrüsendysgenese in Form von genetischen oder epigenetischen Mutationen in den nicht-kodierenden Bereichen des *NKX2-1*-Gens darzustellen. Es sollte eine ca. 800 Basenpaare lange Sequenz mit potenzieller schilddrüsenspezifischer Enhancerfunktion untersucht werden. Im ersten Schritt erfolgte daher die Sanger-Sequenzierung des ca. 800 Basenpaare langen Fragments. Im zweiten Schritt erfolgte dann die epigenetische Methylierungsanalyse der nicht-kodierenden Sequenz downstream des *NKX2-1*-Gens mit potenzieller schilddrüsenspezifischer Enhancerfunktion. Hierbei wurden neun CpG-Positionen eines ca. 300 Basenpaare großen Fragments unter anderen in Schilddrüsendysgenese-Patienten untersucht, das um den SNP rs116909374 gelegen ist.

3.1 Genotypisierung

Analog zu Mutationen in kodierenden Sequenzen kann es durch SNVs zu einem Funktionsverlust von Enhancerelementen und damit auch ihrer entsprechenden Zielgene kommen [34]. Weedon et al. 2014 z.B. konnten zeigen, dass Punktmutationen, die zu einem Funktionsverlust eines pankreasspezifischen Enhancers des *PTFA1*-Gens führen, ursächlich für die isolierte Pankreasagenese sind [21]. Der Begriff „Single nucleotide variant“ (SNV) wird dabei synonym zu dem Begriff Punktmutation verwendet und beschreibt einen einzelnen Basenpaaraustausch in der genomischen DNA, der in Abgrenzung zu SNPs in weniger als einem Prozent der Individuen einer Population auftritt [90]. Nicht nur SNVs, sondern auch größere strukturelle Veränderungen der Basensequenz wie z.B. Deletionen und Insertionen, die zusammenfassend als „Copy number variants“ (CNVs) bezeichnet werden [90], können zu einem Funktionsverlust von Enhancerelementen und so zu einer veränderten Expression der entsprechenden Zielgene führen [91]. Die potentielle schilddrüsenspezifische Enhancerregion sollte daher mittels Sanger-Sequenzierung sowohl auf mögliche SNVs als auch auf kleinere, in der ca. 800 Basenpaare großen PER gelegenen CNVs analysiert werden, die als mögliche Ursache für einen Funktionsverlust des potentiellen schilddrüsenspezifischen Enhancers und damit als Ursache der isolierten Schilddrüsendysgenese in Frage kommen. Ein weiteres Augenmerk sollte auf mögliche Unterschiede in der Allelverteilung des mit dem TSH-Wert assoziierten SNPs rs116909374 und weiteren SNPs in der PER zwischen

gesunden Kontrollen und CH-Patienten gelegt werden, die zu einer Beeinflussung des Enhancers führen könnten.

3.1.1 Genetische Analyse der potentiellen Enhancer Region in drei Kohorten

Als Erstes wurde eine gesunde Kontrollgruppe von 24 Personen analysiert, um einen Überblick über die Verteilung des mit dem Serum TSH-Wert korrelierenden SNPs rs116909374 und möglichen weiteren SNVs oder kleineren CNVs bei gesunden Personen in der PER zu gewinnen. Drei der gesunden Kontrollpersonen waren heterozygote Träger des Minor Allels des SNPs rs116909374. Weitere SNVs oder kleinere CNVs traten in der Kontrollgruppe nicht auf.

Anschließend wurde eine Kohorte mit 14 Hyperthyreotropinämie-Patienten analysiert vor dem Hintergrund der Hypothese, dass alle Patienten die gleiche Ausprägung des mit dem Serum TSH-Wert korrelierenden SNPs rs116909374 besitzen könnten, welche sich von der gesunder Kontrollen unterscheidet und zu einer Funktionsbeeinträchtigung des postulierten schilddrüsenpezifischen Enhancers führt, welche die Ursache dieser milden Form der Schilddrüsenstörung sein könnte. Von den 14 Patienten war ein Patient heterozygoter Träger des Minor Allels des SNPs rs116909374, sodass die oben genannte Hypothese einer milden Schilddrüsenstörung aufgrund einer bestimmten Allel-Ausprägung des SNPs rs116909374 nicht bestätigt werden konnte.

Von den 92 analysierten CH-Patienten waren drei heterozygote Träger des Minor Allels des mit dem TSH-Wert assoziierten SNPs rs116909374. Damit ist ein Unterschied in der Allelverteilung des SNPs rs116909374 im Vergleich zu gesunden Kontrollen unwahrscheinlich. Phänotypisch zeigte sich bei diesen Patienten sowohl eine Athyreose als auch eine Schilddrüsenhypoplasie (siehe Tabelle 3-1). Das Minor Allel des SNPs rs116909374 hat demnach in seiner heterozygoten Form auch keinen Einfluss auf den Schilddrüsenphänotyp innerhalb der CH-Kohorte.

Zusammenfassend traten in allen drei analysierten Kohorten heterozygote Minor Allel-Träger des SNPs rs116909374 auf (siehe Tabelle 3-1). Alle heterozygoten Minor Allel-Träger waren weiblich; allerdings ist die Geschlechterverteilung in der CH-Kohorte nicht zuletzt wegen der deutlichen Mädchenwendigkeit der Erkrankung nicht ausgeglichen [20]. Phänotypisch zeigte sich unter den heterozygoten Minor Allel -Trägern des SNPs rs116909374 ein heterogenes Bild (siehe Tabelle 3-1). Insgesamt gibt es somit keinen

Anhalt dafür, dass sich die Allelverteilung des mit dem Serum TSH-Wert korrelierenden SNPs rs116909374 zwischen den drei Kohorten und damit zwischen unterschiedlichen Schilddrüsenphänotypen unterscheidet. Aufgrund der geringen Kohortengröße ist die Aussagekraft der Ergebnisse zudem deutlich eingeschränkt und die Bestimmung der Minor Allelfrequenz nicht möglich.

Übersicht SNP und SNV Träger

Kohorte	SNV	Allele	Allel Pat.	MAF	MAF Quelle	CpG	Phenotyp	G
Mit dem TSH-Wert korrelierender SNP:								
Kontrollen	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	gesund	w
Kontrollen	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	gesund	w
Kontrollen	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	gesund	w
HTT	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	milde TSH-Erhöhung	m
CH	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	Athyreose	w
CH	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	Hypoplasie	w
CH	rs 116909374	C/T	CT	0,0114/57	1000 Genome	2	Athyreose	w
Weitere								
CH	rs 372840928	C/T	CT	0,0002/1	1000 Genome	4	Hypoplasie	m
CH	rs 77916927	A/G	AG	0,0409/205	1000 Genome	0	Hemithyroidea	w
Neue SNVs:								
CH	dbSNP(147): neu dbSNP(153): rs993803802	G	GA	n.a. 0,000032/4	n.a. TOPMED	1	Athyreose	w
CH	neu	C	CT	n.a.	n.a.	7	n.a.	w

Tabelle 3-1: Phänotyp SNP/ SNV Korrelation

Dargestellt sind die Individuen der Kohorten, die Träger eines SNPs, bekannten SNVs oder neuen SNVs in der analysierten PER sind. Dabei ist zu beachten, dass einer der neuen SNVs in der aktuellsten Version der dbSNP Datenbank (153) inzwischen als SNP rs993803802 hinterlegt ist. Weiter ist der Schilddrüsenphänotyp der SNP/SNV-Träger dargestellt. In Grau hinterlegt sind die Varianten, die in einer Person mit CH zusammen auftraten.

SNV: Single nucleotid variants, Allel Pat.: Allel des Patienten, Hpt. Allel: Hauptallel, MAF: Minor Allel Frequenz, CpG: CpG-Position des SNPs/SNVs, G: Geschlecht, n.a.: nicht angegeben.

3.1.2 Genetische Analyse von Schilddrüsen-DNA gesunder Personen

Neben der DNA aus peripheren Blutzellen (PBL) wurde die DNA aus Schilddrüsengewebe von zwei verschiedenen schilddrüsengesunden Personen analysiert. Es zeigten sich in beiden Proben weder eine Minor Allelausprägung für die bekannten SNPs und SNVs noch neue SNVs oder CNVs. Die Genotypisierung der DNA aus Schilddrüsengewebe ist für die Interpretation der im zweiten Teil folgenden Methylierungsanalyse relevant.

3.1.3 SNPs in der potentiellen Enhancer Region

Innerhalb der ca. 800 Basenpaare großen potentiellen Enhancer Region (PER) wurden zum Zeitpunkt der Analyse in der UCSC Genome Browser Datenbank NCBI GRCh37/hg19 dbSNP(147) im menschlichen Genom 38 SNPs beschrieben; davon waren definitionsgemäß lediglich zwei SNPs mit einer MAF von mehr als einem Prozent. Für viele der übrigen SNPs bzw. SNVs war zu diesem Zeitpunkt keine MAF bekannt.

Eine aktuelle Reanalyse der PER des *NKX2-1*-Gens in der Version der UCSC Genome Browser Datenbank dbSNP(153) ergab mit 188 SNPs und SNVs im menschlichen Genom fast fünfmal so viele SNVs wie bei der ersten Analyse. Bei allen hinzugekommenen Varianten handelt es sich ausschließlich um SNVs. Auch einer der neu entdeckten SNVs dieser Arbeit ist, wie bereits oben beschrieben, in der dbSNP(153) Datenbank inzwischen hinterlegt (siehe Tabelle 3-1). Dieser Befund zeigt die allgemein große Dynamik des Wissenszuwachses in der Genomforschung.

Einzig in der CH-Kohorte traten neben dem mit dem Serum TSH-Wert korrelierenden SNP rs116909374 weitere SNPs und SNVs heterozygot auf. Darunter waren auch zwei neue, bis dahin nicht in der dbSNP(147) Datenbank des USCS Genome Browsers beschriebene heterozygote SNVs, von denen einer der beiden in der dbSNP(153) Datenbank inzwischen als SNV hinterlegt ist. Dabei trat dieser SNV in einem Patienten auf, der gleichzeitig auch heterozygoter Träger des Minor Allels des mit dem Serum TSH-Wertes korrelierenden SNPs rs116909374 war. Phänotypisch hatte dieser Patient eine Athyreose, also die maximale Ausprägung der CH. Allerdings wurde dieser Phänotyp auch bei einem weiteren CH-Patienten gefunden, der außer dem Minor Allel des SNPs rs116909374 keine weiteren SNPs oder SNVs trug. Auch die weiteren in der CH-Kohorte gefundenen heterozygoten SNP- und SNV-Träger zeigen keinen homogenen Schilddrüsenphänotyp (siehe Tabelle 3-1).

3.1.4 SNPs als Ursache einer verminderten Transkriptionsfaktorbindung

Betrachtet man die nur in der CH-Kohorte gefundenen SNPs und hierbei vor allem die zum Analysezeitpunkt neu entdeckten SNVs in dieser Kohorte (siehe Tabelle 3-1), so stellt sich die Frage, ob diese SNPs und SNVs Einfluss auf eine mögliche Enhancerfunktion der Zielregion nehmen könnten. Dabei könnten die SNPs und SNVs eine potentielle Bindung von Transkriptionsfaktoren an die Zielregion negativ beeinflussen und somit deren postulierte schilddrüsenspezifische Enhancerfunktion stören.

Vor allem die mit aktiver Enhancerfunktion assoziierte Histonmodifikation H3K4me1 ist in dem Bereich, in dem die SNPs und SNVs in Patienten mit isolierter Schilddrüsendysgenese in enger Nachbarschaft liegen, besonders ausgeprägt vorhanden (siehe Abbildung 3-1). Dies könnte zu einer funktionell relevanten Veränderung der Chromatinstruktur führen.

DNase I Hypersensitivity Cluster in der potentiellen Enhancer Region deuten auf eine „offene“, als Euchromatin bezeichnete, Nukleosomen geminderte Chromatinstruktur hin, die typisch für aktive Enhancer ist und daher für die Identifikation von regulatorischen Funktionsbereichen der DNA herangezogen wird [38, 48]. Ein DNase I Hypersensitivity Cluster liegt genau in dem Bereich der PER, in dem sich auch drei der vier SNPs und SNVs befinden, die ausschließlich in Patienten mit isolierter Schilddrüsendysgenese zu finden waren (siehe Abbildung 3-1).

Um der oben genannten Frage nachzugehen, wurde die PER mit Hilfe des UCSC Genome Browsers auf die Bindung von Transkriptionsfaktoren untersucht. Insgesamt binden fünf Transkriptionsfaktoren an die Zielregion. Alle Transkriptionsfaktoren binden im Bereich der insgesamt zwei DNase I Hypersensitivity Cluster. Die Euchromatinstruktur, die durch diese Cluster detektiert wird, ist zelltypspezifisch und unterliegt dynamischen Veränderungen während der Entwicklung sowie durch externe Stimuli [92]. FOXA1 gehört zu den bindenden Transkriptionsfaktoren und ist als sogenannter Pionierfaktor durch Bindung an Nukleosomen und Rekrutierung von weiteren Faktoren maßgeblich an der dynamischen „Öffnung“ des Heterochromatins für weitere Transkriptionsfaktoren beteiligt, um letztendlich die Transkription des entsprechenden DNA-Segmentes zu ermöglichen [48, 64, 93]. FOXA1 ist so an der Regulation einer Vielzahl von Geweben während der embryonalen Entwicklung beteiligt. In der Schilddrüse konnte eine FOXA1 Expression in C-Zellen sowie im medullären Schilddrüsenkarzinom nachgewiesen werden. In Follikelzellen zeigt sich hingegen keine FOXA1 Expression. Die Expressionsveränderungen von FOXA1 in nicht neoplastischen Schilddrüsenerkrankungen ist allerdings insgesamt bisher nicht ausreichend untersucht [94].

ESR1 und NR3C1 gehören zu den nukleären Rezeptoren, Transkriptionsfaktoren, die eine wichtige Rolle in der Entwicklung, Homöostase und dem Metabolismus von Organismen spielen [95]. NANOG und POU5F1 sind Transkriptionsfaktoren, die in

embryonalen Stammzellen gefunden wurden und eine Schlüsselrolle in der für sie charakteristischen Selbsterneuerung und Pluripotenz spielen [96]. Eine Bindung von Transkriptionsfaktoren, die an der Schilddrüsenhormonproduktion beteiligt sind, sowie die an der Enhanceraktivierung beteiligte Acetyltransferase P300 [38] konnte allerdings nicht nachgewiesen werden.

In dem Bereich, an dem die meisten Transkriptionsfaktoren potenziell binden, liegen im Abstand von nur wenigen Basenpaaren die zwei in der Zielregion vorkommenden SNPs, darunter der mit dem Serum TSH-Wert korrelierende SNP rs116909374 sowie der zum Analysezeitpunkt neu entdeckte SNV, der inzwischen als SNP rs993803802 hinterlegt ist. Dieser SNP trat in einem Patienten mit Schilddrüsendysgenese auf, der gleichzeitig auch heterozygoter Träger des SNPs rs116909374 ist und somit zwei Varianten im Bereich der potentiellen Transkriptionsfaktorbindung trägt. Eine Beeinflussung der Transkriptionsfaktorbindung scheint daher möglich. Die beiden Varianten dieses Patienten liegen in nächster Nähe zueinander und beide auf CpG-Positionen. Insgesamt vier der in der PER gefundenen SNPs liegen auf CpG-Positionen, die durch das Minor Allel des SNPs verändert werden (siehe Abbildung 3-2).

Die Zerstörung der CpG-Positionen führt zu einer veränderten Methylierung der betroffenen CpG-Positionen. Eine Veränderung der Methylierung kann ebenfalls Einfluss auf regulative Funktionen nehmen, sodass im zweiten Teil dieser Arbeit eine Methylierungsanalyse der PER erfolgte.

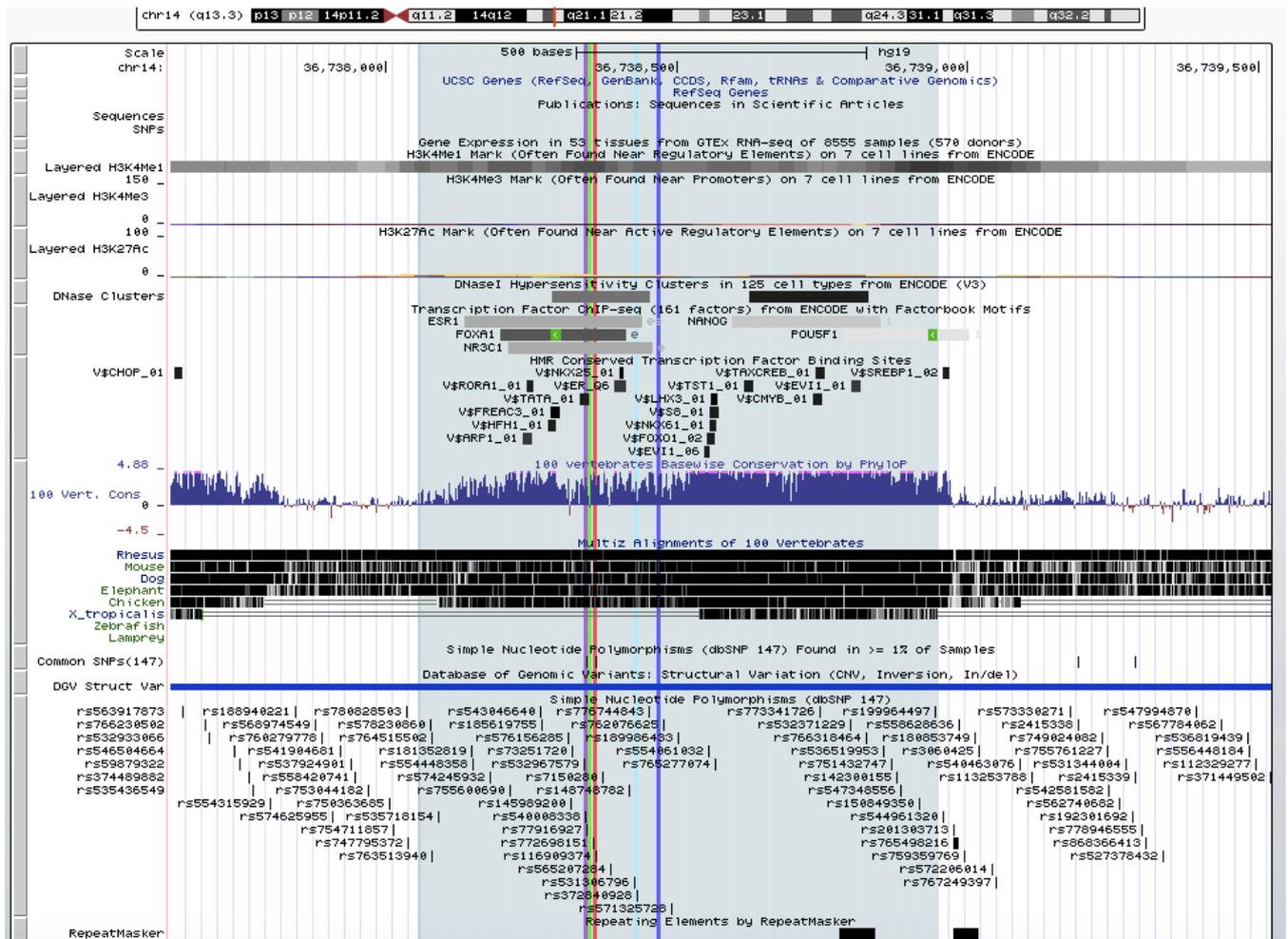


Abbildung 3-1: Auszug aus dem Genombrowser der PER

Auszug aus dem UCSC Genome Browser (NCBI GRCh37/hg19). Der grau hinterlegte Bereich entspricht der PER (chr14:36,738,037-36,738,969). Die senkrechten farbigen Balken markieren die in der CH-Kohorte gefundenen SNPs und SNVs, wobei der rote Balken den mit dem Serum TSH-Wert korrelierenden SNP rs116909374 markiert. Weitergehende Information zu den gefundenen SNPs und SNVs sind in gleicher farblicher Markierung der Abbildung 3-2 zu entnehmen. Weiter sind die an der PER bindenden Transkriptionsfaktoren dargestellt sowie die Transkriptionsfaktorbindungsstellen.

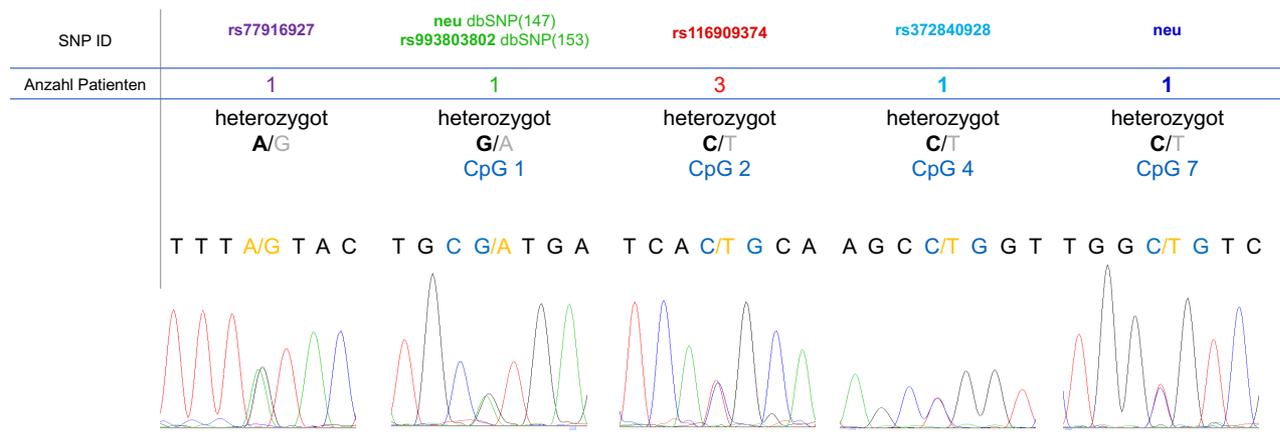


Abbildung 3-2: SNPs und SNVs in der CH-Kohorte

Dargestellt sind die in der CH-Kohorte gefundenen SNPs und neuen SNVs in der Reihenfolge ihres Auftretens in der PER. Weiter ist die Häufigkeit dieser in der CH-Kohorte dargestellt sowie die Lagebeziehung der SNPs und SNVs zu CpG-Positionen der PER.

Zusammenfassend zeigte sich in der genetischen Analyse der PER keine einheitliche SNP-Allelverteilung oder Mutation, die in allen Patienten mit Schilddrüsendysgenese zu finden ist und als alleiniger ursächlicher Faktor für die isolierte Schilddrüsendysgenese in Frage kommt. Allerdings konnten neben dem mit dem TSH-Wert korrelierenden SNP rs116909374 nur in Patienten mit isolierter Schilddrüsendysgenese weitere seltene SNPs und SNVs sowie ein neuer SNV gefunden werden, deren Einfluss z.B. auf die Bindung von Transkriptionsfaktoren nicht auszuschließen ist (siehe Abbildung 3-1 und Abbildung 3-2). Zu beachten ist, dass einer der beiden zum Zeitpunkt der Analyse in der dbSNP(147) Datenbank nicht hinterlegte neue SNV in der aktuellsten Version der Datenbank dbSNP(153) bereits als bekannter SNP rs993803802 geführt wird (siehe Abbildung 3-2).

Das alleinige Auftreten von, neben dem Minor Allel des mit dem Serum TSH-Wert korrelierenden SNP rs116909374, weiteren SNPs und SNVs in der CH-Kohorte könnte durch die ungleiche Kohortengröße bedingt sein, da die CH-Kohorte mit Abstand die größte Kohorte darstellt. Nicht nur die ungleiche Kohortengröße, sondern auch die insgesamt kleine Fallzahl ist bei der Betrachtung der Ergebnisse zu beachten. Eine Aussage bezüglich kohortenspezifischer Unterschiede in der SNP- und SNV-Allelverteilung ist daher nur sehr eingeschränkt möglich.

3.2 Methylierungsanalyse

Epigenetische Faktoren wie die DNA-Methylierung werden als molekulare Ursache der isolierten Schilddrüsendysgenese diskutiert [10]. Eine Störung des postulierten schilddrüsen-spezifischen Enhancers durch Methylierungsveränderungen könnte eine mögliche molekulare Ursache der isolierten Schilddrüsendysgenese sein.

Daher erfolgt im zweiten Teil der Arbeit die Analyse des Methylierungsmusters in der potentiellen Enhancer Region mit Hilfe der Pyrosequenzierung.

3.2.1 Pyrosequenzierung der potentiellen Enhancer Region

Das Grundprinzip der Pyrosequenzierung ist die Sequenzierung durch Synthese. Ein Nachteil dieses Verfahrens ist der relativ kurze DNA-Sequenzbereich von ca. 100 Basenpaaren, der mit diesem Verfahren in einer Messung analysiert werden kann. Limitationsfaktoren sind unter anderem die Hemmung und Verdünnung der beteiligten Enzyme durch Akkumulation von Nebenprodukten [97].

Nach Betrachtung der Verteilung der CpG-Positionen in der PER wurde die Untersuchung auf den mit dem Serum TSH-Wert korrelierenden SNP rs116909374 und acht weitere in der unmittelbaren Nachbarschaft gelegene CpG-Positionen beschränkt (siehe Abbildung 3-3).

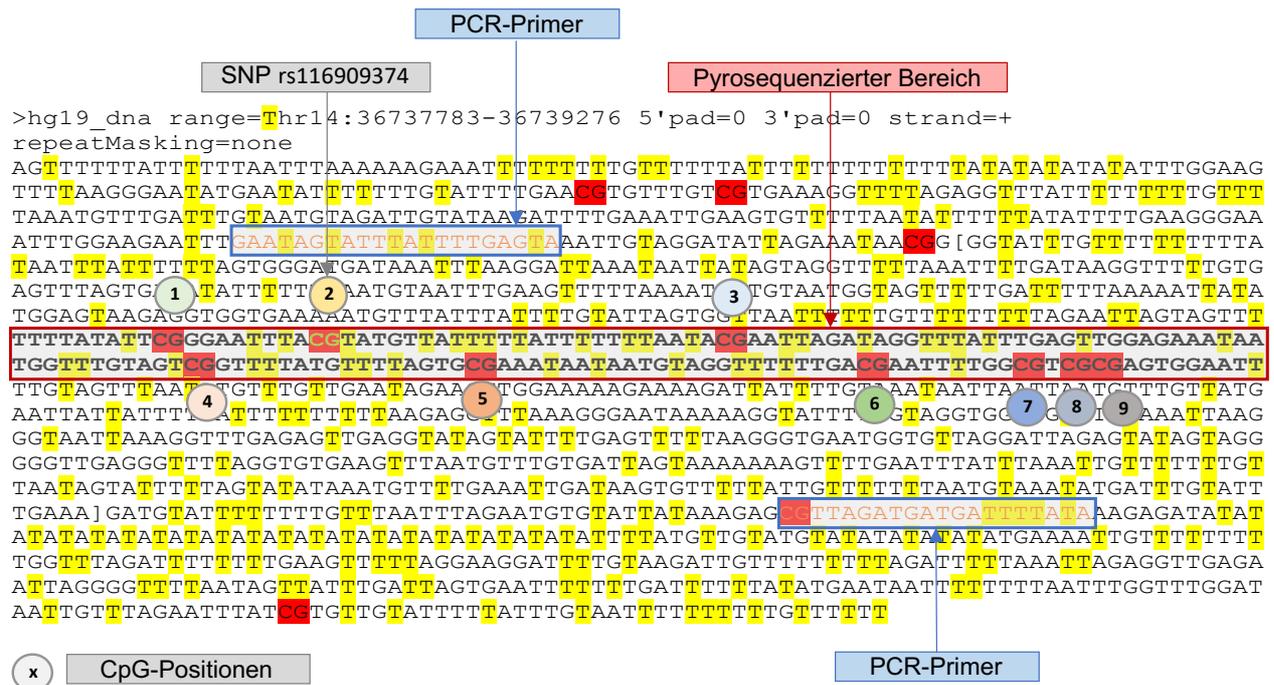


Abbildung 3-3: Ausschnitt aus der DNA-Sequenz der PER nach Bisulfitkonversion.

Gelb markiert sind alle Guanin Basen, die nicht Teil einer CpG-Position sind und daher durch Bisulfitbehandlung zu Thymin konvertieren. Rot markiert sind die CpG-Positionen. Die orangenen Basen markieren die PCR-Primer und zeigen somit die nach Sanger sequenzierte PER (chr14:36,738,037-36,738,969) des *NKX2-1*-Gens. Die grünen Buchstaben markieren den mit dem Serum TSH-Wert korrelierenden SNP rs116909374, die Zahlen die neun analysierten CpG-Positionen der PER. Die rote Umrahmung zeigt den Teil der PER, der pyrosequenziert wurde. Quelle: USCS Genome Browser chr14:36737783-36739276 (NCBI GRCh37/hg19).

In Abbildung 3-4 ist ein Beispielpyrogramm der PER dargestellt. Aufgrund der Länge des in Abbildung 3-3 dargestellten pyrosequenzierten Bereiches von ca. 300 Basenpaaren erfolgte die Pyrosequenzierung jeweils in zwei Teilabschnitten. So wurden in einem Pyrosequenzierungslauf die CpG-Positionen eins bis fünf und anschließend in einem weiteren Pyrosequenzierungslauf die CpG-Positionen sechs bis neun analysiert (siehe Abbildung 3-3).

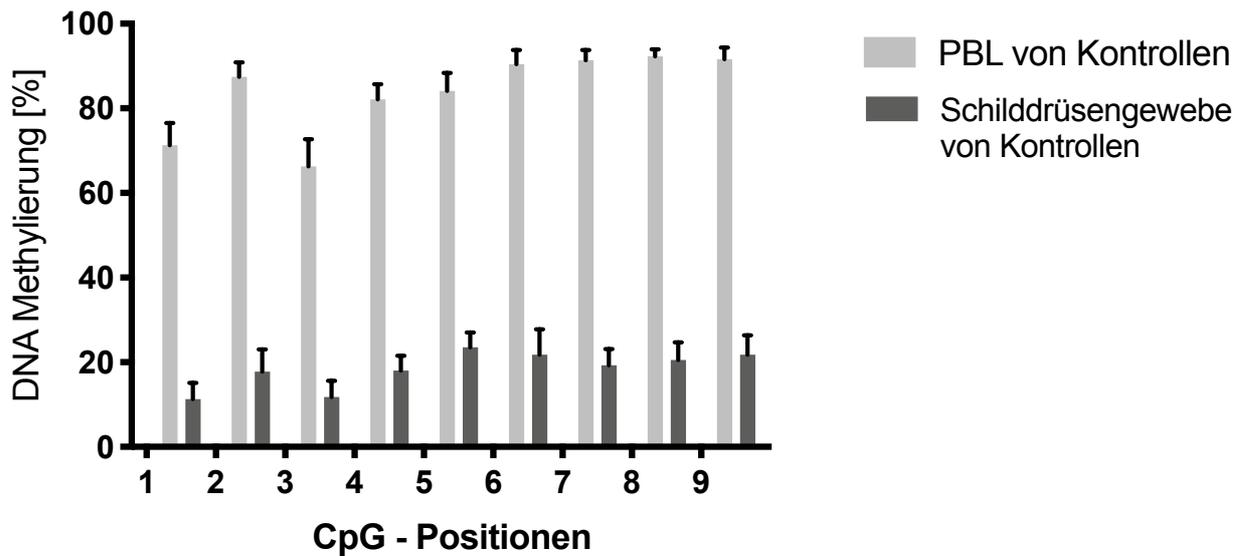


Abbildung 3-5: Gewebespezifisches Methylierungsmuster der PER

Dargestellt ist die gewebespezifische Methylierungsintensität der neun CpG-Positionen in peripheren Blutzellen (PBL) und Schilddrüsengewebe von schilddrüsengesunden Kontrollen. Analysiert wurde das Schilddrüsengewebe von zwei Personen sowie Blut von 30 gesunden Kontrollen. Dargestellt ist der Mittelwert der Methylierung in Prozent mit Standardabweichung.

Die gewebespezifische Hypomethylierung des Schilddrüsengewebes ist ein Hinweis für gewebespezifische regulative Prozesse in der Zielregion [47] und stützt somit die postulierte Annahme einer schilddrüsenspezifischen Enhancerfunktion. Die Analyse von kohortenspezifischen Methylierungsunterschieden müsste daher auch im funktionell relevanten Schilddrüsengewebe erfolgen. Aufgrund der fehlenden Verfügbarkeit von entsprechendem Gewebe konnte eine kohortenspezifische Methylierungsanalyse im Schilddrüsengewebe nicht erfolgen. Problematisch ist zudem, dass Patienten mit einer Athyreose aufgrund des Fehlens von Schilddrüsengewebe von der Analyse ausgeschlossen wären. Daher erfolgte unter der Annahme möglicher gewebeunabhängiger Methylierungsunterschiede zwischen den Kohorten und unter Berücksichtigung der Limitationen in der Aussagekraft der Ergebnisse die kohortenspezifische Methylierungsanalyse aus peripheren Blutzellen.

3.2.3 Kohortenvergleich der Methylierung in der PER

Zur Beantwortung der Frage, ob sich die DNA-Methylierung zwischen gesunden Kontrollen und Patienten mit isolierter Schilddrüsendysgenese in der potentiellen Enhancer Region, aber insbesondere auch vom mit dem Serum TSH-Wert korrelierenden

SNP rs116909374 unterscheidet, erfolgte die qualitative und quantitative Bestimmung der Methylierung an den neun CpG-Positionen mit Hilfe der Pyrosequenzierung. Da das Minor Allel des SNPs rs116909374 mit einem erniedrigten TSH-Wert assoziiert ist [41, 42], wurde auch eine Kohorte mit Hyperthyreotropinämie-Patienten in die Analyse eingeschlossen, um einen möglichen Zusammenhang des SNPs mit einer milden Störung der Schilddrüsenfunktion, die im Falle der Hyperthyreotropinämie durch einen erhöhten TSH-Wert definiert ist, zu analysieren. Insgesamt wurden somit vier verschiedene Kohorten - eine gesunde Kontrollgruppe (K), Patienten mit isolierter Schilddrüsendysgenese (CH), sechs für die isolierte Schilddrüsendysgenese diskordante monozygote Zwillingspaare (MZZP) sowie Patienten mit Hyperthyreotropinämie (HTT) - bezüglich der Methylierung an den neun festgelegten CpG-Positionen analysiert und die Ergebnisse auf ihre Belastbarkeit überprüft (siehe 3.2.6).

Bei der absoluten Anzahl der Patienten pro Kohorte ist zu beachten, dass alle SNP- und SNV-Träger aus der Methylierungsanalyse ausgeschlossen wurden, da der Methylierungsunterschied dieser Personen auf eine genetisch bedingte Zerstörung einer CpG-Position zurückzuführen ist.

3.2.3.1 Mittlere Methylierung der PER

Die Betrachtung des Mittelwertes der Methylierung über alle neun CpG-Positionen zeigt eine statistisch signifikant höhere mittlere Methylierung der CH-Kohorte im Vergleich zu den beiden anderen Kohorten (siehe Abbildung 3-6). Auch in der multiplen Regressionanalyse ist dieser statistisch signifikante Unterschied nachweisbar (siehe Tabelle 3-7 im Abschnitt 3.2.6.5).

Die Kontrollen und die HTT-Patienten zeigen keinen statistisch signifikanten Methylierungsunterschied untereinander.

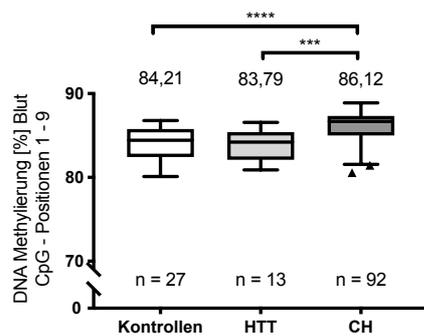


Abbildung 3-6: Kohortenvergleich der mittleren Methylierung der PER

Dargestellt ist die mittlere Methylierung in Prozent über alle neun CpG-Positionen hinweg im Kohortenvergleich. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt.

P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***) < 0,0001(****).

3.2.3.2 Positionsspezifische Methylierungsunterschiede

Die gefundenen kohortenspezifischen Methylierungsunterschiede könnten auf einzelne der neun CpG-Positionen zurückzuführen sein und deren Methylierungsverteilung mit dem Schilddrüsenphänotyp korrelieren. Daher wurde jede einzelne der neun CpG-Positionen auf kohortenspezifische Methylierungsunterschiede analysiert.

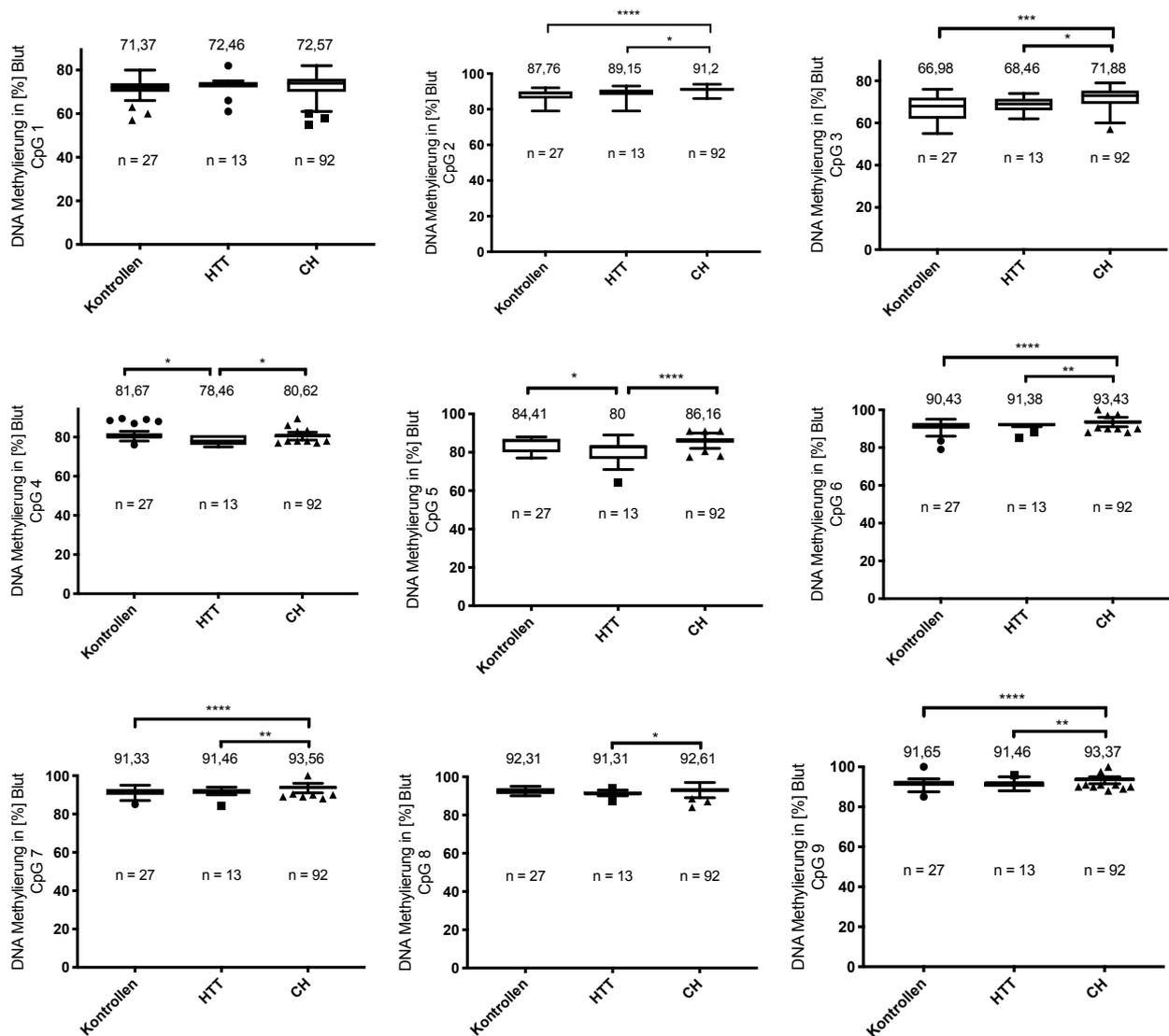


Abbildung 3-7: Positionsspezifischen Methylierungsunterschiede

Dargestellt sind die positionsspezifischen Methylierungsunterschiede der Kohorten an jeder einzelnen der neun CpG-Positionen. Die Zahl über den jeweiligen Graphen gibt den Mittelwert der Methylierung der Kohorte an. Die Zahl n beschreibt die Personenzahl der Kohorte. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***); < 0,0001(****).

An acht der neun CpG-Positionen war die CH-Kohorte konsistent mit den Ergebnissen der mittleren Methylierung über alle CpG-Positionen am höchsten methyliert. Dabei war der Methylierungsunterschied zwischen den Kontrollen und der CH-Kohorte an fünf dieser acht Positionen statistisch signifikant. Einzig an der CpG-Position vier waren die Kontrollen höher methyliert als die CH-Kohorte, wobei dieser Methylierungsunterschied nicht statistisch signifikant war (siehe Abbildung 3-7).

An allen neun CpG-Positionen zeigte sich auch eine Hypermethylierung der CH-Kohorte gegenüber der HTT-Kohorte, die außer an der CpG-Position eins statistisch signifikant waren (siehe Abbildung 3-7).

Zwischen den Kontrollen und der HTT-Kohorte zeigte sich lediglich an den CpG-Positionen vier und fünf ein statistisch signifikanter Methylierungsunterschied mit einer höheren Methylierung der gesunden Kontrollen. Insgesamt war die HTT-Kohorte aber an fünf der neun CpG-Positionen statistisch nicht signifikant höher methyliert als die gesunden Kontrollen (siehe Abbildung 3-7).

Zusammenfassend ist die CH-Kohorte, außer an der CpG-Position vier, somit höher methyliert als die beiden anderen Kohorten. An fünf der neun CpG-Positionen ist dieser Unterschied bezüglich der Kontrollen statistisch signifikant. Bei Betrachtung der HTT-Kohorte ist der Methylierungsunterschied sogar an acht der neun CpG-Positionen statistisch signifikant gegenüber der CH-Kohorte.

Wird der absolute Methylierungsunterschied zwischen den Kohorten betrachtet (siehe Tabelle 3-2), so ist an sechs der neun CpG-Positionen der Methylierungsunterschied zwischen der CH-Kohorte und den Kontrollen größer als zwischen der HTT-Kohorte und den Kontrollen. An fünf dieser CpG-Positionen ist die HTT-Kohorte, wenn auch nicht statistisch signifikant, höher methyliert als die Kontrollen.

Die Methylierung ist somit an fünf der neun CpG-Positionen entsprechend des Schilddrüsenphänotyps in der CH-Kohorte am höchsten, gefolgt von der HTT-Kohorte, und am niedrigsten bei den gesunden Kontrollen, auch wenn der Effekt nicht immer statistisch signifikant ist (siehe Tabelle 3-2).

CpG-Position		Kontrollen vs. CH	HTT vs. CH	Kontrollen vs. HTT
1	Meth. Diff.	-0,75	-0,11	-1,09
	P-Wert	ns >0,1234	ns >0,1234	ns >0,1234
2	Meth. Diff.	-3,44	-2,05	-1,39
	P-Wert	**** <0,0001	* 0,0338	ns >0,1234
3	Meth. Diff.	-4,9	-3,42	-1,48
	P-Wert	***0,0002	*0,0237	ns >0,1234
4	Meth. Diff.	1,05	-2,16	3,21
	P-Wert	ns >0,1234	* 0,0125	* 0,0161
5	Meth. Diff.	-1,75	-6,16	4,41
	P-Wert	ns >0,1234	**** <0,0001	* 0,0242
6	Meth. Diff.	-3	-2,05	-0,95
	P-Wert	**** <0,0001	** 0,0022	ns >0,1234
7	Meth. Diff.	-2,23	-2,1	-0,13
	P-Wert	**** <0,0001	** 0,002	ns >0,1234
8	Meth. Diff.	-1,3	-1,3	0
	P-Wert	ns >0,1234	* 0,0224	ns >0,1234
9	Meth. Diff.	-1,81	-1,91	0,1
	P-Wert	**** <0,0001	** 0,002	ns >0,1234
Gesamt	Meth. Diff.	-1,91	-2,33	0,42
	P-Wert	**** <0,0001	***0,0002	ns >0,1234

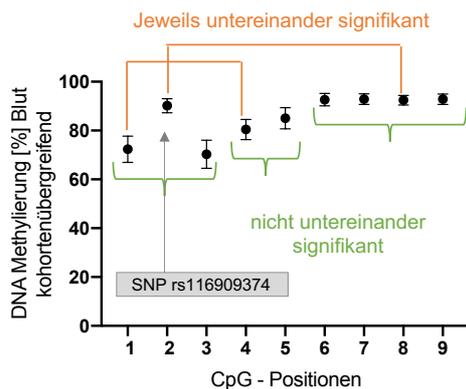
Tabelle 3-2: Absolute Methylierungsdifferenz und Signifikanz

Dargestellt ist die Methylierungsdifferenz der Kohorten an den jeweiligen CpG-Positionen. Zur Bestimmung der Methylierungsdifferenz wurde kohortenspezifisch der Mittelwert der Methylierung für die jeweilige CpG-Position verwendet. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) wurde der nichtparametrische beidseitige Mann-Whitney-Test durchgeführt.

P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***); < 0,0001(****)

3.2.3.3 CpG-positionsspezifische Methylierung

Die Analyse der CpG-positionsspezifischen Methylierung in Abbildung 3-8 zeigt, dass die CpG-Positionen in der unmittelbaren Nachbarschaft des mit dem Serum TSH-Wert korrelierendem SNPs rs116909374 (CpG-Position 2) kohortenübergreifend niedriger methyliert sind als weiter entfernt liegende CpG-Positionen. Die weit entfernt liegenden CpG-Positionen sechs bis neun sind am höchsten methyliert, zeigen eine geringe Varianz und unterscheiden sich im Gegensatz zu den anderen CpG-Positionen nicht signifikant voneinander.



CpG-Positionen	P-Wert
1 vs. 3, 4 vs. 5, 6 vs. 7 vs. 8 vs. 9	ns
1 vs. 4, 2 vs. 8	***
Alle anderen Positionen untereinander	****
	< 0,0001

Abbildung 3-8: Kohortenübergreifende positionsspezifische Methylierung

Dargestellt ist der Mittelwert der Methylierung in Prozent mit Standardabweichung der CpG-Positionen aller analysierten Patienten unabhängig von der Kohorte. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***); < 0,0001(****).

Um die signifikanten positionsspezifischen Methylierungsunterschiede auch kohorten- und gewebespezifisch zu betrachten, erfolgte eine entsprechende Aufstellung der CpG-Positionen nach aufsteigender Methylierung (Abbildung 3-9). Abgesehen von kleineren Abweichungen zeigt das Pfeildiagramm eine relative Konstanz der positionsspezifischen Methylierung über Kohorten und Gewebetypen hinweg.

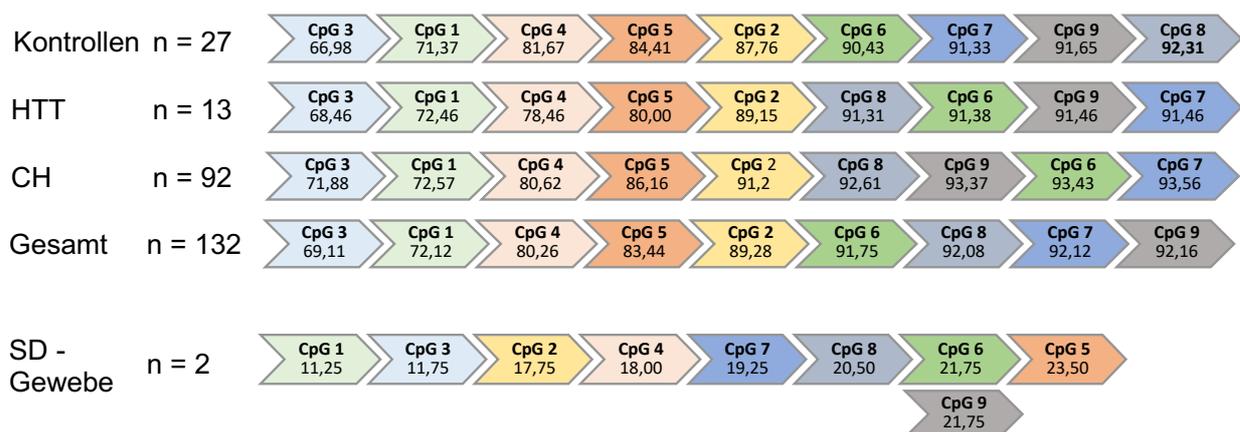


Abbildung 3-9: CpG-Positionen nach aufsteigender Methylierung

Dargestellt ist der Mittelwert der Methylierung in aufsteigender Reihenfolge an den neun CpG-Positionen: kohortenspezifisch, kohortenübergreifend sowie im Schilddrüsengewebe.

HTT: Patienten mit Hyperthyreotropinämie; CH: Patienten Schilddrüsendysgenesie; SD-Gewebe: Schilddrüsengewebe gesunder Kontrollen.

3.2.4 Kohortenvergleich der Methylierung bei monozygoten Zwillingen

Diskordante monozygote Zwillinge sind eine wertvolle Ressource bei der Identifizierung von krankheitsspezifischen epigenetischen Varianten, da eine Keimbahnmutation als Ursache quasi ausgeschlossen ist [86].

Zur weiteren Einordnung der signifikanten Methylierungsunterschiede zwischen den Kontrollen und der CH-Kohorte wurde daher eine Kohorte aus sechs für die CH diskordanten monozygoten Zwillingspaare in die Analyse eingeschlossen.

3.2.4.1 Mittlere Methylierung

Der Vergleich der mittleren Methylierung der neun CpG-Positionen bei diskordanten monozygoten Zwillingen zeigt mit 0,25% eine im Gesamtkontext nicht relevant höhere Methylierungsintensität der CH-Zwillinge im Vergleich zu den gesunden Zwillingen (siehe Tabelle 3-3). Aufgrund der kleinen Fallzahl war eine statistische Signifikanztestung der Ergebnisse nicht möglich.

3.2.4.2 Positionsspezifische Methylierungsunterschiede

An fünf der neun CpG-Positionen sind die CH-Zwillinge höher methyliert als die gesunden Zwillinge (siehe Tabelle 3-3, letzte Spalte). Diese Befunde sind konsistent zu den bisherigen Ergebnissen, welche eine höhere Methylierung unter anderem für diese fünf CpG-Positionen der CH-Kohorte zeigen (siehe 3.2.3.2).

Die größte Methylierungsdifferenz zwischen den gesunden und den CH-Zwillingen zeigt sich an den Positionen eins und drei, welche den mit dem TSH-Wert korrelierenden SNP rs116909374 unmittelbar umgeben (siehe Tabelle 3-3). Aufgrund der kleinen Fallzahl konnte keine statistische Signifikanztestung der Ergebnisse erfolgen. Auch in der Methylierungsanalyse der Kohorten ist die Methylierungsdifferenz zwischen den Kontrollen und den CH-Patienten an der CpG-Position drei am größten (siehe 3.2.3.2). Insgesamt sind die absoluten Methylierungsunterschiede, auch wenn sie tendenziell konsistent zu den Ergebnissen der Kohortenanalyse sind, zwischen den gesunden- und CH-Zwillingen, aber mit einer maximalen mittleren Methylierungsdifferenz von ca.1,3% sehr klein (siehe Tabelle 3-3).

CpG- Position	K - CH	MZZP 1	MZZP 2	MZZP 3	MZZP 4	MZZP 5	MZZP 6	Gesamt K- CH
1	Meth. Diff.	1	-1,5	-1	-1	-2	-1	-0,92
2	Meth. Diff.	0	-1,5	1	0,5	-4	1	-0,67
3	Meth. Diff.	1,5	-5,5	-6	-1	1	0	-1,33
4	Meth. Diff.	0,5	-2	3	-1	2	2	0,75
5	Meth. Diff.	0	1	2	-1,5	0	-1	0,09
6	Meth. Diff.	1	-1	5	1,5	-2	-1	0,41
7	Meth. Diff.	1,5	-1	5	0	1	-2	0,58
8	Meth. Diff.	-0,5	-1,5	-3	0	0	0	-0,83
9	Meth. Diff.	0	-1	3	-1	-2	-1	-0,34
Gesamt	Meth. Diff.	0,44	-1,55	1	-0,39	-0,67	-0,33	-0,25

Tabelle 3-3: Methylierungsdifferenzen monozygoter Zwillingspaare

Die Tabelle zeigt die absolute Methylierungsdifferenz in Prozent zwischen gesunden und CH- Zwillingen für die jeweilige CpG-Position über alle neun CpG-Positionen hinweg sowie pro CpG-Position über alle Zwillingspaare hinweg als Mittelwert der Methylierungsdifferenzen der Zwillingspaare an dieser Position. Die Differenz der Methylierung errechnet sich aus der Subtraktion der Methylierung des CH-Zwillings von der Methylierung des gesunden Zwillings. Ein negativer Wert bedeutet damit eine höhere Methylierung des CH-Zwillings.

MZZP: monozygoten Zwillingenpaar, Meth. Diff.: Methylierungsdifferenz

3.2.4.3 CpG-positionsspezifische Methylierung

Konsistent zu den bisherigen Ergebnissen (siehe Abbildung 3-9) zeigt sich auch bei den monozygoten Zwillingen die niedrigste Methylierung an den zum TSH-Wert assoziierten SNP rs116909374 nahen Positionen eins und drei (siehe Abbildung 3-10), wobei, wie auch im Schilddrüsengewebe, die CpG-Position eins am niedrigsten methyliert ist und nicht die CpG-Position drei (siehe Abbildung 3-9). Zusammenfassend zeigt sich insbesondere für die CpG-Positionen eins bis fünf, abgesehen von kleineren Abweichungen, eine relative Konstanz des positionsspezifischen Methylierungsmusters für die Kohorten, das Schilddrüsengewebe und die monozygoten Zwillinge.

CpG-Position		MZZP 1	MZZP 2	MZZP 3	MZZP 4	MZZP 5	MZZP 6	Gesamt
1	Meth.	70,50	72,75	74,50	66,50	78,00	57,50	69,92
2	Meth.	92,00	90,75	91,50	90,25	91,00	89,50	90,83
3	Meth.	75,25	71,75	74,00	69,50	74,50	60,00	70,83
4	Meth.	79,75	81,00	78,50	79,50	81,00	79,00	79,79
5	Meth.	86,00	85,50	83,00	84,75	83,00	85,50	84,63
6	Meth.	94,50	92,50	90,50	93,75	94,00	93,50	93,13
7	Meth.	94,75	93,50	91,50	94,00	94,50	91,00	93,21
8	Meth.	92,25	91,75	92,50	90,00	92,00	91,00	91,58
9	Meth.	94,00	93,50	91,50	93,50	93,00	92,50	93,00
Gesamt	Meth.	86,55	85,88	85,27	84,63	86,77	82,17	85,21

Tabelle 3-4: Durchschnittliche Methylierung pro Zwillingsspaar pro CpG-Position

Die angegebenen Werte entsprechen jeweils der durchschnittlichen Methylierung in Prozent.

Dargestellt ist die Methylierung pro Zwillingsspaar für jede der neun CpG-Positionen (siehe Spalten) sowie die Methylierung jedes Zwillingsspaars über alle neun CpG-Positionen (siehe letzte Zeile). Weiter ist die Methylierung aller Zwillingsspaare pro CpG-Position dargestellt (siehe letzte Spalte) sowie die Methylierung aller Zwillingsspaare über alle neun CpG-Positionen (siehe letzte Zeile letzte Spalte).

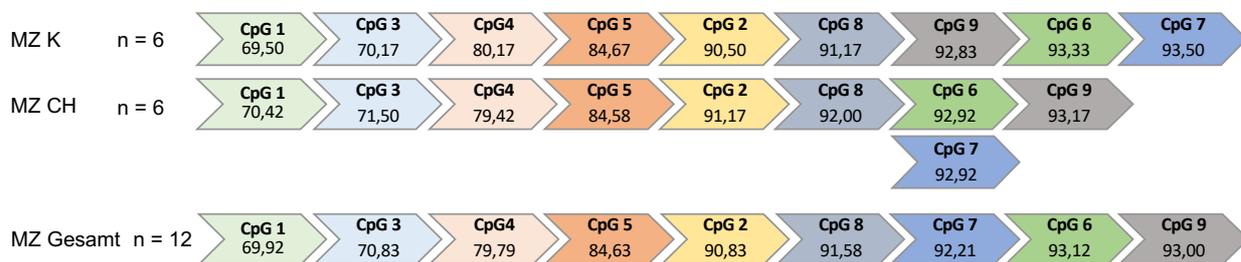


Abbildung 3-10: Monozygote Zwillinge (MZ), CpG-Positionen aufsteigend nach Methylierung sortiert

Dargestellt ist die durchschnittliche Methylierung in Prozent pro CpG-Position der monozygoten Zwillinge in aufsteigender Reihenfolge, sowohl kohortenspezifisch als auch kohortenübergreifend. MZ K: gesunde Kontrollen der monozygoten Zwillinge, MZ CH: an CH erkrankte monozygote Zwillinge, MZ Gesamt: Gesamtheit aller monozygoten Zwillinge.

3.2.4.4 Zwillingsspaarspezifische Methylierung

Betrachtet man die Methylierungsunterschiede der sechs Zwillingsspaare an den neun CpG-Positionen (siehe Abbildung 3-11 und Tabelle 3-4), so fällt auf, dass die größten Methylierungsunterschiede zwischen den Zwillingsspaaren konsistent zu den bisherigen Ergebnissen an den CpG-Positionen eins und drei zu finden sind.

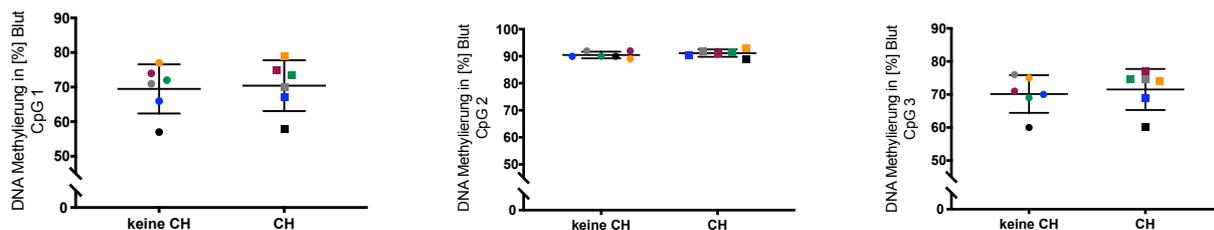


Abbildung 3-11: CpG-positionsspezifische Methylierung der monozygoten Zwillingspaare

Dargestellt ist exemplarisch für die CpG-Positionen eins bis drei die mittlere Methylierung in Prozent der diskordanten monozygoten Zwillinge. Dabei ist vergleichend die Methylierung der gesunde Zwillinge (keine CH) den an CH erkrankten Zwillingen gegenübergestellt. Jedem Zwillingpaar ist hierbei eine Farbe zugeordnet.

Zusammenfassend sind die Methylierungsunterschiede zwischen den CH-Zwillingen und den gesunden Zwillingen sehr gering. Da der Vergleich von diskordanten monozygoten Zwillingen die beste Methode zur Analyse von CH assoziierten Epimutationen ist, sprechen die Ergebnisse dafür, dass sich CH-Patienten in der PER des *NKX2-1*-Gens nicht oder kaum von gesunden Kontrollen unterscheiden.

3.2.5 Geschlechtsspezifische Methylierung

Aufgrund der Mädchenwendigkeit der CH [20] erfolgte eine Analyse bezüglich geschlechtsspezifischer Methylierungsunterschiede in der potentiellen Enhancer Region.

3.2.5.1 Kohortenunabhängige geschlechtsspezifische Methylierung

Die Analyse zeigt, dass Frauen an den neun CpG-Positionen statistisch signifikant höher methyliert sind als Männer (siehe Tabelle 3-5).

	Frauen	Männer	Differenz	P-Wert
Gesamt	85,85	84,74	1,11	0,0011 **
Kontrollen	83,94	84,07	-0,13	>0,1234 ns
HTT	84,11	83,76	0,35	>0,1234 ns
CH	86,41	85,47	0,94	0,0238 *

Tabelle 3-5: Geschlechtsspezifische Methylierungsunterschiede zwischen den Kohorten

Dargestellt sind die Mittelwerte der Methylierung in Prozent für die jeweiligen Kohorten und Geschlechter. Die Differenz ergibt sich aus der Subtraktion des Mittelwertes der Männer vom Mittelwert der Frauen. Ein positiver Wert bedeutet demnach eine höhere Methylierung der Frauen. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) wurde der nichtparametrische beidseitige Mann-Whitney-Test durchgeführt. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***); < 0,0001(****).

3.2.5.2 Kohortenabhängige geschlechtsspezifische Methylierung

Bei Betrachtung der geschlechtsspezifischen Methylierungsunterschiede der Kohorten zeigte sich ein interessantes Ergebnis. So ergab sich einzig für die CH-Kohorte eine statistisch signifikant höhere Methylierung der Frauen. In der HTT-Kohorte zeigte sich eine nicht statistisch signifikant höhere Methylierung der Frauen. Im Gegensatz dazu waren in der Kontrollgruppe die Männer sogar höher methyliert als die Frauen, allerdings ist der Unterschied klein und nicht statistisch signifikant (siehe Tabelle 3-5).

3.2.5.3 Geschlechtsspezifische Methylierungsunterschiede zwischen den Kohorten

Beim Vergleich von Kohorten spezifischen Methylierungsunterschieden innerhalb der Gruppe der Frauen zeigte sich eine statistisch signifikant höhere Methylierung der CH-Kohorte gegenüber den Kontrollen und der HTT-Kohorte (siehe Tabelle 3-6 und Abbildung 3-12). Dies bestätigt die oben genannten Befunde. Ein Vergleich der Kohorten spezifischen Methylierungsunterschiede innerhalb der Gruppe der Männer zeigte keine statistisch signifikanten Unterschiede (siehe Tabelle 3-6 und Abbildung 3-12). Daraus lässt sich schließen, dass der geschlechtsspezifische Methylierungsunterschied letztendlich auf eine statistisch signifikant erhöhte Methylierung der CH-Frauen zurückzuführen ist.

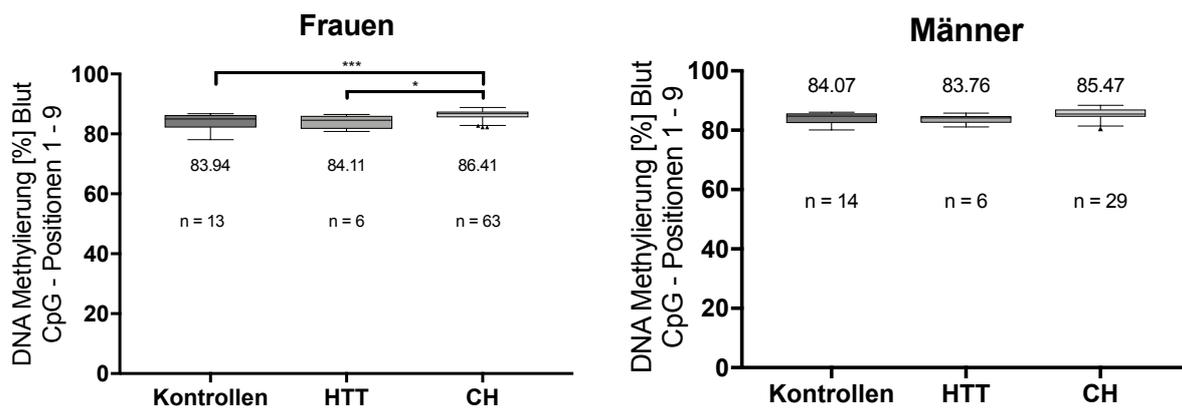


Abbildung 3-12: Geschlechtsspezifische Methylierungsunterschiede

Dargestellt ist geschlechter- und kohortengetrennt die mittlere Methylierung in Prozent über alle neun CpG-Positionen hinweg. Die Zahl über den jeweiligen Graphen gibt den Mittelwert der Methylierung der Kohorte an. Die Zahl n beschreibt die Personenzahl der Kohorte. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***) < 0,0001(****).

	Dif. Berechnung	Differenz	P-Wert	Geschlecht
K vs. HTT	83,94 - 84,11	-0,17	>0,1234 ns	Frauen
K vs. CH	83,94 - 86,41	-2,47	0,0008 ***	Frauen
HTT vs. CH	84,11 - 86,41	-2,30	0,0193 *	Frauen
K vs. HTT	84,07 - 83,76	0,31	>0,1234 ns	Männer
K vs. CH	84,07 - 85,47	-1,40	>0,1234 ns	Männer
HTT vs. CH	83,76 - 85,47	-1,71	>0,1234 ns	Männer

Tabelle 3-6: Geschlechtsspezifische Methylierungsdifferenzen zwischen den Kohorten

Dargestellt ist die geschlechtsspezifische mittlere Methylierung in Prozent der Kohorten.

Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***) < 0,0001(****). K: Kontrollen; HTT: Hyperthyreotropinämie; CH: kongenitale Hyperthyreose; Dif. Berechnung: Differenzen Berechnung; ns: nicht signifikant

3.2.6 Belastbarkeit der Ergebnisse

3.2.6.1 Robustheit der Pyrogramm Daten

Einige der verwendeten Proben wurden mehrfach pyrosequenziert. Dies trifft für insgesamt 28 der analysierten Proben zu. Davon waren 18 Proben von CH-Patienten, drei von an CH erkrankten monozygoten Zwillingen, fünf von gesunden Kontrollen und zwei Proben von Schilddrüsengewebe gesunder Personen. Bei 22 der 28 Proben ist die

vorausgegangene PCR aus der gleichen DNA erfolgt. Bei fünf der Proben ist die PCR aus verschiedenen Blutproben der gleichen Person erfolgt. Bei zwei Personen erfolgte die PCR aus der gleichen Blutprobe und beide Pyrosquenzierungen im gleichen Run.

Zur Überprüfung der Robustheit der Daten wurde die Auswertung der Daten einerseits durch Bestimmung des Mittelwertes aller Ergebnisse des gleichen Patienten als auch durch Auswahl des qualitativ hochwertigeren Pyrogramms durchgeführt. Dabei erfolgte die Qualitätsbewertung des Pyrogramms durch einen Analyse-Algorithmus der entsprechenden Qiagen-Software.

Die Ergebnisse zeigten unabhängig von der Datenauswahl keine signifikanten Unterschiede im Mittelwert der Methylierung über alle neun CpG-Positionen (siehe Abbildung 3-13).

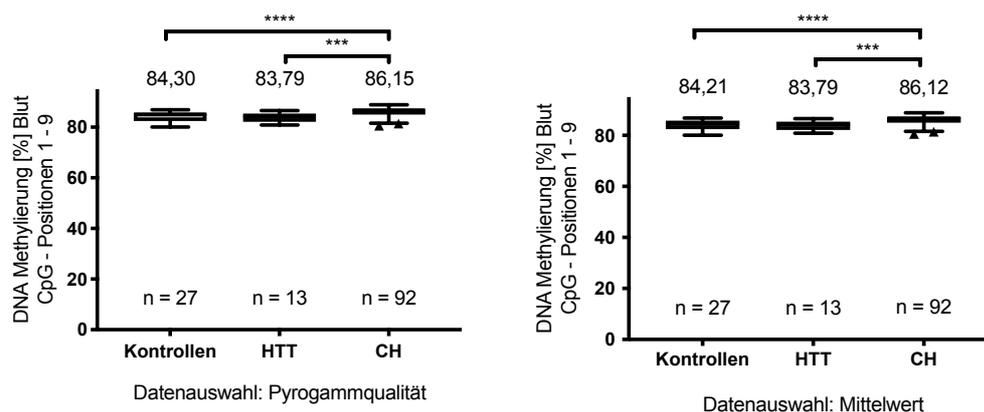


Abbildung 3-13: Signifikanzvergleich des Mittelwertes der Methylierung nach Datenauswahl

Die Zahl über den jeweiligen Graphen gibt den Mittelwert der Methylierung in Prozent der Kohorte an. Die Zahl n beschreibt die Personenzahl der Kohorte. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***) < 0,0001(****).

Auch bei der CpG-positionsspezifischen Analyse hatte die Datenauswahl keinen Einfluss auf die Interpretation der Ergebnisse (siehe Abbildung 3-14).

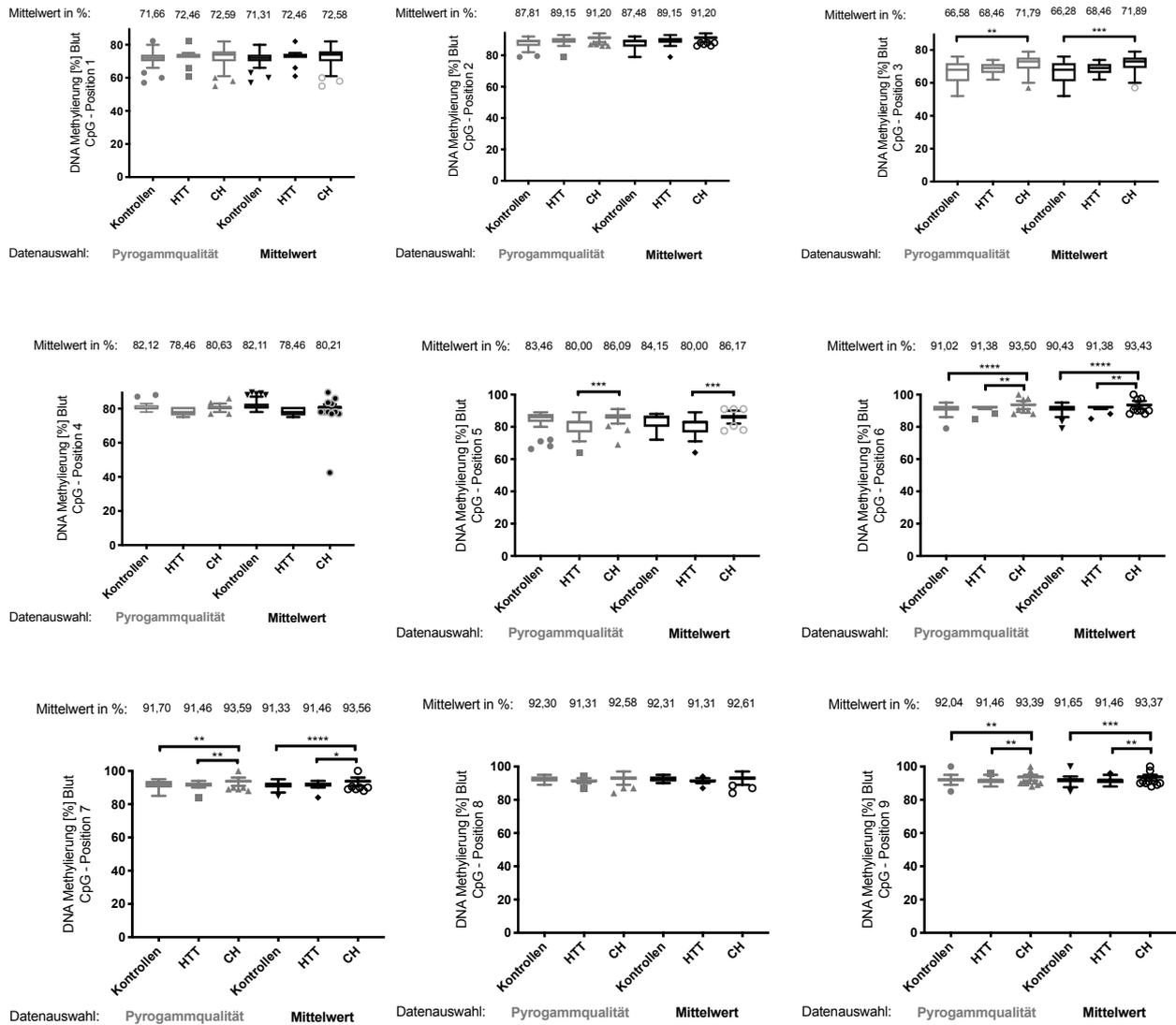


Abbildung 3-14: Signifikanzvergleich des Mittelwertes der Methylierung nach Datenauswahl pro CpG-Position

Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt.

P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***); < 0,0001(****).

Die Daten können somit als robust gewertet werden. Das heißt, die Ergebnisse zeigen keine Abhängigkeit von der verwendeten Datenauswahl.

3.2.6.2 Runspezifischer Methylierungsunterschied

Weiter wurde das gleiche PCR-Produkt in verschiedenen Pyrosequenzierungs-Runs analysiert, um relevante runspezifische Methylierungsunterschiede auszuschließen (siehe Abbildung 3-15, Abbildung 3-16 und Abbildung 3-17).

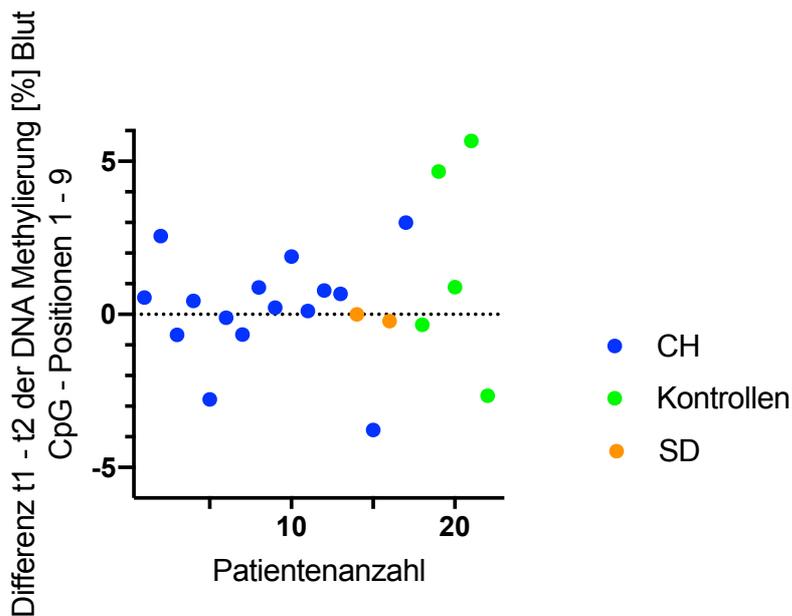


Abbildung 3-15: Runspezifischer Methylierungsunterschied

Dargestellt ist die Differenz des Mittelwerts der Methylierung in Prozent über alle neun CpG-Positionen aller Personen, für die Methylierungsergebnisse aus zwei verschiedenen Pyrosequenzierungs-Runs vorliegen. Dabei wurde jeweils die Differenz des absoluten Mittelwerts der Methylierung der verschiedenen Runs gebildet, wobei der Mittelwert der Methylierung des zeitlich gesehen zweiten Runs vom ersten Run subtrahiert wurde. Die Kohortenzugehörigkeit wurde farblich dargestellt.

Bei der Interpretation der Ergebnisse aus Abbildung 3-15 ist zu beachten, dass eine Wiederholung der Pyrosequenzierungs-Runs nicht selten aufgrund einer schlechten Pyrogrammqualität erfolgt ist und daher die Differenz zwischen den Pyrosequenzierungs-Runs tendenziell überschätzt wird. In der Analyse der Robusheit der Pyrogrammdaten (siehe 3.2.6.1) konnte bereits gezeigt werden, dass die Datenauswahl und damit auch die Differenz der Methylierung verschiedener Pyrosequenzierungs-Runs keine Relevanz für die Interpretation der Daten hat.

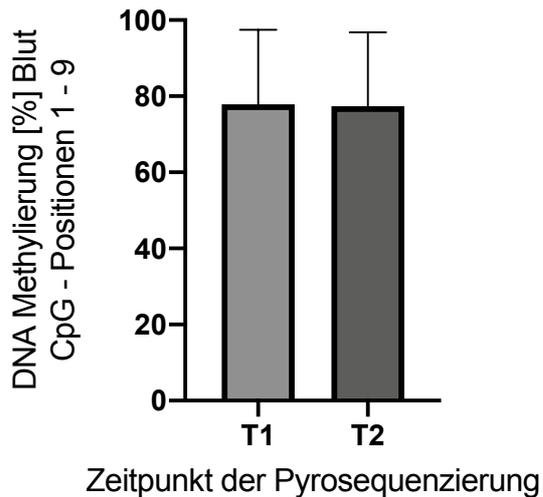


Abbildung 3-16: Kohortenübergreifender Vergleich der mittleren Methylierung verschiedener Pyrosequenzierungs-Runs

Dargestellt ist der Mittelwert der Methylierung mit Standardabweichung über alle neun CpG-Positionen derjenigen Patienten oder gesunden Kontrollenpersonen, für die Daten aus zwei verschiedenen Pyrosequenzierungs-Runs vorhanden sind. Das Säulendiagramm zeigt in T1 den Mittelwert der Methylierung in Prozent des zeitlich gesehen ersten Runs der jeweiligen Patienten und gesunden Kontrollen und T2 entsprechend den Mittelwert der Methylierung in Prozent des zeitlich gesehen zweiten Runs. Dabei war die Anzahl der einbezogenen Personen $n = 22$. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Mann-Whitney-Test. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***) < 0,0001(****).

Auch die kohortenübergreifende Analyse der mittleren Methylierung verschiedener Pyrosequenzierungs-Runs der gleichen Personen konnte keinen signifikanten runspezifischen Unterschied zeigen (siehe Abbildung 3-16).

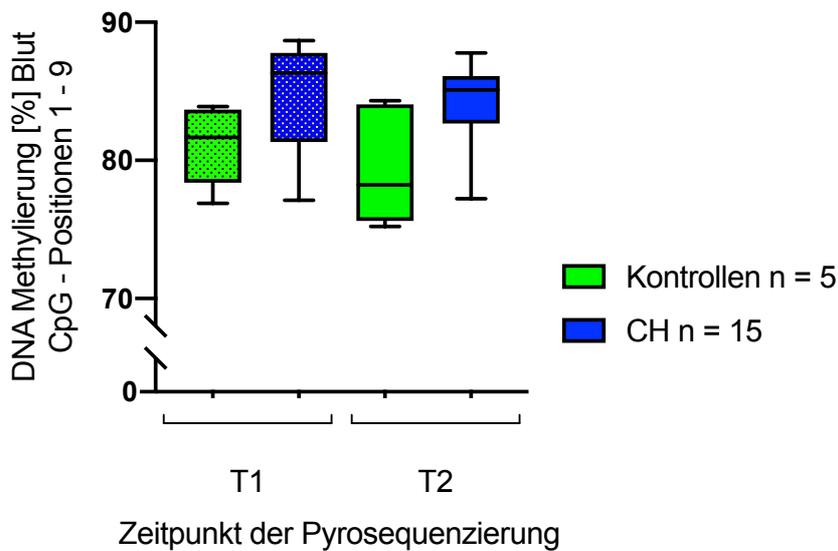


Abbildung 3-17: Kohortenspezifischer Vergleich der mittleren Methylierung verschiedener Pyrosequenzierungs-Runs

Dargestellt ist vergleichend kohortenspezifisch der mittlere Methylierungsunterschied über die neun CpG-Positionen zwischen verschiedenen Pyrosequenzierungs-Runs. Dabei ist T1 zeitlich gesehen der erste Run und T2 entsprechend zeitlich gesehen der zweite Run, in dem DNA der gleichen Person analysiert wurde. N (n) gibt dabei die Anzahl der einbezogenen Personen an. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt. P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***) < 0,0001(****).

Auch die Betrachtung kohortenspezifischer Unterschiede in der mittleren Methylierung zeigt keine signifikanten runspezifischen Unterschiede (siehe Abbildung 3-17).

3.2.6.3 Abhängigkeit der Methylierung vom Alter des Patienten bei der Blutentnahme

Eine Abhängigkeit der Methylierung vom Patientenalter zum Zeitpunkt der Blutentnahme konnte ausgeschlossen werden, da die Korrelation zwischen beiden nicht signifikant ist (siehe Abbildung 3-18).

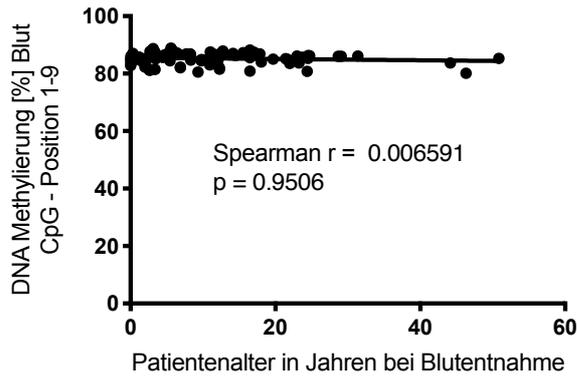


Abbildung 3-18: Korrelation Patientenalter bei Blutentnahme und Methylierung

Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) und Berechnung einer Regressionsgraden wurde zur Interpretation des Zusammenhangs eine nichtparametrische Spearman-Korrelation durchgeführt, die nicht signifikant war.

3.2.6.4 Abhängigkeit der Methylierung vom Alter des gleichen Patienten

Es besteht keine relevante Altersabhängigkeit der Messwerte zu verschiedenen Blutentnahmezeitpunkten des gleichen Patienten (siehe Abbildung 3-19).

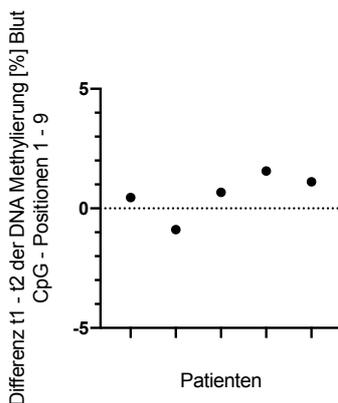


Abbildung 3-19: Abhängigkeiten der Methylierung vom Alter des gleichen Patienten

Abhängigkeit der Güte der Pyrosequenzierung von verschiedenen Blutentnahmezeitpunkten eines Patienten. Dargestellt ist die Differenz des Mittelwertes der Methylierung der analysierten CpG-Positionen verschiedener Blutproben von fünf Individuen.

3.2.6.5 Abhängigkeit der Methylierung vom Alter der Blutprobe

Ein Vergleich des Blutprobenalters der drei analysierten Kohorten zeigt, dass die Blutproben der CH-Kohorte deutlich älter sind als die der HTT-Kohorte und vor allem der Kontrollen (siehe Abbildung 3-20 und Abbildung 3-22). Hätte die Lagerungsdauer der Probe oder ein mit ihr verbundener Faktor einen Einfluss auf die Methylierung, würde das den Vergleich der Methylierung zwischen den Kohorten verfälschen. Eine Veränderung

der Methylierung mit längerer Lagerungsdauer ist nicht anzunehmen. Allerdings wäre es möglich, dass sich die präanalytische Probenqualität über die Zeit verändert hat und dies die Ergebnisse beeinflusst.

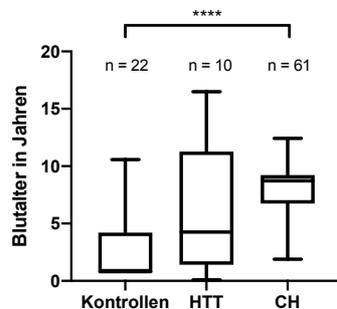


Abbildung 3-20: Kohortenspezifisches Alter der Blutproben

Dargestellt ist kohortenspezifisch das Blutalter der drei analysierten Kohorten. Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) erfolgte der nicht parametrische Kruskal-Wallis-Test mit Korrektur für multiples Testen mittels Dunn's Test. Es wurde die Tukey-Box-Plot-Darstellung gewählt, die den Median, die zwei Quartile, die beiden Extremwerte und eventuelle Ausreißer zeigt.

P-Werte: 0,1234 (ns); 0,0332(*); 0,0021(**); 0,0002(***); < 0,0001(****).

Um auszuschließen, dass die Unterschiede in der Methylierung zwischen den Kohorten auf Unterschiede in der präanalytischen Probequalität zurückzuführen sind, wird getestet, ob das Probenalter einen Einfluss auf die Methylierung hat. Da die vorliegende Arbeit gezeigt hat, dass sich die Kohorten in der Methylierung unterscheiden, werden das Probenalter und die Methylierung für jede Gruppe getrennt betrachtet. Dabei zeigt sich keine signifikante Korrelationen zwischen dem Alter der Blutprobe und der Methylierung (siehe Abbildung 3-21). Die Schätzung ist sehr ungenau in der Kohorte der Kontrollen und der HTT-Kohorte, da diese sehr klein sind, aber deutlich präziser geschätzt in der CH-Kohorte (siehe Abbildung 3-22).

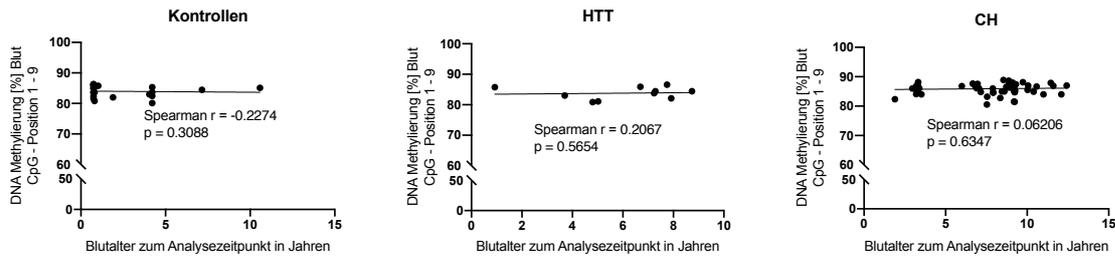


Abbildung 3-21: Abhängigkeit der Methylierung vom Alter der Blutprobe innerhalb der Kohorten

Nach Test auf Normalverteilung (D'Agostino & Pearson omnibus Normalitätstest) und Berechnung einer Regressionsgraden wurden zur Interpretation des Zusammenhangs eine nicht-parametrische Spearman-Korrelation oder eine parametrische Pearson-Korrelation durchgeführt.

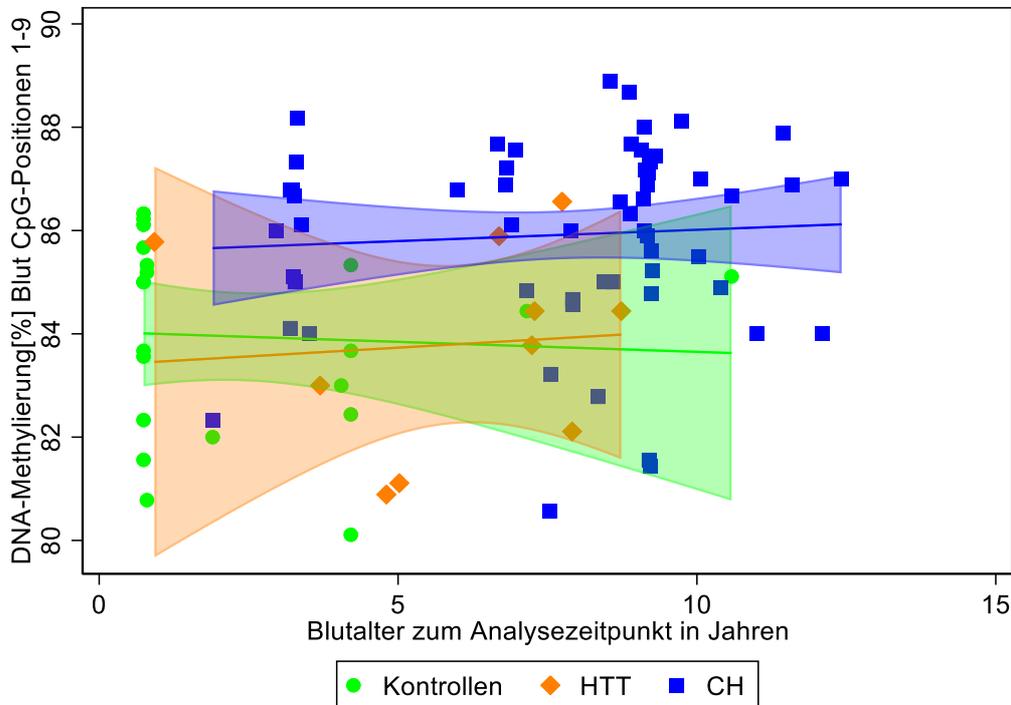


Abbildung 3-22: Kohortenspezifische Abhängigkeit der Methylierung vom Alter der Blutprobe

Zusammenhang vom Alter der Blutprobe und der Methylierung nach Kohortenzugehörigkeit. Dargestellt sind die Datenpunkte mit Regressionsgrade und 95% Konfidenzintervall.

Das Ergebnis bestätigt sich in der multiplen Regression: Wenn man die Kohorten mittels multipler Regression kontrolliert, hat das Blutalter keinen statistisch signifikanten Einfluss auf die mittlere Methylierung (siehe Tabelle 3-7). Die Effektgröße liegt nahe null, ist präzise geschätzt und nicht statistisch signifikant unterschiedlich von null.

Variable	Regressions- koeffizient	Standard- abweichung	P-Wert		95% Konfidenzintervall
Blutalter	0,02701	0,07305	0,7124	ns	-0,1181 bis 0,1722
HTT	-0,256	0,7488	0,7333	ns	-1,744 bis 1,232
CH	1,814	0,6007	0,0033	**	0,6205 bis 3,008
Intercept (Kontrollen)	83,89	0,4264	<0,0001	****	83,05 bis 84,74

Tabelle 3-7: Multiple Regression der Methylierung in Abhängigkeit vom Blutalter und der Kohorte

Weiter zeigt die multiple Regression, dass die mittlere Methylierung der CH-Kohorte auch unter Kontrolle des Blutalters höher ist als die der Kontrollgruppe. Das heißt, die höhere Methylierung der CH-Kohorte ist keine Scheinkorrelation, die durch das unterschiedliche Blutalter der Kohorten entsteht. Dies bestärkt das Ergebnis einer signifikant höheren Methylierung der CH-Kohorte. Dieser Befund kann nicht durch eine Abhängigkeit der Methylierung vom Alter der Blutprobe erklärt werden, da das Ergebnis auch unter Kontrolle des Blutalters robust bleibt.

4 Diskussion

Die kongenitale Hypothyreose ist die häufigste angeborene endokrinologische Erkrankung [1]. Dennoch ist die molekulare Ursache der primären kongenitalen Hypothyreose trotz intensiver Forschung in den meisten Fällen weiter unbekannt. Die häufigste Form der primären kongenitalen Hypothyreose ist die Schilddrüsendysgenese [4], eine Störung der Schilddrüsenentwicklung. In lediglich fünf Prozent der Fälle einer Schilddrüsendysgenese konnten bisher ursächliche autosomal-dominante Mutationen in für Transkriptionsfaktoren kodierenden Genen sowie im TSH-Rezeptor gefunden werden [1]. In diesen Fällen tritt die kongenitale Hypothyreose im Rahmen verschiedener syndromaler Erkrankungen auf. Die Ursache der isolierten kongenitalen Hypothyreose bleibt in den meisten Fällen weiterhin unklar.

Epidemiologische Beobachtungen wie das sporadische Auftreten und die hohe Diskordanzrate bei monozygoten Zwillingen sprechen gegen starke mendelsche genetische Faktoren als primäre Ursache der Schilddrüsendysgenese [10].

Auch für viele andere Krankheiten ist die zugrunde liegende genetische Ursache bisher unklar. Über die Gründe und mögliche Erklärungsansätze wird kontrovers diskutiert.

Ein möglicher Ansatz zur Klärung dieser Diskrepanz ist unter anderem das Einbeziehen von Epimutationen und Mutationen in nicht-kodierenden Regionen als mögliche molekulare Ursache der isolierten Schilddrüsendysgenese. Grundlage für diesen Ansatz ist unter anderem eine Studie des IEPES von Thorwarth et al. 2014, in der gezeigt werden konnte, dass zwei nicht-kodierende Mutationen downstream des *NKX2-1-Gens* die entsprechende syndromale Erkrankung hervorgerufen haben [31]. Als Erklärungsansatz wurde angenommen, dass diese nicht-kodierende Region für regulative Prozesse, z.B. in Form einer schilddrüsen-spezifischen Enhancerfunktion verantwortlich ist. Mutationen und Epimutationen in regulativen nicht-kodierenden DNA-Regionen könnten so eine mögliche Ursache der isolierten Schilddrüsendysgenese sein. Ein solcher Erklärungsmechanismus wurde für die Pankreasagenese gefunden:

Weedon et al. 2014 konnten zeigen, dass ein Großteil der Fälle einer isolierten Pankreasagenese durch nicht-kodierende Mutationen in einem Enhancerelement downstream des für die Pankreasentwicklung essentiellen Gens *PTF1A* verursacht

werden. Wobei, ähnlich wie bei der Schilddrüsendysgenese, Mutationen im *PTF1A*-Gen bereits zuvor als Auslöser der syndromalen Pankreasagenese beschrieben wurden [21].

Ein weiterer Erklärungsansatz für das Fehlen eines klaren genetischen Korrelats von Krankheiten ist das sogenannte „Two-Hit-Modell“, nach dem eine genetische Keimbahnmutation vererbt wird, die erst zusammen mit einer post-zygotischen, gewebespezifischen oder Epimutation zur Störung der Schilddrüsenentwicklung führen könnte [10].

Auch eine Beteiligung von allelspezifischen Methylierungsunterschieden an der Entstehung von Krankheiten, die einerseits durch genetische Varianten wie z.B. SNPs, aber auch durch Imprinting, Umwelteinflüsse und stochastische Ereignisse verursacht werden können, wird vermutet [98].

Bisherige Bemühungen, die molekulare Grundlage der isolierten kongenitalen Hypothyreose zu entschlüsseln, bezogen sich vor allem auf die Analyse von Genen, die an der Schilddrüsenentwicklung beteiligt sind. Aufbauend auf den oben genannten Ergebnissen des IEPes von Thorwarth et al. wurde in dieser Arbeit ein nicht-kodierender Bereich des *NKX2-1*-Gens auf Mutationen, Epimutationen und SNP-Assoziationen analysiert.

4.1 Genotypisierung

4.1.1 Kohortenspezifische SNP-Verteilung in der PER

Genomweite SNP-Analysen konnten eine Korrelation des in der potentiellen Enhancer Region des *NKX2-1*-Gens befindlichen SNPs rs116909374 mit dem TSH-Wert nachweisen [41, 42]. Es wird vermutet, dass bei vielen Erkrankungen, bei denen bisher vergebens nach klassischen mendelschen Mutationen gesucht wurde, nicht-kodierende SNPs entscheidend an der Krankheitsentstehung beteiligt sind [99].

Sifrim et al. 2016 konnten z.B. zeigen, dass Patienten mit *isolierten* kongenitalen Herzdefekten eine Anreicherung von SNVs mit einer MAF < 0,1% in für kongenitale Herzdefekte relevanten Genen aufweisen, die durch gesunde Eltern vererbt werden und zu einer gestörten Expression der Gene führen, die ursächlich für Herzdefekte sind. Hingegen konnten für *syndromale* kongenitale Herzdefekte bisher de novo Mutationen in für die Krankheit relevanten Genen gefunden werden [100].

Auch bei der Entstehung der kongenitalen Hypothyreose wird als prädisponierender Faktor für die Entstehung unter anderem eine vererbte SNP-Kombination vermutet [10]. In dieser Arbeit wurde deshalb mittels Sanger-Sequenzierung die Allelverteilung insbesondere des mit dem Serum TSH-Wertes korrelierenden SNPs rs116909374, aber auch der 37 weiteren zum Analysezeitpunkt in der dbSNP(147) Datenbank hinterlegten SNPs in der PER auf mögliche Verteilungsunterschiede zwischen Patienten mit isolierter kongenitaler Hypothyreose, Patienten mit Hyperthyreotropinämie, einer milden Störung der Schilddrüse, die mit einem erhöhtem TSH-Wert einhergeht, und gesunden Kontrollen analysiert (siehe 3.1.3). In jeder der analysierten Kohorten traten heterozygote Minor Allelträger des SNPs rs116909374 auf. Auch eine einheitliche Allelausprägung des mit dem Serum TSH-Wert korrelierenden SNP rs116909374 in der CH-Kohorte sowie ein schilddrüsenphänotypischer Zusammenhang der SNP-Allelausprägung in der CH-Kohorte konnte nicht gezeigt werden (siehe Tabelle 3-1). Bei der Interpretation der Daten ist zu beachten, dass auf Grund von kleinen Kohortengrößen und der ungleichen Verteilung der Personenzahl die Bestimmung der MAF im Kohortenvergleich nicht zielführend war und daher nicht erfolgt ist. Eine Aussage über kohortenspezifische Unterschiede in der SNP-Verteilung ist daher nicht möglich. Bei größeren Fallzahlen wäre der Vergleich der MAF des mit dem Serum TSH-Wert korrelierenden SNP rs116909374 sowie der weiteren in der PER befindlichen SNPs der verschiedenen Kohorten ein interessanter Aspekt, um mögliche Unterschiede der MAF zwischen den Kohorten zu detektieren. Eine Reanalyse der SNP-Allelverteilung in der potentiellen Enhancer Region des *NKX2-1*-Gens in deutlich größeren und vor allem auch gleich großen Kohorten wäre daher sinnvoll.

Dass aber einzig in der CH-Kohorte von zwei weiteren SNPs heterozygot die Minor Allele zu finden waren und sogar zwei zum Zeitpunkt der Analyse im UCSC Genome Browser dbSNP(147) Datenbank nicht beschriebene neue SNVs gefunden wurden (siehe Abbildung 3-2), von denen einer der SNVs in der neusten Version der Datenbank dbSNP(153), allerdings inzwischen als seltener SNP rs993803802 hinterlegt ist, könnte dennoch ein Hinweis für die Beteiligung von seltenen SNPs an der Krankheitsentstehung z.B. in Form einer möglichen Störung des postulierten schilddrüsen-spezifischen Enhancers durch seltene SNPs sein.

Auch Thorwarth et al. 2010 begründen die Relevanz von CNVs für die Entstehung der Schilddrüsendysgenese mit neu gefundenen nicht in der gesunden Bevölkerung nachweisbaren CNVs in Patienten mit Schilddrüsendysgenese. Da, wie auch in dieser Arbeit die SNVs, die neuen CNVs nicht wiederkehrend auftraten, wurde geschlussfolgert, dass der Schilddrüsendysgenese am ehesten eine heterogene molekulare Ursache zu Grunde liegen könnte, die zu einer krankheitsauslösenden Regulationsstörung führt. Der Schilddrüsenphänotyp der Patienten mit Nachweis einer CNV war vorwiegend eine Athyreose - die schwerste Form der Schilddrüsendysgenese [11]. Auch der Patient in dieser Arbeit, der heterozygoter Träger einer der beiden zum Analysezeitpunkt neuen SNVs und inzwischen heterozygoter Träger des SNVs rs993803802 und zusätzlich heterozygoter Träger des mit dem TSH-Wert korrelierenden SNPs rs116909374 ist, zeigt phänotypisch eine Athyreose. Für den Träger der zum Analysezeitpunkt zweiten neuen SNV und aktuell einzigen neuen SNV ist der Phänotyp leider nicht bekannt. Eine krankheitsauslösende Störung des postulierten schilddrüsen-spezifischen Enhancers durch die in der CH-Kohorte gefundenen SNPs, SNVs und dem neu gefundenen SNV ist in Zusammenschau der Ergebnisse denkbar. Die SNPs, SNVs und der neu gefundene SNV in der CH-Kohorte könnten dabei Folge einer Anreicherung von vererbten SNPs und SNVs sein, die zu einer Störung der postulierten schilddrüsen-spezifischen Enhancerfunktion führen könnte, wie es Sifrim et al. 2016 für kongenitale Herzdefekte zeigen konnten [100]. Bei der Interpretation der Ergebnisse muss allerdings beachtet werden, dass die CH-Kohorte von allen analysierten Kohorten mit Abstand die größte Kohorte ist und die Häufung von SNPs und SNVs sowie die Entdeckung eines neuen SNVs in der CH-Kohorte dadurch beeinflusst sein könnte. Weiter ist der ethnische Hintergrund der einzelnen Personen der Kohorten nicht bekannt, sodass mögliche ethnische Unterschiede der SNP-Allelverteilung nicht berücksichtigt werden konnten. In der neusten Version der dbSNP(153) Datenbank ist die Anzahl der SNPs in der PER des *NKX2-1*-Gens mit insgesamt 188 fast fünfmal so groß wie zum Analysezeitpunkt in der dbSNP(147) Datenbank. Dieser Befund macht eine funktionelle Relevanz der gefundenen SNPs deutlich unwahrscheinlicher.

Um eine Aussage über mögliche kohortenspezifische Unterschiede in der SNP-Allelverteilung treffen zu können, ist eine erneute Analyse der PER in deutlich größeren Kohorten notwendig.

Um den Einfluss der SNP-Verteilung auf die Entstehung der Schilddrüsendysgenese näher zu untersuchen, sind weitere genomweite SNP-Analysen wie z.B. Linkage Disequilibrium Untersuchungen nötig. Linkage Disequilibrium beschreibt dabei eine nicht zufällige Assoziation von Allelen verschiedener Genom Loci, deren Assoziationsfrequenz höher oder niedriger ist, als es bei unabhängigen Allelen zu erwarten wäre [101]. Durch eine derartige Analyse könnte eine krankheitsrelevante, nicht zufällige Assoziation von mehreren SNP-Allelen in Patienten mit kongenitaler Hypothyreose gefunden werden [99].

In der hier betrachteten potentiellen Enhancer Region scheint die quantitative Häufigkeit der durch die 1000 Genomdaten ermittelten SNPs zum Analysezeitpunkt mit 38 für eine nicht-kodierende Region weder ungewöhnlich hoch noch sehr niedrig zu sein. Allerdings sind von den 38 SNPs nur zwei mit einer MAF von über einem Prozent SNPs im eigentlichen Sinne. Bei allen anderen Varianten handelt es sich definitionsgemäß um SNVs. Einer dieser beiden SNPs ist hierbei der mit dem TSH-Wert korrelierende SNP rs116909374. In der neuesten Version der dbSNP(153) Datenbank ist die Anzahl der SNVs in der PER inzwischen auf 186 gestiegen, allerdings zeigt sich die Anzahl der SNPs weiter konstant bei zwei. Die niedrige MAF der übrigen SNVs könnte durch evolutionären Druck bedingt sein. Eine Studie von Wang et al. 2018 zeigt, dass in regulativen nicht-kodierenden DNA-Bereichen die Tertiärstruktur der DNA und damit die Transkriptionsfaktorbindung durch SNPs mit einer kleinen MAF stärker beeinflusst wird als durch SNPs mit einer großen MAF [102, 103]. Umgekehrt bedeutet dies, dass SNPs, die in regulativen nicht-kodierenden Bereichen der DNA einen großen Einfluss auf die Konformation der DNA haben und damit eine potentiell funktionell relevante Transkriptionsfaktorbindung verhindern, einem hohen evolutionären Druck unterliegen und daher selten vorkommen. Es wird vermutet, dass dieser Mechanismus eine große Rolle bei krankheitsspezifischen SNPs spielt [103]. Auch Weedon et al. 2014 beschreiben, dass der für die isolierte Pankreasagenese ursächliche, 400 Basenpaare große, intergenische Enhancer hoch konserviert ist und die drei zuvor in der USCS Genome Browser dbSNP Datenbank beschriebenen SNVs eine MAF < 0,2% aufweisen [21].

4.1.2 Beeinflussung der Transkriptionsfaktorbindung durch SNPs

Hamid et al. 2019 konnten zeigen, dass eine veränderte Bindung von Transkriptionsfaktoren durch das Vorhandensein des Minor Allels eines SNPs für eine

veränderte Expression eines Zielgens verantwortlich sein kann. Diese SNPs werden als regulatorische SNPs (rSNPs) bezeichnet [104]. Bezogen auf die PER könnte es durch die Ausprägung von Minor Allelen verschiedener SNPs und SNVs mit einer geringen MAF zu einer gestörten Transkriptionsfaktorbindung und damit einhergehend zu einer Funktionseinschränkung des potentiellen schilddrüsenspezifischen Enhancers kommen. Die Ergebnisse dieser Arbeit zeigen, dass alle in der CH-Kohorte gefundenen heterozygoten Minor Allele von SNPs ebenso wie der neu entdeckte SNV in unmittelbarer räumlicher Nachbarschaft zu dem mit dem TSH-Wert korrelierenden SNP rs116909374 liegen. In diesem Bereich binden auch die drei Transkriptionsfaktoren ESR1, FOXA1 und NR3C1. Weiter scheint auch genau dieser Bereich um den SNP rs116909374 ein DNaseI-Hypersensitivity-Cluster zu sein, was ein Marker für DNA-Bereiche mit regulativer Funktion im Genom ist [105] (siehe Abbildung 3-1).

Durch die Lokalisation der heterozygoten SNPs und SNVs sowie des neu gefundenen heterozygoten SNVs in direkter Nachbarschaft zu dem mit dem TSH-Wert korrelierenden SNP rs116909374 und dem Bereich von potentiellen Transkriptionsfaktorbindungsstellen ist der oben beschriebene Mechanismus einer Konformationsstörung der DNA durch die SNPs mit nachfolgender Transkriptionsfaktorbindungsstörung nicht auszuschließen.

Dennoch zeigen die Ergebnisse weder einen einheitlichen Haplotyp noch eine signifikante Häufung von seltenen SNP- und SNV-Allelen der CH-Kohorte in der PER des *NKX2-1*-Gens, die eine Beteiligung einer SNP bedingten Transkriptionsfaktorbindungsstörung an der Pathogenese der isolierten Schilddrüsendysgenese wahrscheinlich machen würde (siehe Ergebnisse 3.1.3). Auch ein Zusammenhang der Schwere der phänotypischen Ausprägung der CH mit der Allelausprägung des SNPs rs116909374 konnte nicht gezeigt werden (siehe Tabelle 3-1). Ein möglicher Erklärungsansatz hierfür ist, dass der SNP rs116909374 zu einem von zwei SNPs in der Zielregion mit einer MAF > 1% gehört, der nach der Theorie von Wang et al. 2018 nur geringen Einfluss auf regulative Prozesse in der PER hätte. Andererseits wurde aber eine Assoziation des SNPs rs116909374 mit dem TSH-Wert nachgewiesen [41, 42].

Eine mögliche Erklärung für die Assoziation des T-Allels des SNPs rs116909374 mit einem niedrigen TSH-Wert könnte eine verminderte Bindung von inhibitorischen Proteinen wie z.B. CTCF an das T-Allel sein. Dadurch könnte es zu einer Aktivitätssteigerung des postulierten schilddrüsenspezifischen Enhancers kommen, was eine erhöhte Schilddrüsenhormonproduktion nach sich ziehen könnte. Eine Bindung von

CTCF an die PER geht aus den Genome Browser Daten allerdings nicht hervor (Abbildung 3-1). Es könnte sich aber um andere inhibitorische, bisher nicht beschriebene Faktoren handeln, die Einfluss auf regulative Prozesse in der Zielregion nehmen [106].

Weiter könnte das Fehlen eines einheitlichen Haplotyps in der CH-Kohorte dadurch bedingt sein, dass es sich nicht nur um einen einzigen krankheitsauslösenden SNP handelt. Viel wahrscheinlicher ist es, wie bereits durch Thorwarth et al. 2014 beschrieben, dass es sich bei der Schilddrüsendysgenese um eine heterogene molekulare Ursache handelt. So könnten eine Reihe verschiedener SNPs und SNVs und vor allem auch eine Kombinationen dieser in der Lage sein, die Transkriptionsfaktorbindung und damit die postulierte Enhancerfunktion zu stören.

Eine weitere Ursache für das Fehlen einer einheitlichen SNP- und SNV-Verteilung in der CH-Kohorte, die zu einer möglichen schilddrüsenspezifischen Enhancerstörung führen könnte, ist die Auswahl der 800 Basenpaare großen potentiellen Enhancer Region (PER). Die Auswahl der Region beruht auf der Arbeit von Thorwarth et al. 2014 und weiteren typischen DNA-Markern für regulative Funktion, die bereits in der Einleitung ausführlich besprochen wurden (1.3.2). Das Fehlen einer einheitlichen SNP- und SNV-Verteilung in der CH-Kohorte oder eine größere Anreicherung dieser könnte einerseits dadurch bedingt sein, dass mit der analysierten Region nicht der gesamte für den postulierten Enhancer funktionell relevante Teil der DNA abgebildet wird, andererseits könnte es auch weiter entfernt liegende DNA-Abschnitte geben, die an der Regulation dieses Enhancerelementes beteiligt sind, und deren Störung zu einem Ausfall des vermuteten Enhancerelementes führen.

Zur definitiven Klärung einer Enhancerfunktion einerseits und der potentiellen Störung der Transkriptionsfaktorbindung durch die neuen Varianten und Minor Allele in der CH-Kohorte andererseits sind in vitro Experimente in Schilddrüsenzellen nötig. Für die Bestätigung der postulierten Enhancerfunktion wäre eine funktionelle Charakterisierung in gewebespezifischen Schilddrüsenzellen mittels eines Luciferase Reporterassays durch Klonierung der gesamten PER in einen entsprechenden Vektor notwendig. Um die Relevanz des neuen SNVs und der Minor Allele der SNPs und SNVs für die mögliche Enhancerfunktion zu prüfen, müssten die Varianten mit der gesamten PER in einen entsprechenden Vektor kloniert werden, um anschließend Unterschiede in der

Luciferase-Expression und damit letztendlich in der Transkriptionsrate zwischen den Vektoren mit den jeweiligen Variaten und ohne den jeweiligen Varianten zu detektieren [107, 108].

Um Aufschluss über eine mögliche Interaktionsveränderung der DNA durch die heterozygoten Minor Allele der SNPs und SNVs sowie des neu gefundenen SNVs mit den in der analysierten Region bindenden Transkriptionsfaktoren zu detektieren und deren definitive Bindung an die PER nachzuweisen, sind die Untersuchung spezifischer DNA-Protein-Interaktionen mittels Chromatin-Immunoprecipitation (CHIP) oder mittels Electrophoresis mobility shift assay (EMSA) nötig [106, 109].

Zusammenfassend konnten in der genetischen Analyse der potentiellen Enhancer Region (PER) des *NKX2-1*-Gens als möglicher Hinweis für eine funktionelle Relevanz in Bezug auf die Entstehung der Schilddrüsendysgenese einzig in der CH-Kohorte SNVs, darunter auch ein neuer SNVs, gefunden werden. Allerdings könnten diese Befunde durch die ungleiche Kohortengröße mit einer deutlich größeren CH-Kohorte beeinflusst oder sogar bedingt sein. Eine Reanalyse der PER in insgesamt größeren und gleich großen Kohorten ist daher notwendig. Auch für den Nachweis der postulierten Enhancerfunktion einerseits sowie einer funktionellen Relevanz der SNPs und SNVs sowie der neuen Variante andererseits sind weitere Experimente notwendig.

4.2 Epigenetische Methylierungsanalyse der potentiellen Enhancer Region

Die Methylierung der DNA erfolgt im Allgemeinen an CpG-Positionen, diese befinden sich vorzugweise in sogenannten CpG-Inseln, welche mindestens 200 Basenpaare groß sind und einen CG-Gehalt von mindestens 50% aufweisen. Über die Hälfte der humanen Promotoren sind mit CpG-Inseln assoziiert, wobei diese überwiegend unmethyliert vorliegen. Ein Großteil der gewebespezifischen DNA-Methylierung tritt im Gegensatz dazu an sogenannten CpG-Inselgrenzen (CpG island shore) auf, welche in der nächsten Umgebung von CpG-Inseln liegen und einen geringeren CG-Gehalt aufweisen [110]. Betrachtet man die PER des *NKX2-1*-Gens und ihre nächste Umgebung (siehe Abbildung 3-3), so scheint diese weder eine CpG-Insel noch eine CpG-Inselgrenze zu sein. Allerdings werden regulative Einflüsse der DNA-Methylierung auch außerhalb von CpG-Inseln und CpG-Inselgrenzen beschrieben. So sind z.B. gewebespezifische Gene meist nicht mit einer CpG-Insel assoziiert [51]. Weiter scheint die allelspezifische DNA-

Methylierung mit SNP-Haplotypen assoziiert zu sein, wobei hierbei auch vor allem krankheitsassoziierte SNPs einen möglichen Einfluss auf die DNA-Methylierung haben [111]. Aufgrund des in der PER des *NKX2-1*-Gens gelegenen, mit dem TSH-Wert assoziierten SNP rs116909374, lag daher die Quantifizierung der Methylierungsintensität auch ohne Vorhandensein einer CpG-Insel oder CpG-Inselgrenze nahe.

4.2.1 Gewebespezifität der Methylierung der potentiellen Enhancer Region

DNA-Methylierungsmuster können einerseits ubiquitär im Organismus auftreten, andererseits wird eine Gewebe- und Zelltypspezifität von Methylierungsmustern beschrieben [86, 112]. Gewebespezifische Gene liegen hierbei in den das Gen exprimierenden Zellen unmethyliert und in allen anderen Zellen methyliert vor [51].

Es wird angenommen, dass diese als “Tissue-specific differentially methylated regions” (T-DMRs) bezeichneten DNA-Regionen eine entscheidende Rolle bei der Regulation der gewebespezifischen Transkription von Genen spielen. Weiter konnten T-DMRs in orthologen Sequenzen bei Mäusen und Menschen nachgewiesen werden, was für eine evolutionäre Konservierung dieser Regionen und damit für ihre funktionelle Relevanz spricht [58].

Die in dieser Studie gefundene ausgeprägte Hypomethylierung der PER im funktionell relevanten Schilddrüsengewebe im Verhältnis zur DNA aus peripheren Blutzellen (siehe Abbildung 3-5) spricht für das Vorhandensein einer T-DMR in diesem Bereich der DNA. Viele Studien zeigen Evidenz dafür, dass die Hypomethylierung der DNA, wie sie schilddrüsenspezifisch in der analysierten Region nachgewiesen wurde, an der Entstehung, Aktivität und Aufrechterhaltung von Enhancerelementen beteiligt ist. Hingegen scheint ein hoher Anteil an methylierten CpG-Positionen typischerweise mit einer Transkriptionshemmung assoziiert zu sein [47, 113].

Es konnte gezeigt werden, dass T-DMRs, die vorzugsweise in CpG-Inseln zu finden sind, typischerweise mehrere Kilobasen zum nächsten bekannten Gen entfernt liegen. Es wird daher vermutet, dass T-DMRs eher die Zugänglichkeit von Transkriptionsfaktoren zu Enhancerelementen beeinflussen als an der direkten DNA-Methylierung von CpG-Positionen in der Umgebung von Transkriptionsstartpunkten beteiligt zu sein [58]. Auch für die PER wurde bereits in der genetischen Analyse eine gestörte Transkriptionsfaktorbindung als mögliche Ursache der Enhancerstörung diskutiert (siehe 3.1.3). Die gewebespezifische Hypomethylierung der analysierten Region im für die kongenitale Hypothyreose relevanten Schilddrüsengewebe ist Grund für die Annahme

einer T-DMR in der PER. Die oben vermutete Funktion von T-DMRs bezüglich der Zugänglichkeit von Transkriptionsfaktoren zu Enhancerelementen passt einerseits zu der postulierten schilddrüsen-spezifischen Enhancerfunktion der analysierten Region und ist andererseits konsistent mit dem bereits in der genetischen Analyse diskutierten Mechanismus einer Transkriptionsfaktorbindungsstörung als Ursache der gestörten Enhancerfunktion (siehe 3.1.3).

Die gewebespezifische Hypomethylierung der PER im für die kongenitale Hypothyreose relevanten Schilddrüsengewebe ist daher ein deutlicher Hinweis für die postulierte schilddrüsen-spezifische regulatorische Funktion.

Ein allgegenwärtiges Problem von großen epigenetischen Methylierungsanalysen sind aber gerade T-DMRs wie die der potentiellen Enhancer Region (PER). Ihnen wird einerseits ein großer regulativer Einfluss zugeschrieben, andererseits führt die Gewebespezifität der Methylierung dazu, dass die gewöhnlich aufgrund der relativ leichten Verfügbarkeit verwendete DNA aus peripheren Blutzellen (PBL) in ihrer Aussagekraft bezüglich der qualitativen und quantitativen Methylierungsintensität limitiert ist. Dies unterstützen auch zwei Befunde aus unterschiedlichen Bereichen: Beispielsweise zeigte ein Vergleich des Methylierungsmusters aus verschiedenen Etagen des Respirationstraktes und peripheren Blutzellen (PBL) in Bezug auf Asthmastudien, dass sich mit abnehmender Gewebenähe zum unteren Respirationstrakt das Methylierungsmuster verändert [112]. Die größten Methylierungsunterschiede ergaben sich zu PBL. Resümierend wurde die Verwendung von DNA aus PBL für die Methylierungsanalyse von Erkrankungen anderer Gewebe in Frage gestellt [112]. Zu einem ähnlichen Ergebnis kommen Weedon et al. 2014 bei der Entdeckung eines für die Pankreasagenese relevanten pankreasspezifischen Enhancers. Dabei zeigten sich einzig in Pankreasvorläuferzellen entsprechende DNA-Modifikationen als Hinweis auf eine Enhancerfunktion. In über 100 Zellreihen, darunter auch differenzierten exokrinen Pankreaszellen, fehlten diese Modifikationen [21].

Theoretisch denkbare Alternativen zur Methylierungsanalyse aus peripheren Blutzellen (PBL) stellen daher z.B. gewebespezifische Methylierungsanalysen im Mausmodell dar, da die T-DMRs bei Mäusen und Menschen an orthologen Positionen gefunden werden

[58]. Rückschlüsse aus tierexperimentellen Studien auf humane DNA regulative Prozesse sind allerdings nur eingeschränkt möglich.

Bezüglich der *NKX2-1*-Mutationen z.B. wurden phänotypische Unterschiede zwischen Maus und Mensch beschrieben. Im Menschen sind bisher nur heterozygote *NKX2-1*-Mutationen beschrieben, welche mit einem krankheitsspezifischen Phänotyp einhergehen, wohingegen heterozygote Mäuse keinen krankheitsspezifischen Phänotyp aufweisen [18, 31].

Die Ergebnisse von Weedon et al. 2014 zeigen, dass der pankreasspezifische Enhancer lediglich während der Pankreasentwicklung in pankreasspezifischen Vorläuferzellen aktiv ist und nicht in adulten exokrinen Pankreaszellen. Somit ist neben der Gewebespezifität auch der Zeitpunkt von regulativen Prozessen in der embryonalen Entwicklung relevant. Damit wären nicht nur die Methylierungsanalyse in humanem Schilddrüsengewebe, sondern auch das Wissen über den Zeitpunkt der Etablierung des Methylierungsmuster in der PER des *NKX2-1*-Gens sehr aufschlussreich.

Hierfür müsste das Methylierungsmuster der PER zu verschiedenen Zeitpunkten während der Embryonalentwicklung betrachtet werden, um zu analysieren, wie sich das Methylierungsmuster innerhalb eines Keimblattes verhält, um letztendlich zu ermitteln, zu welchem Differenzierungszeitpunkt das Methylierungsmuster gesetzt wird. Unter der aktuellen deutschen Rechtslage ist die Analyse der entwicklungspezifischen Methylierungsveränderungen von humanem Gewebe nicht möglich.

Eine Möglichkeit, gewebe- und entwicklungspezifische regulative Prozesse in humanem Gewebe dennoch untersuchen zu können, ist allerdings für das tiefere Verständnis der Pathogenese von Krankheiten dringend erforderlich. Eine vielversprechende Möglichkeit, humane, gewebe- und entwicklungspezifische Pathomechanismen von Krankheiten zu untersuchen, scheinen dabei zellkulturelle Experimente mit humanen induzierten pluripotenten Stammzellen (iPSCs) zu sein [114].

Unter der Annahme möglicher gewebeunabhängiger kohortenspezifischer Methylierungsunterschiede erfolgte aufgrund der oben diskutierten allgegenwärtigen Problematik bezüglich der Verfügbarkeit von entsprechendem funktionell relevantem Gewebe für große Kohortenstudien, trotz der genannten Limitationen, die kohortenspezifische Methylierungsanalyse in der vorliegenden Arbeit aus peripheren

Blutzellen (PBL). Speziell bei dieser Arbeit ist zu berücksichtigen, dass bei einer gewebespezifischen Methylierungsanalyse im funktionell relevanten Schilddrüsengewebe die Patientengruppe mit einer phänotypischen Athyreose aufgrund des Fehlens von Schilddrüsengewebe ausgeschlossen wäre.

4.2.2 Abhängigkeit der Methylierung von der Zelllinie

Die DNA-Methylierung zeigt nicht nur gewebespezifische Unterschiede, sondern es besteht auch eine Abhängigkeit von der Zelllinie [115]. Damit ist die Methylierung des häufig verwendeten peripheren Blutes, wie auch in der vorliegenden Arbeit, abhängig von der Leukozytenzusammensetzung [116]. Eine Abhängigkeit der Methylierungsergebnisse von der Leukozytenzusammensetzung konnte aufgrund der fehlenden Verfügbarkeit des Materials nicht erfolgen. Somit konnte die Berücksichtigung eines möglichen zelllinienspezifischen Einflusses auf die Ergebnisse nicht erfolgen. Allerdings scheinen die Methylierungsunterschiede zwischen den Leukozytensubtypen kleiner zu sein als vermutet [98]. Joubert et al. 2012 konnten zum Beispiel zeigen, dass der Methylierungsunterschied zwischen mononukleären Zellen und polymorphnukleären Zellen aus Nabelschnurblut klein war [117].

4.2.3 Methylierungsintensität der PER im Kohortenvergleich

Eine Hypermethylierung ist epigenetisch in der Regel mit einer Hemmung der Expression eines Zielgens verbunden [113]. Futscher et al. 2002 z.B. konnten zeigen, dass die zelltypspezifische Methylierungsintensität des SERPINB5-Promotors für die normale zelltypspezifische Expression des entsprechenden Gens verantwortlich ist. Dabei zeigten Zelltypen mit einer Hypermethylierung des Promotors keine SERPINB5-Expression [118].

In der vorliegenden Arbeit wurde daher die Methylierungsintensität der PER, für die eine regulative Funktion in Form eines schilddrüsenspezifischen Enhancers postuliert wurde, zwischen Patienten mit Schilddrüsendysgenese, Patienten mit Hyperthyreotropinämie, einer milden Störung der Schilddrüsenfunktion mit einer isolierten TSH-Erhöhung und gesunden Patienten verglichen. Die Ergebnisse der vorliegenden Arbeit zeigen eine statistisch signifikante Hypermethylierung der CH-Kohorte in der potentiellen Enhancer Region (PER) im Verhältnis zu den beiden anderen analysierten Kohorten sowohl im Mittelwert der Methylierung über die neun CpG-Positionen (siehe Abbildung 3-6) als auch

an einem Großteil der einzelnen CpG-Positionen (siehe Abbildung 3-7). Mit einer maximalen Methylierungsdifferenz im Mittelwert der Methylierung von ca. vier Prozent zwischen den gesunden Kontrollen und der CH-Kohorte ist eine funktionelle Relevanz der gefundenen Ergebnisse fraglich. Auch die bereits kritisch diskutierte Verwendung von peripheren Blutzellen (PBL) für die Analyse möglicher methylierungsbedingter schilddrüsenspezifischer funktioneller Unterschiede ist bei der Interpretation der Ergebnisse zu beachten. Denkbar ist, dass die statistisch signifikanten kohortenspezifischen Methylierungsunterschiede in peripheren Blutzellen (PBL) aufgrund der Beschreibung der PER als T-DMR im funktionell relevanten Schilddrüsengewebe oder in Schilddrüsenvorläuferzellen noch deutlich ausgeprägter sind als in den zur Methylierungsanalyse verwendeten nicht funktionell relevanten peripheren Blutzellen (PBL).

Allerdings muss kritisch hinterfragt werden, ob bei einer derartig geringen Methylierungsdifferenz im peripheren Blut ein Zusammenhang zur Methylierung im Schilddrüsengewebe anzunehmen ist.

Diskordante monozygote Zwillinge sind aufgrund ihres identischen Genoms eine wertvolle Resource bei epigenetischen Methylierungsanalysen und unter anderem der Grund für die Annahme von Epimutationen bei der Entstehung der kongenitalen Hypothyreose. Bei der Analyse von sechs für die isolierte Schilddrüsendysgenese diskordanten monozygoten Zwillingspaaren ergaben sich keine relevanten Methylierungsunterschiede in der PER des *NKX2-1*-Gens zwischen den gesunden und erkrankten Zwillingen. Dieser Befund macht eine funktionelle Relevanz der in der Kohortenanalyse gefundenen Methylierungsdifferenz unwahrscheinlich. Allerdings ist die Aussagekraft der Ergebnisse aufgrund der kleinen Fallzahl eingeschränkt (siehe 3.2.4).

Um die gefundene signifikante Hypermethylierung der CH-Kohorte in der Kohortenanalyse in Assoziation mit einer krankheitsrelevanten Schilddrüsenstörung sehen zu können, müsste eine erneute Analyse der Methylierung im peripheren Blut in deutlich größeren und vor allem auch gleich großen Kohorten erfolgen. Auch die Reanalyse der Methylierungsdifferenz in einer deutlich größeren Kohorte an diskordanten monozygoten Zwillingspaaren sowie in einer großen Kohortenstudie mit Schilddrüsengewebe von gesunden Kontrollen und CH-Patienten wäre für die Aussagekraft und Interpretation der Ergebnisse wichtig.

Wie auch in der vorliegenden Arbeit sind die statistisch signifikanten Methylierungsunterschiede oft klein und die biologische Bedeutung dieser kleinen Methylierungsunterschiede, unter anderem für die Genexpression- und Regulation, bisher weitgehend unbekannt [98]. Ein weitergehendes Verständnis der funktionellen Prozesse ist daher für die Einordnung und Interpretation der Ergebnisse aus Methylierungsanalysen im Allgemeinen von essentieller Bedeutung.

4.2.4 Altersabhängige Methylierungsveränderungen

Neben den bereits in der Einleitung beschriebenen Methylierungsveränderungen während der Keimzell- und embryonalen Entwicklung konnten auch postnatale altersabhängige Methylierungsveränderungen nachgewiesen werden [119]. Es konnte gezeigt werden, dass es in der späten Kindheit und Adoleszenz zu einem raschen Anstieg der genomweiten Methylierung kommt, die anschließend im Erwachsenenalter stabil bleibt und mit zunehmendem Alter wieder abnimmt [119]. Die dynamische globale Veränderung der Methylierung scheint dabei jeweils spezifisch unterschiedliche, regulative Elemente zu betreffen. Die Abnahme der Methylierungsintensität im höheren Alter z.B. scheint vorwiegend vor allem intergenisch nicht CpG-Insel assoziiert aufzutreten, wohingegen an promotorassoziierten CpG-Inseln eher eine Zunahme der Methylierung zu beobachten ist. Da die meisten CpG-Positionen allerdings außerhalb von CpG-Inseln liegen, kommt es genomweit betrachtet zu einem Methylierungsverlust mit zunehmendem Alter [119].

Die altersabhängigen Methylierungsveränderungen werden dabei zwei epigenetischen Phänomenen zugeordnet. Die epigenetische Drift beschreibt eine mit dem Alter zunehmende Variabilität der Methylierung zwischen verschiedenen Individuen, die sowohl durch stochastische Ereignisse während der Zellteilung als auch durch Umwelteinflüsse erworben ist [119]. Fraga et al. 2005 konnten z.B. zeigen, dass sich das Methylom von monozygoten Zwillingen mit zunehmendem Alter stärker unterscheidet [120]. Die epigenetische Uhr hingegen beschreibt spezifische Gen-Loci, deren altersabhängige Methylierungsveränderungen interindividuell stabil auftreten und mit dem chronologischen Alter zu korrelieren scheinen [119].

Die Möglichkeit von altersabhängigen Methylierungsunterschieden sollte daher bei der Auswertung und Interpretation von Methylierungsanalysen beachtet werden. In der vorliegenden Arbeit variiert das mittlere Alter der Kohorten (siehe Tabelle 2-4); allerdings

zeigte sich keine Abhängigkeit der Methylierung vom Alter der Personen (siehe Abbildung 3-18).

4.2.5 Zusammenhang der Methylierung mit der Schilddrüsenfunktion

Die Analyse der CpG-positionsspezifischen Methylierungsunterschiede der Kohorten zeigt an fünf der neun CpG-Positionen eine nicht statistisch signifikante Methylierungsverteilung mit der höchsten Methylierung der CH-Kohorte und der niedrigsten Methylierung der gesunden Kontrollen (siehe Abbildung 3-7). Diese Methylierungsverteilung würde theoretisch zu einer graduellen Störung der potentiellen Enhancerfunktion durch die quantitative Methylierungshöhe passen (siehe Abbildung 4-1).



Abbildung 4-1: Zusammenhang Methylierungsverteilung und Schilddrüsenfunktion

Theoretisch denkbarer Zusammenhang zwischen der Methylierungsverteilung der Kohorten in der potentiellen Enhancer Region und dem Grad der Schilddrüsenfunktion.

Die Daten dieser Arbeit zeigen allerdings keinen Hinweis für einen derartigen Zusammenhang zwischen der quantitativen Methylierung der PER des *NKX2-1*-Gens und der Schilddrüsenfunktion. Die Möglichkeit einer graduellen Funktionsstörung von regulativen genomischen Prozessen bei der Suche nach der molekularen Ursache phänotypisch klar abgrenzbarer Krankheitsbilder ist aber denkbar und sollte auch bei weiterführenden Analysen auf der Suche der molekularen Ursache der isolierten Schilddrüsendysgenese in Betracht gezogen werden.

4.2.6 CpG-positionsspezifische Methylierungsunterschiede

Eine Hypomethylierung von CpG-Positionen ist im Allgemeinen mit aktiver regulativer Funktion von genomischen Regionen verbunden [113]. Ein Vergleich der kohortenübergreifenden CpG-positionsspezifischen Methylierung der PER ergab eine signifikante Hypomethylierung der den mit dem TSH-Wert korrelierenden SNP

rs116909374 nahen CpG-Positionen im Vergleich zu weiter entfernt liegenden CpG-Positionen (siehe Abbildung 3-8). Auch die Analyse der diskordanten monozygoten Zwillinge zeigt vergleichbare Ergebnisse (siehe 3.2.4.3). Weiter zeigte sich an der dem SNP rs116909374 nahen CpG-Position drei auch die größte Methylierungsdifferenz zwischen gesunden Kontrollen und CH-Patienten (siehe Tabelle 3-2). Die Methylierungsdifferenzen zwischen den gesunden und an CH erkrankten Zwillingen war sogar an beiden dem SNP rs116909374 benachbarten CpG-Positionen am größten. Konsistent mit den Ergebnissen der Kohortenanalyse zeigte die Position drei dabei den größten Methylierungsunterschied (siehe Tabelle 3-3). Die Konsistenz der Ergebnisse zwischen der Kohortenanalyse und der Analyse der für die CH diskordanten monozygoten Zwillingspaare stärkt die Aussagekraft der Ergebnisse, auch wenn die absoluten Methylierungsunterschiede sehr gering sind.

In Zusammenschau sind die Ergebnisse somit hinweisend dafür, dass insbesondere der mit dem TSH-Wert korrelierende SNP rs116909374 und seine unmittelbare Umgebung eine besondere Relevanz bezüglich einer regulativen Funktion der PER spielen könnten. Diese Annahme wird durch die vorausgehenden unveröffentlichten 4C-Experimente des IEPES von Thorwarth et al. im Mausmodell gestützt, die eine Interaktion des SNPs rs116909374 und seiner unmittelbaren Umgebung mit dem *NKX2-1*- und dem *Mbip*-Gen zeigen konnten (siehe Abbildung 1-5). Weiter hat die Analyse der PER im UCSC Genome Browser (siehe Abbildung 3-1) gezeigt, dass in diesem Bereich ein Großteil der an der PER bindenden Transkriptionsfaktoren zu finden sind, darunter auch FOXA1, dessen Expression in C-Zellen der Schilddrüse sowie im medullären Schilddrüsenkarzinom nachgewiesen werden konnte, jedoch nicht in Follikelzellen der Schilddrüse. Allerdings ist die Expressionsveränderung von FOXA1 in nicht neoplastischen Schilddrüsenerkrankungen bisher nicht ausreichend untersucht [94]. Weiter waren in diesem Bereich der PER neben dem SNP rs116909374 zwei heterozygote SNPs auf den CpG-Positionen eins und vier ausschließlich in der CH-Kohorte zu finden (siehe Abbildung 3-2). Allerdings muss hierbei beachtet werden, dass dieser Befund durch die ungleiche Kohortengröße mit einer deutlich größeren CH-Kohorte beeinflusst sein könnte.

In Zusammenschau der Befunde könnte die kohortenübergreifende Hypomethylierung der CpG-Positionen eins und drei, die den mit dem TSH-Wert korrelierenden SNP rs116909374 umgeben, und die an der CpG-Position drei gefundene größte

Methylierungsdifferenz zwischen gesunden Kontrollen und CH-Patienten Ausdruck einer funktionellen Relevanz speziell dieser CpG-Positionen sein.

4.2.7 DNA-sequenzspezifische Methylierungsunterschiede

CpG-Positionen in Säugetieren beeinflussen die Genregulation abhängig von ihrem Methylierungsstatus [121]. Vier der fünf in der CH-Kohorte gefundenen SNPs liegen auf CpG-Positionen (siehe Abbildung 3-2). Diese SNPs werden als CpG-SNPs bezeichnet. Ihre funktionelle Rolle bei der Genexpression besteht im Allgemeinen unabhängig von ihrer Lokalisation im Genom in der Beeinflussung der Bindungsaffinität und Spezifität von spezifischen Faktoren wie z.B. Transkriptionsfaktoren in Abhängigkeit ihres Methylierungsstatus [64, 121]. Ein enges Zusammenspiel zwischen spezifischen Transkriptionsfaktoren und lokalen DNA-Methylierungsmustern wird in vielen Studien beschrieben. Ein Beispiel für die enge funktionelle Beziehung zwischen DNA-Methylierung und Transkriptionsfaktoren ist der viel untersuchte methylierungssensitive Transkriptionsfaktor CTCF, der mit seiner Zinkfingerbindungsdomäne spezifisch an unmethylierte DNA binden kann und so als inhibitorisches Protein unter anderem Enhancer-Promotor-Interaktionen hemmt sowie an der Ausbildung von DNA-Interaktionen über weite Distanzen hinweg durch eine als „DNA-Loop“ bezeichnete Schleifenbildung der DNA beteiligt ist [102].

Bei den drei in dieser Arbeit nachgewiesenen CpG-SNPs sowie dem neu gefundenen SNV kommt es durch das heterozygote Auftreten des Minor Allels zu einer Zerstörung der CpG-Position und somit zu einer Hemimethylierung der CpG-Position. Drei der CpG-SNPs, darunter auch der SNP rs116909374, dessen T-Allel mit einem erniedrigten TSH-Wert assoziiert ist, befinden sich in enger räumlicher Nähe in einem Bereich der PER, an dem verschiedene Transkriptionsfaktoren binden. Neben dem SNP rs116909374, der in jeder der analysierten Kohorte heterozygot auftrat, waren die weiteren CpG-SNPs einzig in der CH-Kohorte zu finden (siehe Abbildung 3-2). Nach den oben beschriebenen allgemeinen Eigenschaften von CpG-SNPs und der engen Beziehung zwischen DNA-Methylierung und Transkriptionsfaktoren liegt es daher nahe, dass die drei unmittelbar benachbarten CpG-SNPs zu einer Beeinflussung der Transkriptionsfaktorbindung führen könnten. Der CH-Patient, der heterozygoter Träger von zwei der unmittelbar benachbart liegenden CpG-SNPs ist (siehe Tabelle 3-1), zeigt phänotypisch eine Athyreose, die schwerste Form der Schilddrüsendysgenese. Unter der Bedingung, dass in einer Reanalyse der SNP-Verteilung in deutlich größeren Kohorten ein vermehrtes

Auftreten von SNPs in der CH-Kohorte nachgewiesen würde, wäre ein Zusammenhang dieser CpG-SNPs mit der Schilddrüsenfunktion denkbar.

4.2.8 Geschlechtsspezifische Methylierungsunterschiede

Mädchen sind zwei bis viermal häufiger von der CH betroffen als Jungen. Die Ursache der Mädchenwendigkeit der Erkrankung ist unbekannt [13, 20]. Eine Beteiligung von epigenetischen Modifikationen bei der Entstehung der CH wird nicht zuletzt wegen der Mädchenwendigkeit der Erkrankung diskutiert [18]. De Felice et al. 2004 postulierten einen DNA-Methylierungsdefekt als einen möglichen Auslöser der Schilddrüsendysgenese [23]. Narumi et al. 2018 konnten in einer daraufhin durchgeführten genomweiten Methylierungsanalyse sowie in der Methylierungsanalyse der an der Schilddrüsenentwicklung beteiligten Transkriptionsfaktor kodierenden Gene *NKX2-1*, *FOXE1* und *PAX-8* keine Methylierungsunterschiede von Patienten mit Schilddrüsenektomie, bei denen die Mädchenwendigkeit besonders ausgeprägt auftritt, und gesunden Kontrollen finden [122]. Allerdings wurde sich auf der Suche nach der molekularen Ursache der Schilddrüsendysgenese bisher und auch in dieser Studie vor allem auf die Analyse kodierender Genbereiche beschränkt. Mögliche Mutationen in Introns oder nicht-kodierenden regulativen Regionen blieben so bisher unbeachtet [23].

In dieser Arbeit wurde hingegen eine konservierte nicht-kodierende Region, für die eine regulative Funktion in Form einer schilddrüsen-spezifischen Enhancerfunktion postuliert wird, analysiert. Die Ergebnisse der geschlechtsspezifischen Methylierungsanalyse zeigten eine statistisch signifikante Hypermethylierung der Frauen mit isolierter Schilddrüsendysgenese im Verhältnis zu den Frauen der beiden anderen Kohorten. Für die beiden anderen Kohorten und die Männer zeigten sich keine signifikanten Methylierungsunterschiede (siehe Tabelle 3-6). Die signifikante Hypermethylierung der PER bei Frauen mit isolierter Schilddrüsendysgenese könnte somit eine funktionelle Relevanz für die Entstehung der Erkrankung haben, aber vor allem auch an der Mädchenwendigkeit der Erkrankung beteiligt sein. Einschränkend ist anzumerken, dass vor allem in der CH-Kohorte der Frauenanteil nicht zuletzt wegen der Mädchenwendigkeit der Erkrankung deutlich größer war und die Ergebnisse dadurch beeinflusst sein könnten. Eine erneute Methylierungsanalyse mit homogeneren Kohorten, insbesondere bezüglich der Patientenzahl und der Geschlechtsverteilung, wäre daher sinnvoll.

4.3 Zusammenfassung und Ausblick

Die molekulare Ursache der Schilddrüsendysgenese ist trotz intensiver Forschung in der Mehrzahl der Fälle weiter unbekannt. Der Schwerpunkt von vielen Studien bezüglich der molekularen Ursache liegt vor allem in der Analyse von kodierenden Genbereichen, die mit der Schilddrüsenentwicklung oder Funktion in Zusammenhang gebracht werden konnten. Eine Aufklärung der molekularen Ursache der Schilddrüsendysgenese für einen Großteil der Patienten gelang dadurch bisher nicht. Andere molekulare Mechanismen, wie z.B. die Beteiligung von regulativen nicht-kodierenden Genbereichen sowie epigenetischen Veränderungen als mögliche Ursache, werden daher in Erwägung gezogen.

In der vorliegenden Arbeit wurde daher ein konservierter nicht-kodierender Bereich des *NKX2-1*-Gens, für den eine schilddrüsenspezifische Enhancerfunktion postuliert wurde, auf mögliche genetische oder epigenetische Störungen untersucht, die als mögliche molekulare Ursache der Schilddrüsendysgenese in Frage kommen könnten.

In der genetischen Analyse konnten SNPs und SNVs gefunden werden, die einzig in der CH-Kohorte aufgetreten sind, darunter auch ein neuer SNV. Allerdings könnten diese Ergebnisse durch die ungleiche Kohortengröße und fehlende Betrachtung ethnischer Unterschiede bedingt sein. Eine mutationsbedingte Störung der potentiellen Enhancer Region (PER) als Ursache der Schilddrüsendysgenese scheint daher unwahrscheinlich, wenn auch eine Mitbeteiligung dieser Varianten an der Entstehung der CH nicht ausgeschlossen ist. Weitere Untersuchungen der kohortenspezifischen SNP-Verteilung in der PER des *NKX2-1*-Gens in deutlich größeren, und bezüglich der Anzahl und des ethnischen Hintergrundes homogenen Kohorten, wären für die Einordnung der vorliegenden Ergebnisse zwingend notwendig.

Eine mögliche Enhancerfunktion der analysierten Region wurde aufgrund von Chromatinstrukturen und vorausgehenden Studien postuliert. Die funktionelle Überprüfung dieser Annahme erfolgte in dieser Arbeit nicht, wäre aber für die Interpretation und Bewertung der gefundenen Ergebnisse von großer Bedeutung. Eine funktionelle Charakterisierung der PER, z.B. mittels Luciferase-Reporterassays, würde Aufschluss über eine funktionelle Relevanz z.B. im Sinne des postulierten Enhancers liefern. Im zweiten Schritt könnten dann die in der CH-Kohorte gefundenen SNP-Varianten und vor allem auch die neu entdeckte Variante auf eine mögliche Modulierung der Enhancerfunktion getestet werden.

In der epigenetischen Methylierungsanalyse konnte eine, wenn auch gering ausgeprägte, statistisch signifikante Hypermethylierung der CH-Patienten in der PER des *NKX2-1*-Gens gezeigt werden, was theoretisch mit einer Inaktivierung des postulierten schilddrüsen-spezifischen Enhancers und einer dadurch bedingten Schilddrüsenfunktionsstörung vereinbar wäre. Durch die ausgeprägte Gewebespezifität der untersuchten Region ist ein Rückschluss dieser geringen Methylierungsunterschiede der DNA aus peripheren Blutzellen auf eine mögliche schilddrüsen-spezifische funktionelle Relevanz nicht möglich.

Weiter zeigte sich in der geschlechtsspezifischen Methylierungsanalyse eine statistisch signifikante Hypermethylierung der Frauen mit isolierter Schilddrüsendysgenese. In der Gruppe der Männer gab es diesen statistisch signifikanten Methylierungsunterschied nicht. Dieser Befund ist mit der Mädchenwendigkeit der Erkrankung gut vereinbar. Rückschlüsse dieser Ergebnisse auf einen möglichen krankheitsspezifischen Zusammenhang sind einerseits wegen der Gewebespezifität der untersuchten Region und andererseits wegen der ungleichen Geschlechterverteilung der Kohorten schwierig. Weitere, vor allem gewebespezifische Methylierungsanalysen sowie eine homogene Geschlechterverteilung der Kohorten sind nötig, um einen möglichen geschlechtsspezifischen Methylierungsunterschied der CH-Patienten in der PER des *NKX2-1*-Gens nachzuweisen.

Sollten sich die gefundenen Methylierungsunterschiede in Reanalysen bestätigen, sind weitere Studien notwendig, um zu untersuchen, ob die gefundenen Methylierungsunterschiede tatsächlich einen Einfluss auf regulatorische Prozesse in der analysierten PER des *NKX2-1*-Gens haben.

Ein Vergleich von DNA-Interaktionen zwischen gesunden Kontrollen und den in der PER hypermethylierten CH-Patienten, z.B. mittels 4C-Experimenten, wäre sinnvoll. Weiter ist eine mögliche veränderte Transkriptionsfaktorbindung in der PER durch die Hypermethylierung und die gefundenen SNP-Varianten in CH-Patienten denkbar. Die vergleichende Untersuchung spezifischer DNA-Protein-Interaktionen zwischen gesunden Kontrollen und den analysierten CH-Patienten, z.B. mittels Chromatin-Immunoprecipitation (CHIP), würde Aufschluss über eine relevante Beeinflussung der Transkriptionsfaktorbindung liefern.

Der Vergleich des Methylierungsmusters in der PER zwischen DNA aus peripheren Blutzellen (PBL) und Schilddrüsengewebe gesunder Personen zeigte eine ausgeprägte gewebespezifische Hypomethylierung im funktionell relevanten Schilddrüsengewebe. Aufgrund der Gewebespezifität des Methylierungsmusters und der damit vermutlich einhergehenden regulativen Prozesse wäre eine gewebespezifische Analyse der PER in Schilddrüsengewebe sinnvoll.

Weiter stellt die Verwendung von induzierten pluripotenten Stammzellen (iPCS) einen vielversprechenden Ansatz für die Analyse humaner gewebespezifischer regulativer Prozesse und vor allem auch deren krankheitsspezifischen Störungen dar.

Zusammenfassend konnten die Ergebnisse dieser Arbeit Hinweise für eine mögliche funktionelle Relevanz der potentiellen Enhancer Region in Bezug auf die molekulare Ursache der CH zeigen. Die funktionellen Prozesse und deren mögliche Störung in der potentiellen Enhancer Region bleiben aber weiter unklar und sind in weiteren Studien zu klären.

5 Literaturverzeichnis

- 1. Gruters, A., H. Biebermann, and H. Krude, *Neonatal thyroid disorders*. Horm Res, 2003. **59 Suppl 1**: p. 24-9.
- 2. Rastogi, M.V. and S.H. LaFranchi, *Congenital hypothyroidism*. Orphanet J Rare Dis, 2010. **5**: p. 17.
- 3. Wassner, A.J. and R.S. Brown, *Congenital hypothyroidism: recent advances*. Curr Opin Endocrinol Diabetes Obes, 2015. **22(5)**: p. 407-12.
- 4. Gruters, A. and H. Krude, *Detection and treatment of congenital hypothyroidism*. Nat Rev Endocrinol, 2012. **8(2)**: p. 104-13.
- 5. Gruters, A., A. Jenner, and H. Krude, *Long-term consequences of congenital hypothyroidism in the era of screening programmes*. Best Pract Res Clin Endocrinol Metab, 2002. **16(2)**: p. 369-82.
- 6. Aleksander, P.E., M. Bruckner-Spieler, A.M. Stoehr, E. Lankes, P. Kuhnen, D. Schnabel, A. Ernert, W. Stablein, M.E. Craig, O. Blankenstein, A. Gruters, and H. Krude, *Mean High-Dose L-Thyroxine Treatment Is Efficient and Safe to Achieve a Normal IQ in Young Adult Patients With Congenital Hypothyroidism*. J Clin Endocrinol Metab, 2018. **103(4)**: p. 1459-1469.
- 7. Kuhnen, P., S. Turan, S. Frohler, T. Guran, S. Abali, H. Biebermann, A. Bereket, A. Gruters, W. Chen, and H. Krude, *Identification of PENDRIN (SLC26A4) mutations in patients with congenital hypothyroidism and "apparent" thyroid dysgenesis*. J Clin Endocrinol Metab, 2014. **99(1)**: p. E169-76.
- 8. Hinton, C.F., K.B. Harris, L. Borgfeld, M. Drummond-Borg, R. Eaton, F. Lorey, B.L. Therrell, J. Wallace, and K.A. Pass, *Trends in incidence rates of congenital hypothyroidism related to select demographic factors: data from the United States, California, Massachusetts, New York, and Texas*. Pediatrics, 2010. **125 Suppl 2**: p. S37-47.
- 9. Harris, K.B. and K.A. Pass, *Increase in congenital hypothyroidism in New York State and in the United States*. Mol Genet Metab, 2007. **91(3)**: p. 268-77.
- 10. Abu-Khudir, R., S. Larrivee-Vanier, J.D. Wasserman, and J. Deladoey, *Disorders of thyroid morphogenesis*. Best Pract Res Clin Endocrinol Metab, 2017. **31(2)**: p. 143-159.
- 11. Thorwarth, A., I. Mueller, H. Biebermann, H.H. Ropers, A. Grueters, H. Krude, and R. Ullmann, *Screening chromosomal aberrations by array comparative genomic hybridization in 80 patients with congenital hypothyroidism and thyroid dysgenesis*. J Clin Endocrinol Metab, 2010. **95(7)**: p. 3446-52.
- 12. Castanet, M., M. Polak, C. Bonaiti-Pellie, S. Lyonnet, P. Czernichow, and J. Leger, *Nineteen years of national screening for congenital hypothyroidism: familial cases with thyroid dysgenesis suggest the involvement of genetic factors*. J Clin Endocrinol Metab, 2001. **86(5)**: p. 2009-14.
- 13. Devos, H., C. Rodd, N. Gagne, R. Laframboise, and G. Van Vliet, *A search for the possible molecular mechanisms of thyroid dysgenesis: sex ratios and associated malformations*. J Clin Endocrinol Metab, 1999. **84(7)**: p. 2502-6.
- 14. Perry, R., C. Heinrichs, P. Bourdoux, K. Khoury, F. Szots, J.H. Dussault, G. Vassart, and G. Van Vliet, *Discordance of monozygotic twins for thyroid dysgenesis: implications for screening and for molecular pathophysiology*. J Clin Endocrinol Metab, 2002. **87(9)**: p. 4072-7.
- 15. Magne, F., R. Serpa, G. Van Vliet, M.E. Samuels, and J. Deladoey, *Somatic mutations are not observed by exome sequencing of lymphocyte DNA from monozygotic*

- twins discordant for congenital hypothyroidism due to thyroid dysgenesis.* Horm Res Paediatr, 2015. **83**(2): p. 79-85.
- 16. Polak, M., S. Sura-Trueba, A. Chauty, G. Szinnai, A. Carre, and M. Castanet, *Molecular mechanisms of thyroid dysgenesis.* Horm Res, 2004. **62 Suppl 3**: p. 14-21.
 - 17. Haerlingen, B., R. Opitz, I. Vandernoot, A. Trubiroha, P. Gillotay, N. Giusti, and S. Costagliola, *Small-Molecule Screening in Zebrafish Embryos Identifies Signaling Pathways Regulating Early Thyroid Development.* Thyroid, 2019.
 - 18. Mio, C., G. Grani, C. Durante, and G. Damante, *Molecular defects in thyroid dysgenesis.* Clin Genet, 2020. **97**(1): p. 222-231.
 - 19. Szinnai, G., *Clinical genetics of congenital hypothyroidism.* Endocr Dev, 2014. **26**: p. 60-78.
 - 20. Gruters, A., H. Krude, and H. Biebermann, *Molecular genetic defects in congenital hypothyroidism.* Eur J Endocrinol, 2004. **151 Suppl 3**: p. U39-44.
 - 21. Weedon, M.N., I. Cebola, A.M. Patch, S.E. Flanagan, E. De Franco, R. Caswell, S.A. Rodriguez-Segui, C. Shaw-Smith, C.H. Cho, H. Lango Allen, J.A. Houghton, C.L. Roth, R. Chen, K. Hussain, P. Marsh, L. Vallier, A. Murray, C. International Pancreatic Agenesis, S. Ellard, J. Ferrer, and A.T. Hattersley, *Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis.* Nat Genet, 2014. **46**(1): p. 61-4.
 - 22. Monti, S., A. Nicoletti, A. Cantasano, H. Krude, and A. Cassio, *NKX2.1-Related Disorders: a novel mutation with mild clinical presentation.* Ital J Pediatr, 2015. **41**: p. 45.
 - 23. De Felice, M. and R. Di Lauro, *Thyroid development and its disorders: genetics and molecular mechanisms.* Endocr Rev, 2004. **25**(5): p. 722-46.
 - 24. Hamdan, H., H. Liu, C. Li, C. Jones, M. Lee, R. deLemos, and P. Minoo, *Structure of the human Nkx2.1 gene.* Biochim Biophys Acta, 1998. **1396**(3): p. 336-48.
 - 25. Santagati, F., K. Abe, V. Schmidt, T. Schmitt-John, M. Suzuki, K. Yamamura, and K. Imai, *Identification of Cis-regulatory elements in the mouse Pax9/Nkx2-9 genomic region: implication for evolutionary conserved synteny.* Genetics, 2003. **165**(1): p. 235-42.
 - 26. Minocha, S., D. Valloton, Y. Arsenijevic, J.R. Cardinaux, R. Guidi, J.P. Hornung, and C. Lebrand, *Nkx2.1 regulates the generation of telencephalic astrocytes during embryonic development.* Sci Rep, 2017. **7**: p. 43093.
 - 27. Krude, H., B. Schutz, H. Biebermann, A. von Moers, D. Schnabel, H. Neitzel, H. Tonnies, D. Weise, A. Lafferty, S. Schwarz, M. DeFelice, A. von Deimling, F. van Landeghem, R. DiLauro, and A. Gruters, *Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency.* J Clin Invest, 2002. **109**(4): p. 475-80.
 - 28. Fernandez, L.P., A. Lopez-Marquez, and P. Santisteban, *Thyroid transcription factors in development, differentiation and disease.* Nat Rev Endocrinol, 2015. **11**(1): p. 29-42.
 - 29. Invernizzi, F., G. Zorzi, A. Legati, G. Coppola, P. D'Adamo, N. Nardocci, B. Garavaglia, and D. Ghezzi, *Benign hereditary chorea and deletions outside NKX2-1: What's the role of MBIP?* Eur J Med Genet, 2018. **61**(10): p. 581-584.
 - 30. Kharbanda, M., P. Hermanns, J. Jones, J. Pohlenz, I. Horrocks, and M. Donaldson, *A further case of brain-lung-thyroid syndrome with deletion proximal to NKX2-1.* Eur J Med Genet, 2017. **60**(5): p. 257-260.
 - 31. Thorwarth, A., S. Schnittert-Hubener, P. Schruppf, I. Muller, S. Jyrch, C. Dame, H. Biebermann, G. Kleinau, J. Katchanov, M. Schuelke, G. Ebert, A. Steininger, C. Bonnemann, K. Brockmann, H.J. Christen, P. Crock, F. deZegher, M. Griese, J. Hewitt, S.

- Ivarsson, C. Hubner, K. Kapelari, B. Plecko, D. Rating, I. Stoeva, H.H. Ropers, A. Gruters, R. Ullmann, and H. Krude, *Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum*. *J Med Genet*, 2014. **51**(6): p. 375-87.
- 32. Barnett, C.P., J.J. Mencil, J. Gecz, W. Waters, S.M. Kirwin, K.M. Vinette, M. Uppill, and J. Nicholl, *Choreoathetosis, congenital hypothyroidism and neonatal respiratory distress syndrome with intact NKX2-1*. *Am J Med Genet A*, 2012. **158a**(12): p. 3168-73.
 - 33. Dale, R.C., P. Grattan-Smith, M. Nicholson, and G.B. Peters, *Microdeletions detected using chromosome microarray in children with suspected genetic movement disorders: a single-centre study*. *Dev Med Child Neurol*, 2012. **54**(7): p. 618-23.
 - 34. Spielmann, M. and S. Mundlos, *Looking beyond the genes: the role of non-coding variants in human disease*. *Hum Mol Genet*, 2016. **25**(R2): p. R157-r165.
 - 35. Dixon, J.R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J.S. Liu, and B. Ren, *Topological domains in mammalian genomes identified by analysis of chromatin interactions*. *Nature*, 2012. **485**(7398): p. 376-80.
 - 36. Sati, S. and G. Cavalli, *Chromosome conformation capture technologies and their impact in understanding genome function*. *Chromosoma*, 2017. **126**(1): p. 33-44.
 - 37. Weinreb, C. and B.J. Raphael, *Identification of hierarchical chromatin domains*. *Bioinformatics*, 2015.
 - 38. Stavreva, D.A. and G.L. Hager, *Chromatin structure and gene regulation: a dynamic view of enhancer function*. *Nucleus*, 2015. **6**(6): p. 442-8.
 - 39. Lupiáñez, D.G., K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki, D. Horn, H. Kayserili, J.M. Opitz, R. Laxova, F. Santos-Simarro, B. Gilbert-Dussardier, L. Wittler, M. Borschiwer, S.A. Haas, M. Osterwalder, M. Franke, B. Timmermann, J. Hecht, M. Spielmann, A. Visel, and S. Mundlos, *Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions*. *Cell*, 2015. **161**(5): p. 1012-1025.
 - 40. Lettice, L.A., S. Daniels, E. Sweeney, S. Venkataraman, P.S. Devenney, P. Gautier, H. Morrison, J. Fantes, R.E. Hill, and D.R. FitzPatrick, *Enhancer-adoption as a mechanism of human developmental disease*. *Hum Mutat*, 2011. **32**(12): p. 1492-9.
 - 41. Gudmundsson, J., P. Sulem, D.F. Gudbjartsson, J.G. Jonasson, G. Masson, H. He, A. Jonasdottir, A. Sigurdsson, S.N. Stacey, H. Johannsdottir, H.T. Helgadottir, W. Li, R. Nagy, M.D. Ringel, R.T. Kloos, M.C. de Visser, T.S. Plantinga, M. den Heijer, E. Aguillo, A. Panadero, E. Prats, A. Garcia-Castano, A. De Juan, F. Rivera, G.B. Walters, H. Bjarnason, L. Tryggvadottir, G.I. Eyjolfsson, U.S. Bjornsdottir, H. Holm, I. Olafsson, K. Kristjansson, H. Kristvinsson, O.T. Magnusson, G. Thorleifsson, J.R. Gulcher, A. Kong, L.A. Kiemeny, T. Jonsson, H. Hjartarson, J.I. Mayordomo, R.T. Netea-Maier, A. de la Chapelle, J. Hrafinkelsson, U. Thorsteinsdottir, T. Rafnar, and K. Stefansson, *Discovery of common variants associated with low TSH levels and thyroid cancer risk*. *Nat Genet*, 2012. **44**(3): p. 319-22.
 - 42. Taylor, P.N., E. Porcu, S. Chew, P.J. Campbell, M. Traglia, S.J. Brown, B.H. Mullin, H.A. Shihab, J. Min, K. Walter, Y. Memari, J. Huang, M.R. Barnes, J.P. Beilby, P. Charoen, P. Danecek, F. Dudbridge, V. Forgetta, C. Greenwood, E. Grundberg, A.D. Johnson, J. Hui, E.M. Lim, S. McCarthy, D. Muddyman, V. Panicker, J.R. Perry, J.T. Bell, W. Yuan, C. Relton, T. Gaunt, D. Schlessinger, G. Abecasis, F. Cucca, G.L. Surdulescu, W. Woltersdorf, E. Zeggini, H.F. Zheng, D. Toniolo, C.M. Dayan, S. Naitza, J.P. Walsh, T. Spector, G. Davey Smith, R. Durbin, J.B. Richards, S. Sanna, N. Soranzo, N.J. Timpson, and S.G. Wilson,

Whole-genome sequence-based analysis of thyroid function. Nat Commun, 2015. **6**: p. 5681.

- 43. Brookes, A.J., *The essence of SNPs*. Gene, 1999. **234**(2): p. 177-86.
- 44. Hayes, B., *Overview of Statistical Methods for Genome-Wide Association Studies (GWAS)*. Methods Mol Biol, 2013. **1019**: p. 149-69.
- 45. Smemo, S., J.J. Tena, K.H. Kim, E.R. Gamazon, N.J. Sakabe, C. Gomez-Marin, I. Aneas, F.L. Credidio, D.R. Sobreira, N.F. Wasserman, J.H. Lee, V. Puvindran, D. Tam, M. Shen, J.E. Son, N.A. Vakili, H.K. Sung, S. Naranjo, R.D. Acemel, M. Manzanares, A. Nagy, N.J. Cox, C.C. Hui, J.L. Gomez-Skarmeta, and M.A. Nobrega, *Obesity-associated variants within FTO form long-range functional connections with IRX3*. Nature, 2014. **507**(7492): p. 371-5.
- 46. Pennacchio, L.A., N. Ahituv, A.M. Moses, S. Prabhakar, M.A. Nobrega, M. Shoukry, S. Minovitsky, I. Dubchak, A. Holt, K.D. Lewis, I. Plajzer-Frick, J. Akiyama, S. De Val, V. Afzal, B.L. Black, O. Couronne, M.B. Eisen, A. Visel, and E.M. Rubin, *In vivo enhancer analysis of human conserved non-coding sequences*. Nature, 2006. **444**(7118): p. 499-502.
- 47. Ehrlich, K.C., H.L. Paterson, M. Lacey, and M. Ehrlich, *DNA Hypomethylation in Intragenic and Intergenic Enhancer Chromatin of Muscle-Specific Genes Usually Correlates with their Expression*. Yale J Biol Med, 2016. **89**(4): p. 441-455.
- 48. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions*. Nat Rev Genet, 2014. **15**(4): p. 272-86.
- 49. Elkon, R. and R. Agami, *Characterization of noncoding regulatory DNA in the human genome*. Nat Biotechnol, 2017. **35**(8): p. 732-746.
- 50. Kim, M. and J. Costello, *DNA methylation: an epigenetic mark of cellular memory*. Exp Mol Med, 2017. **49**(4): p. e322.
- 51. Razin, A. and B. Kantor, *DNA methylation in epigenetic control of gene expression*. Prog Mol Subcell Biol, 2005. **38**: p. 151-67.
- 52. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control*. Nat Rev Genet, 2016. **17**(8): p. 487-500.
- 53. Kouzarides, T., *Chromatin modifications and their function*. Cell, 2007. **128**(4): p. 693-705.
- 54. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications*. Cell Res, 2011. **21**(3): p. 381-95.
- 55. Vakoc, C.R., S.A. Mandat, B.A. Olenchock, and G.A. Blobel, *Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin*. Mol Cell, 2005. **19**(3): p. 381-91.
- 56. Bartke, T., M. Vermeulen, B. Xhemalce, S.C. Robson, M. Mann, and T. Kouzarides, *Nucleosome-interacting proteins regulated by DNA and histone methylation*. Cell, 2010. **143**(3): p. 470-84.
- 57. Reik, W., W. Dean, and J. Walter, *Epigenetic reprogramming in mammalian development*. Science, 2001. **293**(5532): p. 1089-93.
- 58. Brena, R.M., T.H. Huang, and C. Plass, *Toward a human epigenome*. Nat Genet, 2006. **38**(12): p. 1359-60.
- 59. Illingworth, R.S. and A.P. Bird, *CpG islands--'a rough guide'*. FEBS Lett, 2009. **583**(11): p. 1713-20.
- 60. Bird, A., *DNA methylation patterns and epigenetic memory*. Genes Dev, 2002. **16**(1): p. 6-21.

- 61. Luo, C., P. Hajkova, and J.R. Ecker, *Dynamic DNA methylation: In the right place at the right time*. Science, 2018. **361**(6409): p. 1336-1340.
- 62. Weber, M., I. Hellmann, M.B. Stadler, L. Ramos, S. Paabo, M. Rebhan, and D. Schubeler, *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome*. Nat Genet, 2007. **39**(4): p. 457-66.
- 63. Javierre, B.M., A.F. Fernandez, J. Richter, F. Al-Shahrour, J.I. Martin-Subero, J. Rodriguez-Ubreva, M. Berdasco, M.F. Fraga, T.P. O'Hanlon, L.G. Rider, F.V. Jacinto, F.J. Lopez-Longo, J. Dopazo, M. Forn, M.A. Peinado, L. Carreno, A.H. Sawalha, J.B. Harley, R. Siebert, M. Esteller, F.W. Miller, and E. Ballestar, *Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus*. Genome Res, 2010. **20**(2): p. 170-9.
- 64. Zhu, H., G. Wang, and J. Qian, *Transcription factors as readers and effectors of DNA methylation*. Nat Rev Genet, 2016. **17**(9): p. 551-65.
- 65. Ferguson-Smith, A.C., *Genomic imprinting: the emergence of an epigenetic paradigm*. Nat Rev Genet, 2011. **12**(8): p. 565-75.
- 66. Bell, A.C. and G. Felsenfeld, *Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene*. Nature, 2000. **405**(6785): p. 482-5.
- 67. Rakyan, V.K., M.E. Blewitt, R. Druker, J.I. Preis, and E. Whitelaw, *Metastable epialleles in mammals*. Trends Genet, 2002. **18**(7): p. 348-51.
- 68. Morgan, H.D., H.G. Sutherland, D.I. Martin, and E. Whitelaw, *Epigenetic inheritance at the agouti locus in the mouse*. Nat Genet, 1999. **23**(3): p. 314-8.
- 69. Kuhnen, P., D. Handke, R.A. Waterland, B.J. Hennig, M. Silver, A.J. Fulford, P. Dominguez-Salas, S.E. Moore, A.M. Prentice, J. Spranger, A. Hinney, J. Hebebrand, F.L. Heppner, L. Walzer, C. Grotzinger, J. Gromoll, S. Wiegand, A. Gruters, and H. Krude, *Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity*. Cell Metab, 2016. **24**(3): p. 502-9.
- 70. Castillo-Fernandez, J.E., T.D. Spector, and J.T. Bell, *Epigenetics of discordant monozygotic twins: implications for disease*. Genome Med, 2014. **6**(7): p. 60.
- 71. Kaminsky, Z.A., T. Tang, S.C. Wang, C. Ptak, G.H. Oh, A.H. Wong, L.A. Feldcamp, C. Virtanen, J. Halfvarson, C. Tysk, A.F. McRae, P.M. Visscher, G.W. Montgomery, Gottesman, II, N.G. Martin, and A. Petronis, *DNA methylation profiles in monozygotic and dizygotic twins*. Nat Genet, 2009. **41**(2): p. 240-5.
- 72. Rakyan, V.K., H. Beyan, T.A. Down, M.I. Hawa, S. Maslau, D. Aden, A. Daunay, F. Busato, C.A. Mein, B. Manfras, K.R. Dias, C.G. Bell, J. Tost, B.O. Boehm, S. Beck, and R.D. Leslie, *Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis*. PLoS Genet, 2011. **7**(9): p. e1002300.
- 73. Ishmael, F.T. and C. Stellato, *Principles and applications of polymerase chain reaction: basic science for the practicing physician*. Ann Allergy Asthma Immunol, 2008. **101**(4): p. 437-43.
- 74. Garibyan, L. and N. Avashia, *Polymerase chain reaction*. J Invest Dermatol, 2013. **133**(3): p. e6.
- 75. Erlich, H.A., *Polymerase chain reaction*. J Clin Immunol, 1989. **9**(6): p. 437-47.
- 76. Knebelsberger, T. and I. Stoger, *DNA extraction, preservation, and amplification*. Methods Mol Biol, 2012. **858**: p. 311-38.
- 77. Lee, P.Y., J. Costumbrado, C.Y. Hsu, and Y.H. Kim, *Agarose gel electrophoresis for the separation of DNA fragments*. J Vis Exp, 2012(62).

- 78. Prober, J.M., G.L. Trainor, R.J. Dam, F.W. Hobbs, C.W. Robertson, R.J. Zagursky, A.J. Cocuzza, M.A. Jensen, and K. Baumeister, *A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides*. *Science*, 1987. **238**(4825): p. 336-41.
- 79. Heather, J.M. and B. Chain, *The sequence of sequencers: The history of sequencing DNA*. *Genomics*, 2016. **107**(1): p. 1-8.
- 80. Ehrich, M., S. Zoll, S. Sur, and D. van den Boom, *A new method for accurate assessment of DNA quality after bisulfite treatment*. *Nucleic Acids Res*, 2007. **35**(5): p. e29.
- 81. Clark, S.J., J. Harrison, C.L. Paul, and M. Frommer, *High sensitivity mapping of methylated cytosines*. *Nucleic Acids Res*, 1994. **22**(15): p. 2990-7.
- 82. Grigg, G. and S. Clark, *Sequencing 5-methylcytosine residues in genomic DNA*. *Bioessays*, 1994. **16**(6): p. 431-6.
- 83. Clark, S.J., A. Statham, C. Stirzaker, P.L. Molloy, and M. Frommer, *DNA methylation: bisulphite modification and analysis*. *Nat Protoc*, 2006. **1**(5): p. 2353-64.
- 84. Grigg, G.W., *Sequencing 5-methylcytosine residues by the bisulphite method*. *DNA Sequence*, 2009. **6**(4): p. 189-198.
- 85. Hajj, N.E., J. Kultz, and T. Haaf, *Limiting Dilution Bisulfite Pyrosequencing(R): A Method for Methylation Analysis of Individual DNA Molecules in a Single or a Few Cells*. *Methods Mol Biol*, 2015. **1315**: p. 221-39.
- 86. Rakyan, V.K., T.A. Down, D.J. Balding, and S. Beck, *Epigenome-wide association studies for common human diseases*. *Nat Rev Genet*, 2011. **12**(8): p. 529-41.
- 87. Ronaghi, M., M. Uhlen, and P. Nyren, *A sequencing method based on real-time pyrophosphate*. *Science*, 1998. **281**(5375): p. 363, 365.
- 88. Ronaghi, M., S. Shokralla, and B. Gharizadeh, *Pyrosequencing for discovery and analysis of DNA sequence variations*. *Pharmacogenomics*, 2007. **8**(10): p. 1437-41.
- 89. Qiagen. *Pyrosequencing — the unique detection and quantification technology*. 2013 [cited 2020 22/05/2020]; Available from: <https://www.qiagen.com/be/resources/technologies/pyrosequencing-resource-center/technology-overview/>.
- 90. Haraksingh, R.R. and M.P. Snyder, *Impacts of variation in the human genome on gene regulation*. *J Mol Biol*, 2013. **425**(21): p. 3970-7.
- 91. Spielmann, M. and E. Klopocki, *CNVs of noncoding cis-regulatory elements in human disease*. *Curr Opin Genet Dev*, 2013. **23**(3): p. 249-56.
- 92. Safe, S., *Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions*. *Vitam Horm*, 2001. **62**: p. 231-52.
- 93. Swinstead, E.E., V. Paakinaho, D.M. Presman, and G.L. Hager, *Pioneer factors and ATP-dependent chromatin remodeling factors interact dynamically: A new perspective: Multiple transcription factors can effect chromatin pioneer functions through dynamic interactions with ATP-dependent chromatin remodeling factors*. *Bioessays*, 2016. **38**(11): p. 1150-1157.
- 94. Nonaka, D., *A study of FoxA1 expression in thyroid tumors*. *Hum Pathol*, 2017. **65**: p. 217-224.
- 95. Jimenez-Panizo, A., P. Perez, A. Rojas, P. Fuentes-Prior, and E. Estebanez-Perpina, *Non-canonical dimerization of the androgen receptor and other nuclear receptors: implications for human disease*. *Endocr Relat Cancer*, 2019.
- 96. Blinka, S. and S. Rao, *Nanog Expression in Embryonic Stem Cells - An Ideal Model System to Dissect Enhancer Function*. *Bioessays*, 2017. **39**(12).

- 97. Mashayekhi, F. and M. Ronaghi, *Analysis of read length limiting factors in Pyrosequencing chemistry*. Anal Biochem, 2007. **363**(2): p. 275-87.
- 98. Heijmans, B.T. and J. Mill, *Commentary: The seven plagues of epigenetic epidemiology*. Int J Epidemiol, 2012. **41**(1): p. 74-8.
- 99. Hunt, R., Z.E. Sauna, S.V. Ambudkar, M.M. Gottesman, and C. Kimchi-Sarfaty, *Silent (synonymous) SNPs: should we care about them?* Methods Mol Biol, 2009. **578**: p. 23-39.
- 100. Sifrim, A., M.P. Hitz, A. Wilsdon, J. Breckpot, S.H. Turki, B. Thienpont, J. McRae, T.W. Fitzgerald, T. Singh, G.J. Swaminathan, E. Prigmore, D. Rajan, H. Abdul-Khaliq, S. Banka, U.M. Bauer, J. Bentham, F. Berger, S. Bhattacharya, F. Bu'Lock, N. Canham, I.G. Colgiu, C. Cosgrove, H. Cox, I. Daehnert, A. Daly, J. Danesh, A. Fryer, M. Gewillig, E. Hobson, K. Hoff, T. Homfray, A.K. Kahlert, A. Ketley, H.H. Kramer, K. Lachlan, A.K. Lampe, J.J. Louw, A.K. Manickara, D. Manase, K.P. McCarthy, K. Metcalfe, C. Moore, R. Newbury-Ecob, S.O. Omer, W.H. Ouwehand, S.M. Park, M.J. Parker, T. Pickardt, M.O. Pollard, L. Robert, D.J. Roberts, J. Sambrook, K. Setchfield, B. Stiller, C. Thornborough, O. Toka, H. Watkins, D. Williams, M. Wright, S. Mital, P.E. Daubeney, B. Keavney, J. Goodship, R.M. Abu-Sulaiman, S. Klaassen, C.F. Wright, H.V. Firth, J.C. Barrett, K. Devriendt, D.R. FitzPatrick, J.D. Brook, and M.E. Hurler, *Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing*. Nat Genet, 2016. **48**(9): p. 1060-5.
- 101. Slatkin, M., *Linkage disequilibrium--understanding the evolutionary past and mapping the medical future*. Nat Rev Genet, 2008. **9**(6): p. 477-85.
- 102. Shoemaker, R., W. Wang, and K. Zhang, *Mediators and dynamics of DNA methylation*. Wiley Interdiscip Rev Syst Biol Med, 2011. **3**(3): p. 281-98.
- 103. Wang, X., T. Zhou, Z. Wunderlich, M.T. Maurano, A.H. DePace, S.V. Nuzhdin, and R. Rohs, *Analysis of Genetic Variation Indicates DNA Shape Involvement in Purifying Selection*. Mol Biol Evol, 2018. **35**(8): p. 1958-1967.
- 104. Hamdi, Y., M. Leclerc, M. Dumont, S. Dubois, M. Tranchant, G. Reimnitz, P. Soucy, P. Cassart, M. Ouimet, D. Sinnett, M.L.L. Chaieb, and J. Simard, *Functional Analysis of Promoter Variants in Genes Involved in Sex Steroid Action, DNA Repair and Cell Cycle Control*. Genes (Basel), 2019. **10**(3).
- 105. Goldstein, I. and G.L. Hager, *Dynamic enhancer function in the chromatin context*. Wiley Interdiscip Rev Syst Biol Med, 2017.
- 106. Allen, E.K., A.G. Randolph, T. Bhangale, P. Dogra, M. Ohlson, C.M. Oshansky, A.E. Zamora, J.P. Shannon, D. Finkelstein, A. Dressen, J. DeVincenzo, M. Caniza, B. Youngblood, C.M. Rosenberger, and P.G. Thomas, *SNP-mediated disruption of CTCF binding at the IFITM3 promoter is associated with risk of severe influenza in humans*. Nat Med, 2017. **23**(8): p. 975-983.
- 107. Nair, A.K. and L.J. Baier, *Using Luciferase Reporter Assays to Identify Functional Variants at Disease-Associated Loci*. Methods Mol Biol, 2018. **1706**: p. 303-319.
- 108. Promega, *Biolumineszente Reporterassays*. 2017.
- 109. Remes-Lenicov, F., K.X. Jacobsen, A. Rogaeva, M. Czesak, M. Hadjighasem, M. Daigle, and P.R. Albert, *Frontiers in Neuroscience*
- *Identification of Novel Transcriptional Regulators in the Nervous System, in Serotonin Receptors in Neurobiology*, A. Chattopadhyay, Editor. 2007, CRC Press/Taylor & Francis Taylor & Francis Group, LLC.: Boca Raton (FL).
- 110. Portela, A. and M. Esteller, *Epigenetic modifications and human disease*. Nat Biotechnol, 2010. **28**(10): p. 1057-68.

- 111. Do, C., C.F. Lang, J. Lin, H. Darbary, I. Krupska, A. Gaba, L. Petukhova, J.P. Vonsattel, M.P. Gallagher, R.S. Goland, R.A. Clynes, A. Dwork, J.G. Kral, C. Monk, A.M. Christiano, and B. Tycko, *Mechanisms and Disease Associations of Haplotype-Dependent Allele-Specific DNA Methylation*. *Am J Hum Genet*, 2016. **98**(5): p. 934-955.
- 112. Brugha, R., R. Lowe, A.J. Henderson, J.W. Holloway, V. Rakyan, E. Wozniak, N. Mahmud, K. Seymour, J. Grigg, and S.O. Shaheen, *DNA methylation profiles between airway epithelium and proxy tissues in children*. *Acta Paediatr*, 2017. **106**(12): p. 2011-2016.
- 113. Costello, J.F. and C. Plass, *Methylation matters*. *J Med Genet*, 2001. **38**(5): p. 285-303.
- 114. Banovich, N.E., Y.I. Li, A. Raj, M.C. Ward, P. Greenside, D. Calderon, P.Y. Tung, J.E. Burnett, M. Myrthil, S.M. Thomas, C.K. Burrows, I.G. Romero, B.J. Pavlovic, A. Kundaje, J.K. Pritchard, and Y. Gilad, *Impact of regulatory variation across human iPSCs and differentiated cells*. *Genome Res*, 2018. **28**(1): p. 122-131.
- 115. Ji, H., L.I. Ehrlich, J. Seita, P. Murakami, A. Doi, P. Lindau, H. Lee, M.J. Aryee, R.A. Irizarry, K. Kim, D.J. Rossi, M.A. Inlay, T. Serwold, H. Karsunky, L. Ho, G.Q. Daley, I.L. Weissman, and A.P. Feinberg, *Comprehensive methylome map of lineage commitment from haematopoietic progenitors*. *Nature*, 2010. **467**(7313): p. 338-42.
- 116. Houseman, E.A., K.T. Kelsey, J.K. Wiencke, and C.J. Marsit, *Cell-composition effects in the analysis of DNA methylation array data: a mathematical perspective*. *BMC Bioinformatics*, 2015. **16**: p. 95.
- 117. Joubert, B.R., S.E. Haberg, R.M. Nilsen, X. Wang, S.E. Vollset, S.K. Murphy, Z. Huang, C. Hoyo, O. Midttun, L.A. Cupul-Uicab, P.M. Ueland, M.C. Wu, W. Nystad, D.A. Bell, S.D. Peddada, and S.J. London, *450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy*. *Environ Health Perspect*, 2012. **120**(10): p. 1425-31.
- 118. Futscher, B.W., M.M. Oshiro, R.J. Wozniak, N. Holtan, C.L. Hanigan, H. Duan, and F.E. Domann, *Role for DNA methylation in the control of cell type specific maspin expression*. *Nat Genet*, 2002. **31**(2): p. 175-9.
- 119. Jones, M.J., S.J. Goodman, and M.S. Kobor, *DNA methylation and healthy human aging*. *Aging Cell*, 2015. **14**(6): p. 924-32.
- 120. Fraga, M.F., E. Ballestar, M.F. Paz, S. Ropero, F. Setien, M.L. Ballestar, D. Heine-Suñer, J.C. Cigudosa, M. Urioste, J. Benitez, M. Boix-Chornet, A. Sanchez-Aguilera, C. Ling, E. Carlsson, P. Poulsen, A. Vaag, Z. Stephan, T.D. Spector, Y.Z. Wu, C. Plass, and M. Esteller, *Epigenetic differences arise during the lifetime of monozygotic twins*. *Proc Natl Acad Sci U S A*, 2005. **102**(30): p. 10604-9.
- 121. Izzi, B., M. Pistoni, K. Cludts, P. Akkor, D. Lambrechts, C. Verfaillie, P. Verhamme, K. Freson, and M.F. Hoylaerts, *Allele-specific DNA methylation reinforces PEAR1 enhancer activity*. *Blood*, 2016. **128**(7): p. 1003-12.
- 122. Narumi, S., K. Matsubara, T. Ishii, and T. Hasegawa, *Methylome analysis of thyroid ectopy shows no disease-specific DNA methylation signature*. *Clin Pediatr Endocrinol*, 2018. **27**(4): p. 235-238.

6 Eidesstattliche Versicherung

„Ich, Sarah Assadeh Flemmerer, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Genetische und epigenetische Analyse einer konservierten nicht-kodierenden Region downstream des *NKX2-1* Genes in Patienten mit kongenitaler Hypothyreose“/ „Genetic and epigenetic analysis of a conserved non-coding region downstream of the *NKX2-1* gene in patients with congenital hypothyroidism“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

[Für den Fall, dass Sie die Forschung für Ihre Promotion ganz oder teilweise in Gruppenarbeit durchgeführt haben:] Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.org) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum

Unterschrift

Anteilserklärung an etwaigen erfolgten Publikationen

[Sollten bereits Teile aus Ihrer Monographie publiziert worden sein, dann müssen Sie dies im Vorwort nach dem Deckblatt erklären und diese Anteilserklärung ausfüllen.

Die Anteile an den etwaigen Publikationen sind so deutlich und detailliert zu erklären, dass es der Promotionskommission und den wissenschaftlichen Gutachtern ohne Probleme möglich ist zu erkennen, was Sie selbst dazu beigetragen haben. Wünschenswert wäre ein konkreter Bezug zur Publikation wie z. B.: „aus meiner statistischen Auswertung sind die Tabellen 1, 4, 47 und 60 entstanden.“]

Sarah Assadeh Flemmerer hatte folgenden Anteil an den folgenden Publikationen:

- Keine -

Unterschrift, Datum und Stempel des/der erstbetreuenden Hochschullehrers/in

Unterschrift des Doktoranden/der Doktorandin

7 Curriculum vitae

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

8 Danksagung

Ich möchte mich bei meiner Familie, und hierbei vor allem bei Klaus Louis, Anne Christine Holtmann und meinem Mann Fabian Louis für ihre Unterstützung bedanken.

Weiter möchte ich mich bei dem gesamten Team des Instituts für experimentelle pädiatrische Endokrinologie, und hierbei insbesondere bei Rita Oeltjen bedanken.

Ein besonderer Dank gilt Prof. Heiko Krude, PD Peter Kühnen und Dr. med. Anne Thorwarth, die mir einerseits die Möglichkeit zu dieser spannenden und bereichernden Arbeit eröffnet und mich andererseits stets bei der Umsetzung unterstützt haben.